Essays in Transport Economics - Challenges in Modeling Travel Mode Choice and User Benefits

Artikler i transportøkonomi -
Utfordringer i modellering av transportmiddelvalg og brukernytte

Philosophiae Doctor (PhD) Thesis

Stefan Flügel

School of Economics and Business
Faculty of Social Science
Norwegian University of Life Sciences

Ås 2015

Thesis number 2015:61
ISSN: 1894-6402

ISBN: 978-82-575-1300-9

Acknowledgment

Working with and writing this PhD thesis has been a great experience that has enriched me, certainly as a researcher but also as a person. I am very grateful for all the support and contributions I got throughout my PhD.

In particular, I want to thank my supervisor Ståle Navrud. Ståle was a great support for me from the first day (giving me a guided tour through the University campus) to the last day of my PhD (commenting on the introduction part). I am very grateful for his encouragements, scientific contributions and practical support. I also want to thank my co-superiors Olvar Bergland and Lasse Fridstrøm for their support and helpful comments. I am especially grateful to Lasse for initiating my PhD and integrating it in the TEMPO project. This project financed the high-speed data collection and organized many interesting workshops and conferences that put me in contact with several international experts.

This thesis profits greatly from the contributions of my co-authors: Askill H. Halse, Ståle Navrud, Juan de Dios Ortúzar, Luis Rizzi, Farideh Ramjerdi, Knut Veisten, Marit Killi and Rune Elvik. Thank you all for your significant contributions, the inspiring discussions and the good teamwork!

I would like to thank the School of Economics and Business for financing and administrating my PhD. Many thanks also go to my fellow PhD students for good discussions and social gatherings.

I also want to thank all my colleagues at the Institute of Transport Economics, TØI. I'm particular grateful to Farideh Ramjerdi, Knut Veisten and Harald Minken for introducing me to transport economics, stated preferences studies and discrete choice modeling and to Anne Madslien, Kjell Werner Johansen and Lasse Fridstrøm for their support and trust (giving me great freedom and flexibility in my research).

Furthermore, I want to thank Daniel Sperling, Pat Mokhtarian and my fellow students at UC Davis for an inspiring and enjoyable research visit.

Finally, I want to thank my wife Liv and my whole family in Germany and Norway for their support and love. I became father to Klara and Mikkjel during my PhD, making this period the most memorable time of my life.

Ås, January 2015.

Stefan Flügel

Summary

The subject of this PhD thesis is transport economics. The thesis concerns the modeling and application of transport-related choice data, in particular data from choice experiments, and contributes on the field of specifying utility function in travel mode choice and estimating and applying willingness-to-pay (WTP) measures and user benefits for economic appraisal. The thesis consists of four self-contained essays and an introduction.

Essay 1 is about the curvature of marginal utility functions of Level-of-Service attributes in travel mode choice models. It presents the concept of self-selection to attribute values in travel mode choice models which is argued to be a potential explanation for counter-theoretical empirical results estimated on cross-sectional data. Analyzing stated choice data of high speed rail in Norway, we find some empirical support for that controlling for unobserved taste heterogeneity in estimation can retrieve curvatures for the travel costs attribute suggested from microeconomic theory.

Essay 2 uses the same choice data as Essay 1 and concerns utility specification as well. However, the topic here is the random part of utility, and in particular the correlation structure among the travel mode alternatives. The essay contributes to the current literature by identifying and discussing the limitations and caveats in deriving the error structure of the forecasting model from the estimation models based on binary stated choice data between travel's current mode and a new alternative (here: high-speed rail). The paper provides empirical illustrations of how information from revealed choice data among current modes, and advanced discrete choice models (cross-nested logit model with random coefficients) can be utilized. The essay provides strong arguments for constructing mode choice experiments with at least three travel alternatives.

In Essay 3 we analyze choices made by cyclists in different types of choice experiments and elicit their WTP for cycling facilities such as separated cycling path and reduction of crossings. The novel element of this paper is that we include a casualty risk attribute. With a pooled estimation model, we can then elicit how much of the user benefits of cycling facilities

are connected to casualty risk reduction. We find that WTP is close to halve when controlling for casualty risk. Recognizing this can avoid double-counting in economic appraisals.

Finally, Essay 4 provides a discussion about opportunities and challenges of including information about user type- and mode effects on between-mode differences in value of travel time savings in project appraisal. In this context, I argue that the proposed approach of "mode effect dependent equity value", which acknowledges mode effects due to comfort difference of travel modes but controls for user type effects due to self-selection, may help to provide optimal standards in economic appraisal. In a stylized case study and using Norwegian data, I illustrate how the ranking of projects can be affected by the choice of different approaches.

Sammendrag

Temaet for denne doktorgradsavhandlingen er transportøkonomi. Avhandlingen handler om modellering og anvendelse av transportrelaterte valgdata, spesielt data fra valgeksperimenter, og bidrar på feltet knytet til spesifisering av nyttefunksjon i transportmiddelvalgmodeller og estimering og anvendelse av betalingsvillighet og brukernytte i samfunnsøkonomisk analyser. Avhandlingen består av fire selvstendige artikler og en innledning.

Artikkel 1 handler om formen av den marginale nyttefunksjon til kostnads- og tidsattributter i transportmiddelvalgmodeller. Begrepet "selvseleksjon til attributtverdier" presenteres, og vi argumenterer dette for å være en potensiell forklaring på empiriske resultater estimert på tverrsnittsdata som strider mot mikroøkonomisk teori. I analysen av utalte valgdata (stated choice data) for høyhastighetstog i Norge finner vi noe empirisk støtte for at man får estimert den teoretisk forventede formen for kostnadsattributtet når man kontrollerer for uobservert heterogenitet i folks preferanser.

Artikkel 2 bruker de samme data som Artikkel 1 og handler også om spesifikasjon av den marginale nyttefunksjonen. Imidlertid er temaet her den tilfeldige delen av nyttefunksjonen, nærmere bestemt korrelasjonsstrukturen blant reisemiddelalternativer. Artikkelen diskuterer begrensninger og mulige feil ved å utlede korrelasjonensstruktur i en prognosemodell fra estimeringsmodeller basert på binære uttalte valg mellom den reisendes nåværende transportmiddel og et nytt alternativ (som her er høyhastighetstog). Artikkelen gir en empirisk illustrasjon av hvordan informasjon fra avslørte valgdata (Revealed Preference Data) blant dagens transportmidler og avanserte diskrete valgmodeller (kryss-nestede logitmodeller med tilfeldige koeffisienter) kan utnyttes. Artikkelen gir sterke argumenter for å konstruere eksperimenter av transportmiddelvalg med minst tre reisealternativer.

I artikkel 3 analyserer vi valgene som blir gjort av syklister i ulike typer valgeksperimenter, og beregne betalingsvilligheten til to sykkelfasiliteter: adskilte sykkelstier og reduksjon av kryssinger. Denne artikkelen tilføres ny kunnskap på dette feltet ved at vi inkluderer

ulykkesrisiko som attributt i et valgeksperiment for sykling. Med en samlet estimeringsmodell, kan vi dermed finne ut hvor mye av brukernytten til sykkelfasiliteter er knytet til reduksjon av ulykkesrisiko. Vi finner at verdiene er nær halvert når man kontrollerer for ulykkesrisiko. Dette kan være viktig informasjon for å unngå dobbelttelling i samfunnsøkonomiske analyser.

Artikkel 4 diskuterer muligheter og utfordringer ved å inkludere informasjon om brukertype- og transportmiddeleffekter i transportmiddel-spesifikke tidsverdier når man evaluere transportprosjekter. Jeg argumenterer her for at "transportmiddelspesifikke enhetsverdier", som erkjenner transportmiddeleffekter grunnet komfortforskjellen av reisemidler, men kontrollerer for brukertypeeffekter på grunn av selvseleksjon, kan bidra til å gi optimale standardverdier til bruk i samfunnsøkonomisk analyser. I en stilisert case studie og ved bruk av norske data illustrerer jeg hvordan rangeringen av prosjekter kan påvirkes av valg av ulike tilnærminger til transportmiddel-spesifikke tidsverdier.

Table of Content

1. Introduction

The introduction chapter is intended to give an overview over the thesis and to establish a thematic connection between the essays. It provides some background knowledge and describes briefly some common concepts, theories, data and models. By an upfront discussion of challenges in modelling travel mode choices and user benefits, the motivation for the topic of the thesis is presented. Finally, the introduction points to the particular contributions of the essays in relation to the existing literature.

1.1. Overall Topic

This thesis contains four essays in Transport Economics, where I employ economic principles and the theory of economic behaviour to study transportation processes. More specifically, the thesis considers: (i) modelling of utility functions underlying traveller's choice of transportation mode, and (ii) topics in the estimation and application of user benefits for economic appraisals. The former directly relates to the demand side of the travel market while the latter relates to information needed for resource allocation within the travel market.

In addition to the transport economics context, the essays also have in common that they depart from conceptual or methodological challenges related to discrete choice analysis; either in model specification or in applications. An important topic throughout the thesis is different types of self-selection of (heterogeneous) travellers to travel modes, and how this affects the results from choice experiments. In this connection, accounting for heterogeneity between and/or within user groups is a challenge that is approached by appropriate model formulations (in particular models that include randomly distributed coefficients) and/or and joint estimation of different data sources.

As my PhD project was part of the research project TEMPO (Fridstrøm and Alfsen 2014) on sustainable transportation, it should not come as a surprise that the first three essays are empirical analyses related to environmentally sustainable travel modes, i.e. high-speed rail in the first two essays and cycling in the third. Notwithstanding, the methodological challenges discussed in the thesis are by no means limited to environmentally sustainable travel modes. In fact, they seem - in their conceptual core - relevant not only to transportation but also to other research fields.

Despite the thematic overlaps the essays are self-contained, and can be read independently of each other.

Box 1 provides the most important expressions and their abbreviations.

ASC: alternative specific constant	MRS: marginal rate of substitution
BCT: Box-Cox transformations	NL: nested logit
CBA: cost-benefit analysis	OCT: opportunity cost of travelling
CE: choice experiments	OD: origin-destination
CNL: cross-nested logit	RC: revealed choice
GC: generalized costs	RP: revealed preference
GEV: generalized extreme value	RUM: random utility models
HL: heteroskedastic logit	SC: stated choice
HSR: high-speed rail	SP: stated preference
IIA: independence of irrelevant alternatives	SSAV: self-selection to attribute values
iid: independent and identical distributed	SSTM: self-selection to travel modes
LoS: level-of-service	TH: taste heterogeneity
ML: mixed logit	VTTS: value of travel time savings
MNL: multinomial logit model	WTP: willingness-to-pay

Box 1: Important abbreviations in introduction section

1.2 Transport Economics

Particularities of Transport Economics

Like other applied fields of economics, transport economics is concerned with the demand, supply and allocation of goods/products/resources in this specific sector. The product of transportation processes are displacements of individuals or things, i.e. movements of passengers and freight. Unlike traditional consumer products, the key dimensions of transportation are space and time. To the extent that standard economic theory (i.e. classical consumer and production theory) do not explicitly account for space and time, the theory of transport economics is filling a gap by making space and time explicit on the supply/production and demand/consumption sides.

Jara-Diaz (2007, p. 12) specifies the product vector of transportation processes as

(1) $$Y = \{y_{od}^{kt}\} \in R^{KxNxT},$$

where each component y_{ij}^{kt} represents the flow of type *k* (specifying - at least - what is transported, and by which transport mode) from origin *o* to destination *d* (OD pair *od*) within period *t* (*K*, *N* and *T* are the number of flow types, the number of OD pairs, and the number of time periods, respectively).

Regarding the demand side of person transport, which is the focus of this thesis, two specific elements of transport economics may be highlighted: i) there is no particular demand for transportation itself, rather a demand for different activities that are spatially separated; and hence the demand for travel is derived from the demand of activities, and ii) the consumption of transport process requires traveler's own time and (unlike for other products) time is the single most important element for transport products (Jara-Diaz, 2007, p. 7). This makes the economic theory of time allocation, initiated by Becker (1965) and further developed by e.g. Oort (1969), DeSerpa (1971), Small (1982) and Jara-Diaz and Guevara (2003) an important conceptual element of transport economics.

Time Allocation Models and the Value of Travel Time Savings

Time allocation models suggest that travel should be seen in relation to other activities and within an integrated framework of all-day time allocation. They also provide the foundations for economic valuation of time.

Consider the basic model by DeSerpa (1971); in an adjusted form focusing on travel application (Jara-Diaz and Guevara, 2003):

(2) $$\max \; U(\boldsymbol{X}, \boldsymbol{T}) \qquad \text{s.t.:}$$

(3) $$I_f + wT_w - P'X - c_R \geq 0 \;\rightarrow \lambda$$

(4) $$\tau - \sum_i T_i = 0 \;\rightarrow\; \mu$$

(5) $$T_j - T_j^{min} \geq 0 \;\rightarrow \mathrm{K}_j \qquad \forall j = R, W_f$$

(6) $$T_i - h_i(X) \geq 0 \;\rightarrow \mathrm{K}_i \qquad \forall i \neq R, W_f$$

A representative agent maximizes the utility derived from the consumption of goods (set ***X***) and time spend on activities (set ***T***) subject to several constraints: The income constraint (Eq. 3) imposes that fixed income (I_f) plus the wage eared (wage rate, *w*, times the allocated working time, T_w) must be greater or equal the expenditures for goods consumption ($P'X$) and travel activities (c_R). The time constraint (Eq. 4) states that the time spend in all activities equals the total time of the considered period (τ). In this model, there are additional time restrictions put on single activities depending on whether the length of the activity can be adjusted freely by the agent or not. For travel activities (T_R) and fixed working time (T_{W_f}) there exists an exogenous time limit (T_j^{min}) which the duration of the activity cannot fall below. For other activities like leisure activities, that limit is endogenous and is modeled as a function of consumption goods *X*. When consumer goods that facilitate the activity are no longer available or no longer affordable, the activity cannot be conducted for the desirable duration of time.

The Lagrange multipliers λ and μ represent the marginal utility of income and of (unspecified) time respectively. μ may be interpreted as the marginal utility one would get if the daily time budget of 24 hours would be (marginally) raised. The parameter ratio μ/λ is referred to as the *value of time as a resource* and reflects the opportunity cost of time. In the given model it can be shown to equal $w + \frac{\frac{\partial U}{\partial T_w}}{\lambda}$ (DeSerpa, 1971, Jara-Diaz 2007), that is the nominal wage plus the marginal (dis)utility of working. The opportunity cost of time, and in a transportation perspective, the opportunity costs of travelling (OCT), are usually lower than the wage rate as the marginal utility of time spend working ($\frac{\partial U}{\partial T_w}$) is in general assumed negative.

For leisure activities, K_i will be zero because an agent is (by definition of a leisure activity) putting more time than needed into a leisure activity (making constraint (5) non-binding).

The ratio K_j/λ is the value of saving time in an undesirable activity *j*. It can be shown to be the sum of the value of time as a resource and the value of assigning time, i.e. the scaled dis-utility of time spend in activity *j*.

(7) $$\frac{K_j}{\lambda} = \frac{\mu}{\lambda} - \frac{\frac{\partial U}{\partial T_j}}{\lambda} = w + \frac{\frac{\partial U}{\partial T_w}}{\lambda} - \frac{\frac{\partial U}{\partial T_j}}{\lambda}.$$

For travel activities, Eq.7 - which goes back to Oort (1969) - is a decomposition of the Value of Travel Time Saving (VTTS). In *Essay 4*, this decomposition is used as a starting point for an interpretation of user-type and mode effects of differences of the VTTS across various travel models.

Economic time allocation models provide a conceptual framework to understand travel demand and valuation of time. However, they are not widely used in practise to predict actual travel demand and traveller's valuation of time.

Utility of Discrete Travel Alternatives

In practise, travel demand modelling and valuation of time is to a large extent based on the analysis of choices made between discrete alternatives, as these model formulations are typically easier to specify and calibrate than the continuous and budget constrained maximisation problem in time allocation models. This makes discrete choice models a central tool in transportation analysis and planning.[1] After defining the relevant decision makers *n*, and choice sets, a crucial step is to specify the utility functions of each alternative *i* available in the choice set (representing for instance different travel routes or different travel modes). Utility functions are typically subdivided in a deterministic and a random part (see section 1.5 for underlying theories and details of utility specifications).

(8) $$Z_{ni} = V_{ni}(\boldsymbol{\theta_{ni}}, \boldsymbol{A_{ni}}) + \varepsilon_{ni}$$

[1] It is also interesting to note that leading researchers in the field of discrete choice modeling like Daniel McFadden, Moshe Ben-Akiva, David Hensher, Kenneth Train, Juan de Dios Ortúzar, Chandra Bhat, and more recently Joan Walker, Stephane Hess, Michel Bierlaire and Mogens Fosgerau all have their main applications in the field of transport.

The random (unsystematic) part, $\varepsilon_{ni},$ is undoubtedly important, especially in the prediction of market shares (see *Essay 2*). However, it is the deterministic (systematic) part of utility, $V_{ni},$ that contains the economic mechanism in travellers´ decision making There the attributes of alternatives $\boldsymbol{A_{ni}}$, such as travel time and travel cost, are specified together with the parameter $\boldsymbol{\theta_{ni}}$ representing the impact of these attributes on utility.

The deterministic utility function can be derived as the conditional indirect utility function from time allocation models, i.e. a function representing the utility level reached given the choice of an alternative *i* (e.g. Blayac and Causse 2001). Doing so, the functional form assumed for *U* in Eq. 3 has implication for the functional form of V_i.

Let *k* be the index for attributes in vector $\boldsymbol{A_{ni}}$. Then the marginal utility of specific attribute a_{nik} on utility V_{ni} is defined as $\frac{\partial V_{ni}}{\partial a_{nik}}$.

Inspired by a quote of Koppelman (1981, p.131), we argue in *Essay 1* that if standard economic assumptions are imposed for the arguments in the utility function in time allocation models (desirability and convexity assumptions, see e.g. Mas-Colell *et al* (1995), the marginal utility associated with travel time and travel cost (and other Level-of-Service attributes that are reducing travellers income or leisure time budget) should be negative and increasing (or constant), i.e. $\frac{\partial V_{ni}}{\partial a_{nik}} < 0$ and $\frac{\partial^2 V_{ni}}{\partial {a_{nik}}^2} \geq 0.$

Marginal Rate of Substitution and Willingness-To-Pay

The concept of decomposing utility into attributes (prices, travel time and qualitative attributes) is an important concept in transport modelling. This idea from consumer theory goes back to Griliches (1961) and Lancaster (1966). It implies that utility is compensatory, i.e. that a traveller can be offset for a worsening of one attribute with an improvement in another attribute.

The marginal rate of substitution (MRS) is a measure for this trade-off. It is calculated as the ratio of marginal utility values of two attributes:

(9) $$MRS_{ni,k=1,k=2}(a_{ni1}, a_{ni2}) = \frac{\partial V_{ni}(a_{ni1}, a_{ni2})}{\partial a_{ni1}} / \frac{\partial V_{ni}(a_{ni1}, a_{ni2})}{\partial a_{ni2}}$$

A nice feature of the MRS is that the overall scale of utility (the absolute numerical utility value), which is arbitrary in the ordinal utility concept of economics, cancels out. Therefore the MRS has a direct interpretation in that it represents the trade-off rate two attributes can be substituted with each other, such that the utility of traveller *n* is kept constant.

In case k=2 is the monetary cost of the alternative (c_i) and in case k=1 is an economic bad $(\frac{\partial V_{ni}}{\partial a_{ni1}} < 0)$, Eq. 9 is readily interpreted as the amount of money a traveller *n* is willingness-to-pay (WTP) to for a marginal reduction in the attribute.[2] If that attribute is travel time, the MRS represents the Value of Travel Time Saving (VTTS). Note, that the VTTS (or other WTP) calculated by Eq. 9 may depend on traveller *n*, alternative *i,* and on the absolute attribute value of travel cost and time (other attributes).

MRS can also be calculated for more qualitative characteristics of the alternatives.[3] In *Essay 3,* for instance, we calculate trade-offs between cycling time and the share of a cycling route being separated from car traffic. The MRS is then interpreted as the additional minutes a traveller is willing to cycle, in order to increase the share of cycling route being on a separated cycling path by one percent point.

In some instances it may be meaningful to normalize the marginal utility of travel cost to the value one. Then the utility function may be written in "WTP-space" in which case it represents the generalized costs of a travel alternative (GC_{ni}).

User Benefits and Social Welfare

For an efficient resource allocation, e.g. the ranking of transport projects, it is important to be able to compare the welfare change for society associated with different improvements (or worsening) in the transportation sector. A crucial step in this quest is the identification and quantification of user benefits.

---

[2] In correspondence with classical economic theory, the essays in this thesis do not make a difference between WTP and willingness-to-accept (WTA).
[3] For purely nominal attributes, binary dummy variables (0/1) are usually applied. The marginal change is then defined as a change of the dummy value from 1 to 0.

There are different interpretations and measures of user benefits.[4] I concentrate here on the compensating variation concept (Hicks 1956) which I also adopt in *Essay 4.* It is originally expressed for changes in price, but can also be applied to changes in product quality, e.g. in terms of travel time (Jara-Diaz 2007, p.98). The idea is to find the change in income needed to exactly offset the changes in price and/or travel time.

Building on a general expression of the compensating variation (CV) by Small and Rosen (1981), Jara-Diaz (1990) derives a compact expression of CV as an (aggregate) measure of user benefits (*B*). With the notation used above, it may be written as:

(10) $$B = CV = \frac{1}{\lambda} \sum_i \Delta V_i \bar{D}_i$$

where $\bar{D}_i$ is the expected numbers of users of alternative *i* and ΔV_i is the local variation in (indirect) utility expressed in local variation of monetary cost (c_i) and other attributes (a_{ik})

(11) $$\Delta V_i = \frac{\partial V_i}{\partial c_i} \Delta c_i + \sum_k \frac{\partial V_i}{\partial a_{ik}} \Delta a_{ik}$$

Given that the marginal utility of travel cost equals the marginal utility of income (with reversed sign), i.e. $\frac{\partial V_i}{\partial c_i} = -\lambda$, we can reformulate Eq. 9 to (compare Jara-Diaz 1990):

(12) $$B = CV = -\sum_i \Delta c_i \bar{D}_i + \sum_k WTP_k \sum_i \bar{D}_i \Delta a_{ik}$$

with $WTP_k = -\frac{\partial V_i}{\partial a_{ik}} / \frac{\partial V_i}{\partial c_i}$.

When a project affects only the travel time of one travel alternative, Eq. 12 can be written as:

(13) $$CV_i^{\Delta t} = \bar{D}_i VTTS_i \Delta t_i.$$

Note that Eq. 13 assumes a common VTTS for all users. This may be relaxed by letting VTTS vary over users, in which case aggregated user benefits can be written as:

---

[4] Probably the simplest and most widely applied method is the rule-of-half, where the user benefits from changes in the generalized cost (of one OD pair and one mode) is approximates as: $\Delta B \approx \frac{1}{2} * (M^0 + M^1) * (GC^0 + GC^1)$, with M^0, M^1 are the total amount of trips (for that mode between the OD) before and after the change. The rule-of-half is an approximation of the Marshallian consumer surplus (Marshall 1920, Hotelling 1938) that is - according to Jara-Diaz (2007, p. 87) - a quite arbitrary measure of user benefits.

(14) $$B_i^{\Delta t} = CV_i^{\Delta t} = \sum_n VTTS_{ni}\Delta t_{ni}.$$

From a social welfare perspective however, not only user benefits matter. It matters also how user benefits affect the utility of each individual and how society weights the utility of each member.

From a general formulation of the social welfare function,

(15) $$W_s = W_s(U_1, \dots, U_n, \dots, U_N),$$

where individual utility is a function of goods X_n, which again is a function of generalized income I_n and prices P

(16) $$U_n(X_n) = U_n(X_n(I_n, P)),$$

Gálvez and Jara-Díaz (1998) state that a welfare change resulting from a change in user benefits (dB_q) can be expressed as:

(17) $$dW_s = \sum_q \Omega_q \lambda_q dB_q$$

where $\Omega_q = \frac{\partial W_s}{\partial U_n}$ is a normative social weight put on individual *n* and λ_n is - as before - the marginal utility of income. From Eq. 17 it is evident that standard cost-benefits analysis (CBA) that take un-weighted sums over user benefits in their calculation implicitly set at $\Omega_q = 1/\lambda_q$. This means, applying higher normative weights for persons with low marginal utility of income (typically wealthy persons). This has some implication for the discussion in *Essay 4*.

1.3 Travel Model Choice

Determinants of Travel Mode Choice

From an economic perspective travel mode choice is a rational choice that is largely explained by the monetary costs (ticket prices, fuel costs, road tolls etc.) and the travel times of different travel modes multiplied by the corresponding VTTS. By using WTP measures, it is also possible to convert other Level-of-Service (LoS) attributes, as the number of departures a day or waiting times, on a monetary scale, after which it is possible to analyze mode choice based on the generalized costs of all travel modes in the choice set. From this perspective the main challenge is (only) to measure/predict travel times and LoS, and to assign/estimate meaningful VTTS and other WTP measures.

In reality however, travel mode choice will depend on many (non-economic) factors such as the situational context (e.g. trip purpose, amount of baggage, weather conditions, car availability) or scheduling considerations (the traveller´s own, and that of other household members). In addition mode choice will be affected by habits, taste and personal characteristic of the travellers (e.g. age, gender, income, lifestyle or psychological elements as the perceived safety of modes or their green image).

To the extent that some of these elements are observable it is possible to account for them in travel mode choice models by model segmentation (as typically done for trip purposes), through choice set definitions (e.g. taking into account the varying availability of travel modes), or by incorporating elements in the deterministic utility function (e.g. using socio-demographic variables as interaction terms or explanatory variables). However, many factor may be unobserved (or insufficiently measured), and/or their impact on utility may be unknown. In this case they are omitted from the systematic part of the model, and affect the unobserved (random) part of the utility function (Eq. 8). There are choice models that can control for unobserved taste heterogeneity among travellers by assuming that parameters in the deterministic utility functions are not fixed but randomly distributed (e.g. Train 2009). Such models (formally introduced in section 1.5.) are utilized in *Essay 1, 2 and 3*[5].

Travel mode choice is also dependent on long term decisions as where to live and work, and whether to buy a car or not. This is a challenge for the prediction of long term travel mode choice as one first has to forecast (or assume) future land use and mobility pattern. Another complicating factor related to predicting travel mode choice - and travel demand in general - is that generalized costs, in particular travel times (i.e. characteristics of the supply side of the travel market), may dependent on the travel demand. This is particularly true for urban

[5] The empirical data in *Essay 4* is also based on such a model.

car traffic where the choice of mode (and route) of all travellers affects travel times due to congestion (which then again may affect mode choice).

Travel Mode Choice as an Integrated Part of Travel Demand

Modeling travel demand includes analyzing the following behavioral elements of travelers: whether to travel at all (trip frequency), where to travel (destination choice), when to travel (departure time choice), by which mode to travel (travel mode choice) and which route to take (route choice).

Rather complicated transport model systems are therefore used to predict travel demand. They typically consist of several integrated model components representing the different choice elements of travel demand, and typically iterate with a network model that calculates the physical conditions (congestion patterns and travel times) that emerge at the travel supply side given the predicted travel demand (see e.g. Flügel *el al* 2014).

To give some examples, the classical four step model (e.g. Ortúzar and Willumsen, 2011, p. 21) models travel mode choice as the third model step. This model component splits OD-matrices (containing the total demand for each OD pair obtained from the first two steps) into travel modes taking into account the travel times and other LoS that are iteratively calculated in the fourth model component (the network model). In the Norwegian transport models for regional and national person transport (RTM5, Madslien *et al* (2005) and NTM5, Hamre *et al* (2002)), travel mode choice is modelled together with destination choices in a multinomial or nested logit model. This acknowledges a direct relationship between the questions where to travel and how to travel. As the classical four step models, the Norwegian transport models are static, and do therefore not model departure time choice. Further, it operates with aggregated numbers for OD-pair representing geographical zones and model mode choice in a deterministic fashion; i.e. they apply a closed form logit formula to split demand in travel modes. In the dynamic and fully disaggregated model system MATSim (Raney and Nagel 2006, Nagel and Flötteröd 2012) mode choice is modelled via the selection and re-planning of traveller's (all-day) travel plans. Mode choice is modelled together with departure time and route choice in a stochastic fashion, and is simulated for each traveller.

Environmentally Sustainable Travel Models

Sustainable transportation is broadly understood as "satisfying current transportation and mobility needs without compromising the ability of future generations to meet these needs" (WCED 1987, Black 1996). Environmentally sustainability implies that transportation "does not endanger public health or ecosystems and meets needs for access consistent with (a) use of renewable resources below their rates of regeneration, and (b) use of non-renewable resources below the rates of development of renewable substitutes" (Wiederkehr et al. 2004). Departing form this definition, I define environmentally sustainable travel modes in a somewhat simplifying way as those forms of transportation that can largely be made with renewable energy resources. Besides walking and cycling, this includes for Norway (where electricity mainly stems from hydro-energy) trains, metro and electric cars.

Current and expected future CO_2 emissions per passenger kilometre of different travel modes (Figure 1) is often used as a measure of the environmentally sustainability of different travel modes.

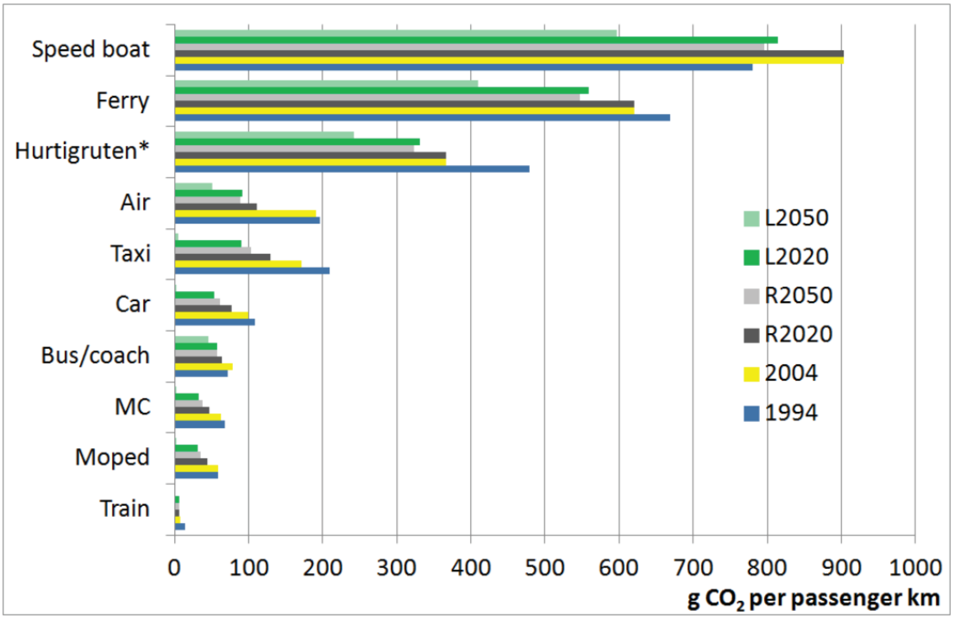


Figure 1: Historic and expected CO_2 emission rates for domestic Norwegian travel (R: reference scenario; L: low emission scenario). Source: Fridstrøm 2013 (based on numbers reported in Thune-Larsen et al. 2009).

Among the listed travel modes in Figure 1, trains have the lowest CO2 emission rates in Norway. In 2004 it was 8, 100 and 191 g CO_2 per passenger-km for train, personal car and airplanes, respectively. The emission rates for air and car are expected to drop considerable in the future (the low emission scenario assumes a full electrification of the car park in 2050).

Norwegian's Travel Mode Choice

The left panel of Figure 2 below gives the total picture of travel mode choice of Norwegians (including trips abroad). Car is the dominated choice of transport mode, followed by walking. Public transport and cycling have relatively low market shares. Air traffic has a very low share measured in trips, but a considerable share of 26% measured in passengers-km (see middle panel Figure 2). The Norwegian contribution to global warming is calculated to be even higher for air traffic than for car traffic (see right panel Figure 2).[6]

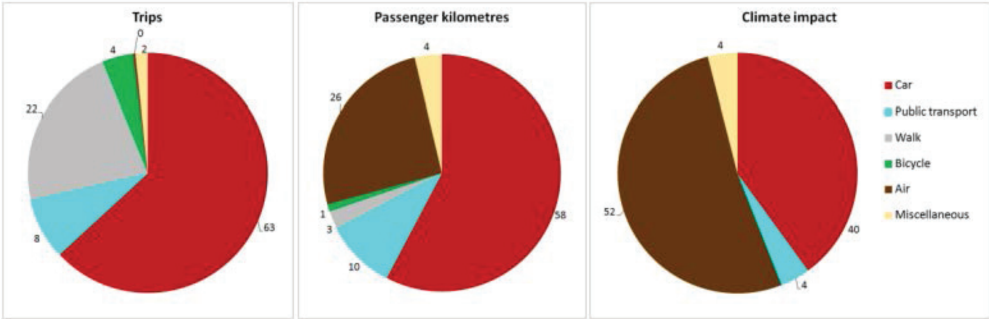


Figure 2: Norwegian travel mode choice and its effect on global warming; numbers from 2009 (source Fridstrøm and Alfsen 2014).

Great differences of CO2 emission rates across transport modes (Figure 1) motivate policy measures that improve services of environmentally sustainable travel alternatives or give incentives for travel mode shifts from rather carbon intensive modes to less carbon intensive modes.

For short distance travel, increasing the share of cycling is a political goal in Norway (e.g. Norwegian Public Roads Administration 2012).The provision of cycling facilities, e.g. by increasing the amount of cycling paths separated from motorized traffic, and is an important measure towards obtaining this goal (e.g. Lea *et al* 2012). *Essay 3* provides estimates of the cyclists' valuation of two types of cycling facilities.

For long distance travel between the largest cities in Norway, train services are rather poor (slow and with few departures a day) and air is the dominant mode of transportation; especially for business trips (Denstadli and Gjerdåker 2011). The Norwegian National Rail Administration has assessed the prospects of a high-speed rail (HSR) lines (Norwegian National Rail Administration 2012). The empirical background of *Essay 1 and 2* also relates to HSR. The study utilizing data from an independent stated choice study (see next chapter).

---

[6] Apart from higher emission rates in air compared to car (Figure 1), this is mainly related to contrails and cirrus cloud formed by aircrafts (Fridstrøm and Alfsen 2014).

1.4 Data

Different Types of Choice Data

The empirical analyses in this thesis are all performed on choice data. In choice data one has typically some information about the decision maker (the traveller), the alternatives that he/she had to choose from, some characteristics (attributes) of the alternatives and a nominal variable indicating which alternative was chosen. In *Essay 1 and 2,* choice sets consist of travel mode alternatives, while choice sets in the data underlying *Essay 3 and 4* consist of travel route alternatives for a given travel mode (part of the data in *Essay 3* also involves mode choice).

Two general types of choice data can be distinguished: (i) revealed preference (RP) data of actual choices made in real life situations, also referred to as revealed choices (RC); and (ii) stated choice (SC) data, i.e. hypothetical choices that are conducted in choice experiment surveys. SC is a particular type of stated preference (SP) data.

Some choice experiments are framed around an actual trip, usually the last relevant trip the respondent made in real life, i.e. a RC. This type of choice experiments is referred as pivoted experiments, and it typically conditions the choice sets in SC on the RC choice sets and/or bases (pivots) attribute values of alternatives in SC on the actual attribute values in real life.[7] This is done to make the choice task more realistic and relevant for the respondents. The conditioning of SC experiments on RC data involves some challenges in inference and in application due to different forms of self-selection (see more in section 1.6.). This is the topic of *Essay 1 and 4* (and it also has implications for *Essay 2*).

All SC data sets used in this thesis are based on pivoted designs. Most of them consist of binary choice task, which constitutes a limitation of the choice set compared to real choice sets (where travellers usually choose among more than two travel modes or travel routes).

[7] Train and Wilson (2008) distinguish between pivoted experiments and "sp-off-rp"; the latter being a special case of pivoted experiments where one alternative in SP is identical to the chosen RP alternative, and where the number of alternatives in the SP task corresponds to alternatives in the RP task.

The use of binary choices has some advantages in that it reduces the complexity of the choice tasks for the respondents. However, for some application it might involve severe information losses (see *Essay 2*).

In SC studies, the researcher can construct the choice task as she wishes, which makes it possible to elicit preference for new attributes or even new alternatives. While the possibility of observing choices in new choice situation is often the reason for SC studies in the first place, the hypothetical context of the choice situations lacking real-world consequences for the respondents is - at the same time - SC's greatest disadvantage.

There is an inherent uncertainty about the external validity of SC, and for new choice situation with non-yet-existing attributes or alternatives, there is typically little relevant real-world data to compare the results with. Conceptually it seems important to distinguish between three uncertainties in the utility functions estimated with SC:

(1) Are the marginal utility values in indirect utility function, V_{ni}, and thereby WTP values elicited in SC, non-biased? E.g. it is possible that respondents willingly over- or understate their implicit VTTS by choosing in a particular manner.

(2) Are the error variances (relative size of the random term) estimated on SC of the same (similar) size as in RC? E.g. error variance might be higher in SC due to fatigue of respondents, or lower due to fewer measurement errors in the attribute values.

(3) Are the error covariance among alternatives estimated on SC corresponding to RC? This issue is seldom discussed in the literature. It seems that the first paper that rigorously discussed it is Yánez *et al* (2010). *Essay 2* provides a discussion in the special case for different binary SC with one common alternative.

Arguably, (1) is most severe and can only be checked if good RC data is available for comparison. (2) is typically not an issue for WTP-studies as the error variances itself does not affect trade-offs between attributes. For prediction of choice behaviour, (2) is a challenge. However, given RC data, and given that (1) holds for at least one attribute, it is possible to assess and control for different error variance with joint RC-SC models (Morikawa (1989), Ben-Akiva and Morikawa (1990), Cherchi and Ortúzar (2006)).

Regarding RC data the following challenges can be mentioned:

a) It is often unobserved/not reported which other alternatives - besides the chosen alternative, which is typically observed in the field or reported in surveys - a traveller considered when making their RC. This makes the generation of choice sets an important part of RC analysis.
b) For some alternatives, in particular the non-chosen alternatives, it might be difficult to get precise data on (E.g. in the RC data for *Essay 1 and 2*, attribute values for the non-chosen alternatives in RC were not observed, and the necessary information had to be derived from external data sources).
c) The choice situations in real life are made in an uncontrolled environment, and many unobserved factors may influence the RC.
d) For parameter inference it is often a problem with RC data in that there is little variation and/or high correlation in explanatory variables.
e) Often one observes only one RC for a given traveller.

In particular related to d) and e), the advantages of SC are quite substantial and may motivate SC studies even if there is RC data available.

In SC studies the researcher has control over the variation and correlation of attributes characterising the alternatives. There exist different methods to do this, the classical being the orthogonal designs, where attributes values are constructed and combined in a way such that explanatory variables are uncorrelated. More recently so-called efficient designs (Rose et al 2008, Ortúzar and Willumsen 2011, p. 108) have become more popular. The general idea is to find a design that is likely to produce the "lowest" variance-covariance matrix of the estimation model given a specific choice model and prior value of the parameters. A common measure to determine the most efficient design is the "D-error" that is based on the determinant of the variance-covariance matrix (Atkinson and Donev, 1992).

In SC it is also possible to let respondents make repeated choices. In this case a (pseudo)[8] panel structure of the data can be constructed. Besides having practical advantages (e.g. reducing the data collection cost per observation), such data has advantages in estimation as e.g. accounting for unobserved taste heterogeneity gets more effective (see section 1.5).

High-Speed Rail data set (*Essay 1 and 2*)

In *Essay 1 and 2*, the main data stems from SC study about High-Speed Rail (HSR) collected in the TEMPO-project (Fridstrøm and Alfson 2014) by the Institute of Transport Economics (TØI) in 2010. As a research economist at TØI, I was personally involved in the design of the survey and choice experiments.[9]

In the following, I present some more background information about the data in addition to the general description given in *Essay 1 and 2*. More details about the data collection, survey designs and descriptive statistics are given in the conference paper Flügel and Halse (2012a) and the working documents of Halse (2012) and Flügel and Halse (2012b).

The two corridors for the SC study are "Oslo-Bergen" and "Oslo-Trondheim" (Figure 3).

[8] "Pseudo" in the sense that there is no real time dimension in these SC studies.
[9] Askill H. Halse (TØI) did most of the work related to the survey design, the technical implementation of the survey and the data collection. Regarding the generation of the efficient designs we got great support from Juan de Dios Ortúzar, Luis Rizzi and Julián Arellana.

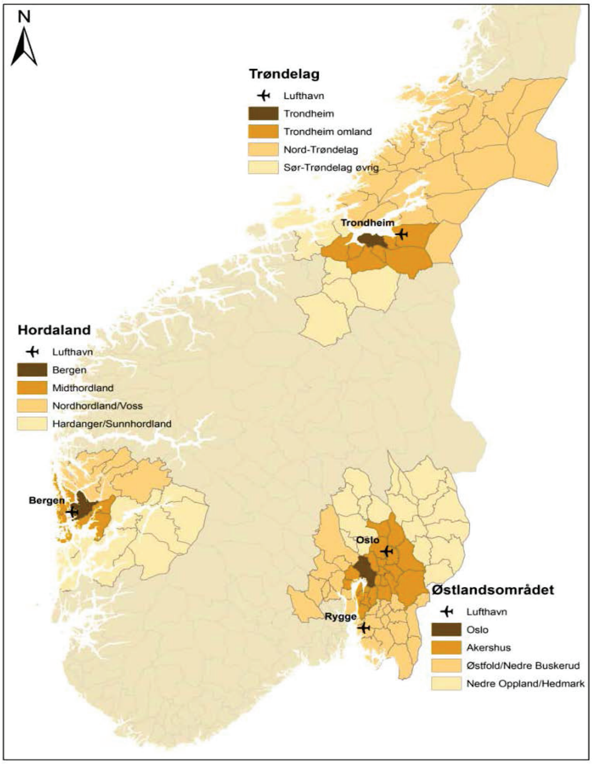


Figure 3: Illustration of corridors "Oslo-Bergen" and Oslo-Trondheim" (Denstadli and Gjerdåker 2011)

Prior to the SC study, a large-scale on-board[10] revealed preference (RP) study was conducted with the primary purpose to assess the current market of the main long distance corridors in Norway (Denstadli and Gjerdåker 2011). At the end of the questionnaire, participants of the RP study were asked to leave their e-mail address in order to receive an invitation for another survey concentrating on HSR.[11] The travellers that left a valid e-mail

[10] Car travelers were stopped on mountain passes; and bus, train and air passengers were asked to fill out the 2-page pen and pencil questionnaire on-board or while waiting to board theses transport modes.

[11] The sampling procedure in RP is choice-based, i.e. the probability that a traveler is included in the survey depends on the choice of transport mode - the dependent variable of the study. As the choice alternatives in the choice experiments in SC are conditioned on the RP-choices through the pivoted design, the dependent variable in SC also depends on the sampling probability. Bierlaire *et al* (2008) showed that estimates on choice based samplings are consistent except the alternative specific constants (ASC) for multinomial logit models and a certain subgroup of generalized extreme value (GEV) models called "block-additive GEV". We decided to use exogenous weights in any case (before we decided which choice model to apply) because there is no real market data for HSR for which the ASC could be adjusted. A concern we did not really attend too much to is a strong

address and the necessary information about their trip (ca. 25%) were invited to participate in a self-administered online-questionnaire. We conducted two pilot surveys, a bigger first pilot (N=221) and a smaller second pilot (N=67) before the main survey (N=605). For the analysis, we used data from all three surveys. The response rate for the SC-data (invited/completed questionnaires) was 33%.

The SC questionnaire consisted of (Flügel and Halse 2012b)[12]:

- an introduction presenting the purpose of the study and the reference trip which the respondents were to recall
- additional questions about the reference trip
- questions about how the respondents would had planned the trip if they were to do it by HSR instead[13]
- the choice experiments (CE1 and CE2)
- control questions about choice task interpretation and choice behavior
- questions about how often the respondents would travel along the corridor, with and without high speed rail
- questions about travel preferences and everyday behavior

In the choice experiments respondents were asked to recall the trip they reported in the RP study and to choose the travel mode given that a HSR option would have been available. Each respondent got 14 choice tasks where travel alternatives were characterize by the following six Level-of-service (LoS) variables: i) total travel costs per person, ii) in-vehicle travel time, iii) travel time to station/airport ('access time'), iv) travel time from station/airport ('egress time'), v) frequency (number of departures per day), and vi) the share of the ride spent in tunnels ('tunnel share'). In the first 8 choices (CE1) the choice was between the

hypothesis of selection bias when passing from the RP to the SC-questionnaire. People with strong opinions about HSR (pro and con) should be more inclined to leave their e-mail addresses and complete the questionnaire.

[12] Note that basic background variables about the travellers were already reported by the RP study.

[13] For instance which HSR station (we provided a list of possible stations) they would have departed from.

current mode with the reported LoS values and a HSR option (see example in Figure 4). In the consecutive six choices (CE2), the respondent could also state that he would not travel by either of the two modes (Figure 5). This "opt-out-alternative" was included as the LoS values for the current mode were varied in CE2 such that none of the two travel modes might have been relevant.

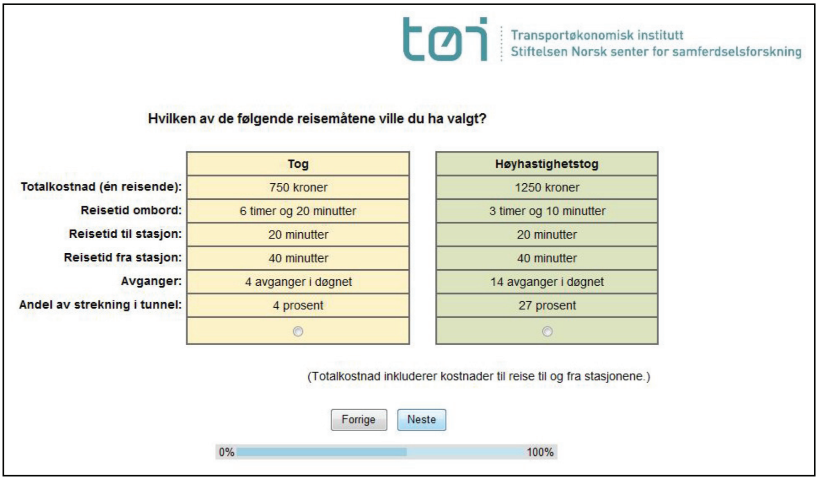


Figure 4: Presentation format, choice experiment 1 (CE1), example for train-user

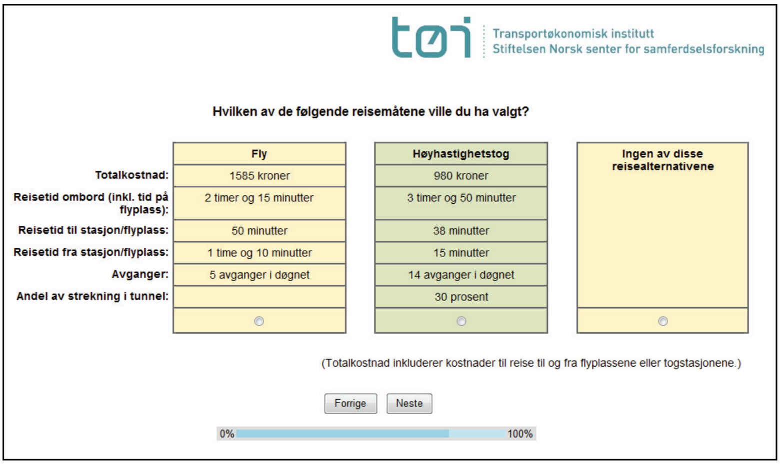


Figure 5: Presentation format, choice experiment 2 (CE2), example for air-user

Tables A1 and A2 in *Essay 1* provide details on how different LoS values were constructed. The combination of attribute values was made on a random basis in pilot 1, and according to an efficient design minimizing the D-error (mentioned above) in pilot 2 and the main survey.

For parameter estimation, the data was subdivided in leisure trip and work-related trips. All models in *Essay 1 and Essay 2* are performed on the bigger leisure subsample. Basic descriptive statistics of the general choice behaviour are presented in Table A3 in *Essay 1*.

Both *Essay 1* and *Essay 2* include analyses of revealed choices (RC). Unfortunately, neither the RP study (nor the SC study) asked for personal specific LoS variables for non-chosen alternatives in real life. For the modelling of RC we therefore derived the values of LoS-variable from the Norwegian National Transport Model (Hamre *et al* 2002) given geographical information we had for the trip reported in RP. This makes the underlying data of LoS-attributes in SC (reported, personal based) and RC (derived, zonal based) quite different.

Data from the "Norwegian Valuation Study" (*Essay 3 and 4*)

Essay 3 and 4 make use of data from the "Norwegian Valuation Study" (Samstad *et al* 2010a), a large-scale SP study conducted by the Institute of Transport Economics (TØI) and Sweco to elicit WTP values for different aspects of transportation. As a research assistant and later as a research economist at TØI, I was personally involved in this project.[14]

The main study consisted of a two-step web survey where respondents initially were recruited from a large and representative consumer panel administrated by Synovate Norway (formerly MMI (Markeds- og Mediainstituttet) and now part of the Ipsos Group). In the first round of data collection (wave 1) respondents were asked about a recent trip they did in real life, and then went through several pivoted route choice experiments related to their current

---

[14] My main role was to support Farideh Ramjerdi and Knut Veisten, two of the project leaders for different parts of the project with analysis of the pilot data, adjusting the experimental results based on the pilot results and the data procession as well as the estimation of discrete choice models in the main survey.

mode.[15] In wave 1, also referred to as the "Norwegian Value of Time Study", the main objective was to elicit WTP value related to travel time savings, travel time reliability and different comfort factors (Ramjerdi *et al* 2010). The second wave recruited respondents from the first wave, and concentrated on different aspects related to traffic safety and health (Veisten *et al* 2010, Flügel *et al* 2010a).

The initial data collection was in 2009. Because of a mistake in the coupling between wave 1 and 2, which made most of the results in wave 2 unreliable, the whole data collection was repeated in a slightly changed format in 2010.[16]

The analysis in *Essay 3* is based on the subsample of current cyclists (those respondents that reported that their last trip (over 10 minutes of duration) was made by cycling). It combines wave 1 and wave 2 data in the 2010 data collection. The information from the three choice experiments (one mode choice experiments and two route choice experiments) are used to obtain the results. For more information about the data, we refer to the detailed description in *Essay 3*.

Essay 4 includes no empirical analysis in itself but discusses the application of results obtained in the conference paper Flügel *et al* (2011) that uses data from the wave 1 data collection of 2009. The data contains binary route choice experiments that were made by of different user groups for different transport modes. For details related to the experimental design and data we refer to Ramjerdi *et al* (2010) and Flügel *et al* (2011). A detailed description of the whole data collection of the Norwegian Valuation Study is reported in Samstad *et al* (2010b).

[15] A few experiments were also framed around alternatives modes such to be able to estimate mode- and user- type-effects (see detailed discussion in *Essay 4*).
[16] Within the "Norwegian Valuation Study" only the 2010 data of wave 2 was analyzed (results of the "Norwegian Value of Time Study" were based on 2009 data only), but for some papers (like Essay 3) the 2010 data of wave 1 was also utilized.

1.5 Discrete Choice Modelling and Estimation

Random Utility Models

The framework for all discrete choice models in this thesis is that of random utility models (RUM).The general functional from of a RUM was already given in Eq. 3. The idea to decompose the (latent) indicator that explains choices in an observable part and a random part goes back to Thurstone (1927), who used this approach to model the outcome of psychological experiments where respondents had to compare two stimuli. The name RUM goes back to Marschak (1960) who called the indicator "utility" and introduced the concept to economics.

The general choice rule is that traveller *n* choices the alternative, denoted *i,* that generates the highest utility value (in consistency with Eq. 3 abbreviated here as *Z*) among all alternative *j* in the choice set C_n:

$$Z_{ni} > Z_{nj} \ for\ all\ j\ \in\ C_n \tag{18}$$

The typical assumption in economics and transport modelling is that the decision maker (traveller) knows each alternative's utility value and chooses strictly (rationally) according to these values such to maximise utility. RUM is therefore also referred to as random utility maximisation (e.g. McFadden 2000).[17]

The researcher can only observe the deterministic parts (V_{nj}) of utility and has to treat the remaining part (ε_{nj}) as random. The error term is always conditioned on the deterministic utility function. The better the deterministic utility function includes and measures the relevant attributes of alternatives and choice situations and the better it accounts for the preference of the decision maker for those attributes, the lower will the relative impact of the error term be (i.e. the less stochastic will choices appear to the researcher).

[17] There exist alternative concepts and choice rules, as the one proposed by Chorus *et al*. (2008) where decision makers make choices such to minimize regrets.

Conceptually it can be important to indentify different error sources. The following sources can be distinguished (Manski 1973): unobserved attributes, unobserved taste variations, measurement errors and instrumental/proxy variables. The latter two sources related to that variables included in the utility function are not precisely measured or represented by other (instrument) variables. The former two relate to the fact that there are factors not observed by the researcher which affect choices. "Unobserved attributes" means there are attributes describing the alternatives and the choice situations that are missing, while "unobserved taste variation" relates to subjective preferences and other unknown factors that vary across decision makers. Distinguishing between these latter two sources can be important as discussed in *Essay 2.* [18]

The probability of traveller *n* choosing alternative *i* can be calculated as follows (e.g. Train 2009, p. 15):

$$
\begin{aligned}
(19) \qquad P_{ni} &= Prob\big(Z_{ni} > Z_{nj}\ \forall\, j \neq i\big) \\
&= Prob\big(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}\ \forall\, j \neq i\big) \\
&= Prob\big(\varepsilon_{nj} - \varepsilon_{ni} > V_{ni} - V_{nj}\ \forall\, j \neq i\big) \\
&= \int I\big(\varepsilon_{nj} - \varepsilon_{ni} > V_{ni} - V_{nj}\ \forall\, j \neq i\big) f(\varepsilon_n) d\varepsilon_n\,,
\end{aligned}
$$

where $f(\varepsilon_n)$ is the joint density of the random vector $\varepsilon_n = \langle \varepsilon_{n1}, \dots, \varepsilon_{nJ} \rangle$ and I(.) is an indicator function, that is one if the statement in parentheses is true and zero otherwise.

Luce's Choice Axiom

The choice axiom by Luce (1959) is another concept that has been important for the development of discrete choice models. The axiom states that the relative probability of two alternatives *i* and *j* should be identical for all choice sets that contain *i* and *j*. This implies the property of the independence of irrelevant alternatives (IIA) which can be stated formally as:

[18] It is also possible to interpret the random part of being on the decision maker's part as in Thurstone's original model. The random part may then represent the "psychological state" of the decision maker. Conceptually it is however much more convenient to assume rational choice behavior on part of the decision makers and to explain the occurrence of apparent irrational choice (i.e. in choice the alternative *i* chosen even though $V_{ni} < V_{nj}$) due to unobserved factors or measurement errors.

(20) $$\frac{P_n(i|\tilde{C}_n)}{P_n(j|\tilde{C}_n)} = \frac{P_n(i|C_n)}{P_n(j|C_n)}, \quad i,j \in \tilde{C}_n \subseteq C_n$$

With a constant utility framework (corresponding to a situation where Z_{nj} in Eq. 18 is completely observed without a random element) Luce (1959) showed that his axiom holds when choice probabilities are calculated:

(21) $$P_n(i|C_n) = \frac{Z_{ni}}{\sum_{j \in C_n} Z_{nj}}.$$

Logit Models Applied in Thesis

The distribution assumption of the error terms and their variance-covariance structure, i.e. the joint density of $f(\varepsilon_n)$ in Eq. 19 defines the specific choice model.

All choice models in this thesis are logit models. Hence, they depart from the random elements being Gumbel (or "type I extreme value") distributed:

(22) $$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}}.$$

For the case error terms have identical variance $\sigma^2 = \frac{\pi^2}{\lambda^2 6}$ and are uncorrelated (e.g. if they are independent and identical, i.i.d., Gumbel distributed), McFadden (1974) showed that the following choice probabilities (via Eq. 19) are obtained:

(23) $$P_{ni} = \frac{e^{\lambda V_{ni}}}{\sum_{j \in C_n} e^{\lambda V_{nj}}},$$

where the scale variable λ can not be identified from the parameters in V and is therefore normalised, typically to value 1. Originally called the conditional logit model this model became popular under the name: multinomial logit model (MNL). With this model, McFadden linked RUM to Luce choice axiom as this model implies the IIA property (seen by comparing Eq. 23 with Eq. 21).

The i.i.d. assumption in the MNL can be relaxed in different ways. One relaxation is to allow for heteroskedastic error variances either related to alternatives (e.g. Train 2009, p. 92) or by subgroups of the data. The latter model, referred to as the heteroskedastic logit (HL) logit model is applied in *Essay 2* (and *Essay 3*). In this model the choice probabilities of alternative *i* being chosen by respondent *n* belonging to subgroup g_n are given as:

(24) $$P_{ni} = \frac{e^{v_{g_n} V_{ni}}}{\sum_{j \in C_{g_n}} e^{v_{g_n} V_{ji}}} \; for\ all\ i \in C_{g_n}$$

This model is particular useful in parameter estimation when data stems from different choice data sets as it is possible to allow for different error variances in the different subgroups of the pooled data.

Another relaxation of the i.i.d. assumption of the MNL is to allow the error terms of alternatives to be correlated. A class of logit models that achieves this are generalized extreme value models, GEV, (McFadden 1978) of which the nested logit NL model (Williams 1977; Daly and Zachary 1978) is the most prominent model. In the NL are alternatives grouped in non-overlapping nests *m* (i.e. each alternative can only enters one nest). The nesting structure has to be defined by the researcher prior to estimation. The choice probability of the NL is:

(25) $$P_{ni} = \frac{(\sum_{j=1}^{Jm} e^{\mu_m V_{nj}})^{\frac{\mu}{\mu_m}}}{\sum_{m=1}^{M} (\sum_{j=1}^{Jm} e^{\mu_m V_{nj}})^{\frac{\mu}{\mu_m}}} \frac{e^{\mu_m V_{nj}}}{(\sum_{j=1}^{Jm} e^{\mu_m V_{nj}})}$$

where μ_m are scale parameters applied to alternatives in nest *m* and μ represent the scale parameters for choice between nests (typically normalised to 1). It can be shown (e.g. Bhat 1997) that the correlation between the utilities of two alternatives *i* and *j* is given by:

(26) $$Corr(Z_i, Z_j) = \left(1 - (\frac{\mu}{\mu_m})^2 \right) d_{ij}$$

where d_{ij} is one when *i* and *j* belong to nest *m* and zero otherwise. The theory of GEV imposes some restrictions on the scale parameter. In the NL, we need $\mu_m \geq \mu > 0$. This implies that the utility of the nested alternatives must be positively correlated. Correlated error terms among alternatives imply that these nested alternatives are closer substitutes to each other (compared to non-nested alternatives). The typical interpretation is that nested alternatives share unobserved attributes (Williams 1977).

An extension of the NL models is the cross-nested logit model, CNL (Vovsha 1997, Bierlaire 2006). It allows alternatives to enter several nests and can therefore imply a richer correlation structure among alternatives. For more details about the CNL we refer to *Essay 2* in which the NL and CNL models play an important role.

Another relaxation of the i.i.d. assumption is to allow for correlation among repeated choices from a given respondent. This is particularly useful in SC data where choice experiments are typically repeated several times (the pseudo panel structure mentioned in section 1.4). The most popular choice model that can deal with inter-personal correlation of error terms is the mixed logit model.

The general expression of the probability function of mixed logit (ML) models is as follows (e.g. Train 2009, p. 135):

(27) $$P_{ni} = \int L_{ni}(\boldsymbol{\beta}) f(\boldsymbol{\beta}|\boldsymbol{\theta}) d\boldsymbol{\beta}$$

with

(28) $$L_{ni}(\boldsymbol{\beta}) = \frac{e^{V_{ni}(\boldsymbol{\beta})}}{\sum_j e^{V_{nj}(\boldsymbol{\beta})}}$$

and $f(\boldsymbol{\beta}|\boldsymbol{\theta})$ being the density function given the to-be-estimated parameters $\boldsymbol{\theta}$ that describe this density. The set of β constitute additional random elements (besides ε_{ni}) that are integrated out when calculating probabilities.[19] $f(\boldsymbol{\beta}|\boldsymbol{\theta})$ is referred to as the mixing distribution.

The researcher can specify whatever mixing distribution she prefers. E.g. assuming a coefficient(s) $\boldsymbol{\beta}$ to be normal distributed with mean *b* and covariance *W* one obtains:

(29) $$P_{ni} = \int \frac{e^{V_{ni}(\beta)}}{\sum_j e^{V_{nj}(\beta)}} \Phi(\boldsymbol{\beta}|\, b, W) d\beta$$

b and *W* can then be estimated from the data. In *Essay 1* we also apply lognormal mixing distributions.

[19] Note that Eq. 27 (and Eq. 29 below) involves multiple integrals when more than one coefficient is assumed randomly distributed.

For repeated choices it is typically assumed that parameters vary over persons but are constant for all choice of a person. In this case, Eq. 28 is specified as being the probability of a particular sequence of choices, **i**={i_1, ..., i_t , .., i_T}, being made:

(30) $$L_{ni}(\beta) = \prod_{t=1}^{T} \frac{e^{V_{nit}(\beta n)}}{\sum_j e^{V_{njt}(\beta n)}} .$$

For interpretation it makes a difference which coefficients of the utility function are assumed random, e.g. if β related to marginal utilities of attributes or to additional error components (independent of a specific attribute). This is discussed further in the next subsection.

ML models are powerful models that can approximate any RUM (McFadden and Train 2000). Hence they can also be used to relax the IIA assumption underlying the MNL model and are thereby alternative models to the NL and the CNL. However, ML models are not as tractable as choice probabilities (and logsum formulas) cannot be calculated by a closed form and need to be simulated. This is probably the main reason why ML models are not (yet) frequently applied in transport model systems.

Specification of Utility Functions

The specification of utility functions is obviously a crucial step in discrete choice analysis, and finding and testing for the most appropriate specification can be a time-consuming activity in practice.

Important elements of utility functions are alternative specific constants, $\beta_{0,i}$ (of which one needs to be normalized). The alternative specific constant (ASC) can be interpreted as the average effect of the alternative *i*'s unobserved factors. Note that in route choice experiments where alternatives are typically unlabelled (just referred to as "Route A" or "Route B"), the ASCs are often found to not be significantly different from each other. This makes sense as respondents usually have no preference for "A" or "B" per se. In mode choice experiments, the ASCs are typically distinct from each other and important for forecasting.

Utility specification involves the selection of explanatory variables (attributes of alternatives, variables describing the choice situation and/or the decision makers) and their parameterisation in the deterministic utility function. In this connection one question relates to which parameters should be modelled as generic for all alternatives and which ones should be modelled alternative-specifically[20].

Another question relates to the assumption about the functional relationship between attributes and utility. As described in more detail in *Essay 1*, one can distinguish between three basic types of parametric specification of V_{ni}:

1. Linear-in-parameter and linear-in-variable specifications. For example:

$$V_{ni} = \beta_{0,i} + \sum_k (\beta_k A_{nik}) \ . \tag{31}$$

This is by far the most popular specification and it has the nice feature that marginal utilities are directly given by the β_k such that MRS (and WTP) measures can simply be calculated as parameter ratios.

2. Linear-in-parameter but non-linear-in-variable specifications where estimated parameters β_k are combined linearly but where variables (attributes) are non-linearly transformed by some assumed transformation parameters T_k. For example:

$$V_{ni} = \beta_{0,i} + \sum_k (\beta_k A_{nik}^{T_k}) \ . \tag{32}$$

3. Non-linear-in-parameter specifications as for instance the Box-Cox transformation, BCT (Box and Cox 1964):

$$V_{ni} = \beta_{0,i} + \sum_k (\beta_k A_{kni}^{(\lambda_k)}) \tag{33}$$

with

$$A_{kni}^{(\lambda_k)} = \begin{cases} \frac{A_{nik}^{\lambda_k} - 1}{\lambda_k} & \lambda_k \neq 0 \\ log A_{nik} & \lambda_k = 0 \end{cases} \tag{34}$$

[20] This may be decided purely on statistical grounds, e.g. with likelihood ration (LR) test one can test if the assumption of generic marginal utilities over alternatives holds. But, the decision can also be made conceptually. E.g. parameters in the model in *Essay 2* are kept generic because the information about alternative specific variable only steam from particular subgroups, something which might be undesirable for forecasting models that are meant to predict generic choice behaviour. The cost coefficient is normally held generic over alternatives as it seems not plausible that the marginal utility of money depends on in which travel mode the money is expended on (even if this might be indicated statistically on a specific data set.)

A λ_k(or T_k) that is estimated (assumed) to be different from unity implies that the marginal utility of the attribute k varies systematically with the size of that attribute. The general advantage of BCT (over Eq. 32) is that one can estimate the curvature of the functional relationship between attribute and utility from the data.

It is also possible to model the marginal impact as varying over respondents. This can for example be achieved by a parameterisation of β_k with observable variables using fixed coefficients or by assuming β_k to be randomly distributed.[21] In the latter case, mixed logit models are usually used, and are then referred to as random coefficient models. How marginal utility values are assumed to vary over respondents (systematically or randomly) also has implications for how MRS and WTP measures are derived.[22]

Note that is it sometimes also possible and meaningful to specify the whole model in WTP space as done in the "integrated approach" to VTTS (Fosgerau *et al* 2006), which was applied in the conference paper by Flügel *et al* (2011) that provides the empirical results discussed in *Essay 4*.

The inclusion of additional error components is another important feature of the specification of utility functions. Dependent on the specification, the error components may capture correlation between alternatives or correlation between repeated choices of the respondent. In *Essay 1* a simple error components model is used to connect SC made by the same respondents, and to account for preferences for alternative per se (i.e. independent the concrete attribute values) differs across respondents.

[21] It is also possible to combine both, i.e. to parameterize β_k with observed variable and a random term. This is done in the estimation model underlying the results in *Essay 4*.
[22] In case of a parameterisation of β_k (the one in the numerator and/or denominator) a MRS is calculated for each respondent by enumerating the vector that is used to parameterize the coefficients. Then summary statistics as the average value (or percentile) are typically reported. In case of random coefficients and in case only the coefficient in the numerator is modelled random, the distribution assumption of that coefficient yields the distribution assumption of the MRS measure and the mean value of the MRS can be calculated as the mean value of the nominator divided by the fixed value of the denominator. If also the denominator follows a random distribution one needs to simulate the resulting distribution of the MRS (e.g. Sillano and Ortúzar 2005). In practise, the simulation will often produce outliers and distributions should be truncated (at some arbitrary value) before mean values are calculated. In Essay 3 we conveniently assume fixed denominators in the analysis.

Estimation Technique

All estimation models in this thesis are based on classical inference with maximum likelihood.[23] This standard approach involves that one finds the set of parameters such that the likelihood of the observed choices is maximised given the probability functions and the data.

The general form of the likelihood function for N independent choices is:

$$(35) \qquad L = \prod_{n=1}^{N} \prod_{i \in C_n} P_{ni}{}^{y_{ni}}$$

or in the more convenient log-transformed version:

$$(36) \qquad LL = \sum_{n=1}^{N} \sum_{i \in C_n} y_{ni} \, lnP_{ni} \; .$$

with y_{ni} equalling 1 if alternative *i* was chosen and equalling zero otherwise.

For linear-in-parameter logit models with fixed coefficients as in Eq. 31 and Eq.32, the likelihood function is globally concave (McFadden 1974) such that a unique solution can be found with standard numerical methods. In non-linear-in-parameter models as in Box-Cox-Transformation model or mixed logit models, the likelihood function is normally non-concave which involves the danger that the numerical method gets stack in a local maximum.

In mixed logit models, the probability functions (Eq. 27) that enter to log-likelihood function has no closed-form and needs to be simulated by taking draws from the mixing distribution. The simulated log-likelihood function is then given as (Train 2009, p.144):

$$(36) \qquad SLL = \sum_{n=1}^{N} \sum_{i \in C_n} y_{ni} \, ln\check{P}_{ni}$$

with

$$(37) \qquad \check{P}_{ni} = \frac{1}{R} \sum_{r=1}^{R} L_{ni}(\beta^r)$$

where *r* is the index of the draw from density $f(\boldsymbol{\beta}|\boldsymbol{\theta})$ and $L_{ni}(\beta^r)$ is the logit formula (Eq. 28) calculated at that draw.

[23] In an unpublished working paper I employed Bayesian estimation methods, which has some advantages over the classical estimation techniques, in particular related to the inference of ratios (WTP) of two random coefficients. See also Bergland and Flügel (2013) for analysis with Bayesian estimation methods on the SC data on HSR.

For the estimation of mixed logit models in this thesis we use Halton draws (Halton1960). All models are estimated with the software package Biogeme (Bierlaire 2003, 2008).

1.6 Introduction of Main Challenges covered in the Thesis

Taste Heterogeneity (TH)

Taste of people differs. Even the abstract time allocation model described in section 1.2 illustrates this fact by showing that the VTTS depends on the wage rate (Eq. 7), a variable that naturally differs among travellers. Taste heterogeneity (TH) is typically also found empirically; e.g. in the Norwegian Value of Time Study the estimated distribution of VTTS for different travel mode and distances had 99-percentiles of factor 5-10 compared to the corresponding mean values (Ramjerdi et al 2010, p. 239). The major part of this variation is unobserved, i.e. not explained by typical background variables such as age and gender. This is consistent with the typical finding - also seen in this thesis - that the inclusion of randomly distributed coefficients improves the goodness-of-fit of choice models considerably (and typically much more than by the inclusion of interaction effects of background variables that are typically observed/reported in surveys). Note that preferences, as for time savings, do also differ within a person on a day to day bases or from situation to situation (e.g. depending on the perceived time pressure in a particular situation and other factors what are largely unobserved by the researcher).

In the context of predicting choice behaviour it is important to take TH into account even if one is only interested in average effects. This is because choice models as the logit model are non-linear models implying that the probability calculated at an average utility level differs in general from the average probability calculated at varying utility levels (see e.g. Ben-Akiva and Lerman 1985, p. 136).

In the context of eliciting WTP and user benefits, it is notable that models that do not account for TH can only provide a fixed value (point estimate of WTP). This value is typically interpreted as the average value for the sample, and may - given that the sample is representative - be used in cost-benefit-analysis. Note that the point estimate of WTP will often differ from the mean of the distribution estimated by random coefficient mixed logit model. As the latter model is superior in explaining choices it is generally recommended to derive WTP value from mixed logit models.[24] This approach is therefore chosen in *Essay 3*.

In the other essays, the challenges of TH are more subtle. In *Essay 2*, different degree of TH in different subgroups of the (pooled) data, is argued to impact the relative sizes of the estimate group scale parameters (Eq. 24). In *Essay 1 and 4*, TH is argued to be a driver for self-selection to transport modes and attribute value ranges that affect results in pivoted choice experiments (see next subsection).

Self-selection

We refer to self-selection as a mechanism in which persons allocate to different groups according to their own preference/abilities.

While there are some groups you can't self-select to (e.g. group "male" versus group "female"), self-selection is truly a part of everyday life (at least in societies where people can largely choose according to their own will): What kind of work to pursue or which travel mode to choose will largely dependent on one own preferences/abilities.

Analysing behaviour (or eliciting WTP) in various groups prone to self-selection one will often find differences because the members of the groups are not randomly allocated but allocated through a mechanism that depends on person's preferences/abilities being directly related to or correlated with the analysed behaviour (or WTP).

Let us give an example of relevance for this thesis. [25] The value of saving travel time in air traffic (VTTS in air) is likely to be higher than the corresponding value in bus traffic (VTTS in bus) when the groups these values are elicited on depend on the real world mode choice, i.e.

[24] A challenge with mixed logit models is that the researcher has to assume the mixing distribution. Estimated mean values typically depend on the distribution assumption (Fosgerau 2005).
[25] Self-selection is also the driver in the classical case of sample selection bias (Heckman 1979). A typical example for self-selection in transport studies is residential self-selection (e.g. Mokhtarian and Cao 2008).

if the value elicited in air (bus) is based on travellers that usually take air (bus). Differences in VTTS are likely because air users indicate with their travel mode choice (air preferred over bus) that they have relative high preference for fast travel modes: a preference which is likely to be correlated with the preference for travel time savings. Given the natural self-selection to travel modes (SSTM) in real life and the study design (using only current air users to elicited VTTS in air), the differences in VTTS for various travel modes will depend on difference in the preferences of user groups (taste heterogeneity across user groups). This may or not may be desirable in economic appraisal as discussed in-depth in *Essay 4*.

Regarding the design of SC studies, the researcher may have the possibility to avoid that the self-selection in real life carries over to the choice experiments. E.g. she can design route choice experiments for travel modes independent of respondent's mode choice (e.g. letting all respondents make choices between two bus routes independent if the respondent took bus in real life or not). This is however seldom done. Two main reasons may be distinguished:

(i) The researcher is often interested in finding representative preferences (e.g. WTP) for the actual users of that travel mode. In the Norwegian Valuation Study, for instances, the choice experiments involving cycling (used in *Essay 3*) were only given to travellers that cycle in real life.[26]

(ii) The typical intention with pivoting choice experiments is to make the SC more realistic and meaningful (e.g. air users might regard hypothetical choices between two bus services as irrelevant and they might therefore put little effort in the choices or even drop out of the survey).

In the data underlying the empirical discussion in *Essay 4,* respondents went through two types of route choice experiments one in their current mode and one in their first-best

[26] In *Essay 3* we thus can provide a representative WTP for cycling facilities for current cyclists. For economic appraisals of certain projects however, it would have been important to get also information about the WTP of potential cyclist (which potential differs due to self-selection). However the pivoted design based on the current mode did not contain this information.

alternative mode. In this case the experimental design enables to control for self-selection to the current mode, but, as argued in the essay, the self-selection problematic is not completely eliminated with this data because there is also a self-selection to the alternative mode (e.g. car-users that would alternatively take the plane might differ from car-users that would alternatively take the bus).

Self-selection is also an issue in the SC data for HSR described in section 1.4. Here self-selection to travel modes (SSTM) in real life carried over to choice experiments in that respondents got binary choice task between their current mode and HSR (omitting all non-chosen alternatives in real life from the choice sets in SC). A challenge in *Essay 2* relates to the objective of building a generic forecasting model from that data. When estimating a joint model (on the pooled data of all binary choice experiments), the coefficient that are specified as alternative-specific for current modes will one be inferred from the choice behaviour of the underlying user-group. E.g. an alternative-specific marginal utility for alternative air will only depend on preferences of (current) air users. This is problematic when one aims for a forecasting model with full choice set that is meant to apply to all future decision makers (independent of the current self-selection). As already mentioned in footnote 20, the approach chosen in *Essay 2* (*Essay1*) was to use generic marginal coefficient (despite that this is likely to be inferior in explaining choice in the binary choice experiments).

In *Essay 1* we introduce another type of self-selection where the allocation is not going into distinct groups but in overlapping value ranges. This "self-selection to attribute values" (SSAV) is relevant when attribute values of alternatives are based on real-life decision. In the experimental design of HRS study, for instance, the pivoted design implied not only that choice sets dependent on the current travel mode choice but also that attribute values that respondents were confronted with in the choice experiments depend (were pivoted on) the actual attribute values of the chosen travel mode in real life. E.g. a car driver that made a lot of pauses on his trip from Oslo and Trondheim resulting in a relative long travel time in the actual car trip (say 10 hours) would get choice tasks between car and HSR where the car alternative is characterised with travel time of about 10 hours. For the estimation on the cross-sectional data, the problem that emerges is that the marginal utility of travel time estimated at around 10 hours will be relative low because the estimation at this value is mainly based on those respondents that travelled slowly in their actual trip indicating that these respondents have low marginal utility from reductions in travel time. As argued in *Essay 1*, this SSAV may therefore be an explanation for counter-theoretical findings of estimated utility functions with decreasing marginal dis-utility of travel time and travel cost.

Econometrically, self-selection may result in endogeneity in the choice model, i.e. a situation where explanatory variables (as travel time and travel cost) are endogenous and correlated with the unobserved factors that underlie the random terms in RUM. Endogeneity will in general lead to inconsistent estimation of coefficients and is therefore a challenge in many pivoted choice experiments. Only when the unobserved factors influencing the real world-choices are assumed to apply additively and with same size to all alternatives in SC, the use of fixed coefficient logit model may yield that the endogeneity in the SC attributes "takes a form that cancels out of the behaviour model" (Train and Wilson, 2008, p.196). Thus self-selection and the resulting endogeneity should be considered in more general models (e.g. random coefficient models) and in cases the assumption about additively applying unobserved factors is not reasonable (as in case of SSAV as argued in *Essay 1*).

Estimation on combined data sets

Many models in this thesis are estimated on combined data sets. In our case, different types of stated choice experiments involving different choice sets and/or attributes are analysed in one single estimation model. Doing so, some coefficients are typically specified to be generic for some or all subsamples (otherwise there is no point in doing a joint estimation). The very motivation for this approach differs from application to application.

Regarding the HSR data set (*Essay 1 and 2*), pooling different binary choice data allows for estimating a common utility function for HSR as needed for a generic forecasting model with full choice sets. *Essay 2* also involves an estimation model where the (combined) SC is pooled with RC data. In our case this is done with the primary goal of getting sufficient information about the inter-alternative error structure.

In *Essay 3*, two route choice experiments, one with and one without an additional casualty risk attribute, are pooled together in order to be able to test if the assumption of generic marginal utilities of cycling facilities (independent of the inclusion of the casualty risk attributes) holds.

In the estimation model in Ramjerdi *et al* (2010) and Flügel *et al* (2011) that provides the empirical results for *Essay 4,* route choice experiments for different user groups and involving different transport modes are jointly analysed. The generic coefficients are included to control for different elements of the experimental design (e.g. the absolute size of travel time savings in the choice experiments, a design variable that is arbitrary chosen by the researcher and may differ between subsamples).

A common feature of joint estimation models is to allow for heteroskedastic error variances across the subsamples, acknowledging that error variance might differ across subsamples. This is applied in *Essay 3,* and is discussed in detail in *Essay 2*.

Decomposition of user benefits

The second half of the thesis discusses the quantifications and applications of different components of user benefits associated with cycling facilities (*Essay 3*) and travel time savings in different travel modes (*Essay 4*).

The main topic in *Essay 3* is the quantification of user benefits associated with two cycling facilities: i) the provision of cycling paths that are separated from motorized traffic, and ii) the reduction of crossings where cyclist potentially have of stop and wait. There are different benefits for cyclists associated with such improvements of the infrastructure: i) it may reduce travel times for cyclists, ii) it may increase cyclist's traffic safety, and iii) it may increase the convenience and comfort of cycling. Eliciting WTP for cycling facilities it is important to get information about the size of these different components in order to be able to avoid double-counting of user benefits in cost-benefit analysis (CBA). E.g. if the user benefits from reduced travel time and increased traffic safety is already included in the CBA calculus by the VTTS and the Value of Statistical Life (VSL), then one should include a WTP for cycling facilities that only relates to the remaining user benefits (comfort, convenience, risk reduction of minor accidents etc). *Essay 3* provides a description of an experimental design and statistical analysis that provide WTP for cycling facilities with and without controlling for casualty risk (but in both cases time savings are controlled for).

Essay 4 concerns the two components that make up the differences in VTTS across travel modes used to calculate user benefits in the economic appraisal /CBA of transport project. Differences in VTTS can be decomposed in mode and user-type effects (Wardman 2004). The former relates to differences in the comfort level of travel modes, and the latter relates to differences between user-groups; e.g. the official VTTS for long distance travel is 204 NOK/h for air and 98 NOK/h in long-distance train trips (Samstad *et al* 2010a). For project appraisal it may be important information how much of this difference is associated with the two components, e.g. how much of the higher VTTS in air is due to that travelling by air is more uncomfortable than travelling by train (independent of the user-group), and how much is due to that typical air-users have higher preferences for time savings (independent of in which travel mode these savings are experienced in).

1.7 Contributions of Essays

Contributions of Essay 1

The title of *Essay 1* is: "How to explain decreasing marginal dis-utility of travel time and cost? – Self-selection to attribute values in travel mode choice models." The essay is co-authored with Askill H. Halse and Ståle Navrud, and is submitted to *Transportation* (3.12.2014).[27]

[27] An earlier version of this paper was presented at the Kuhmo Nectar Conference on Transportation Economics 2012 in Berlin. The paper was also presented at a workshop on discrete choice modeling organized by Jürgen Meyerhoff (Berlin 2013).

The overall contribution of Essay 1 is the introduction of the concept of self-selection to attribute values (SSAV) as an explanation for theoretically contradicting decreasing marginal dis-utility of travel time and cost in mode choice models.

This research departs from by an apparent - but rarely discussed - contradiction that empirical results that relax the linear-in-variable assumption, e.g. by Box-Cox Transformations, often suggest decreasing marginal dis-utility for travel cost and travel time (Gaudry 2010) while economic theory suggest increasing marginal disutility (Koppelmann 1981). As an explanation, we suggest SSAV (see introduction in section 1.6.). This self-selection problem has - to our knowledge - not been discussed in the mode choice literature before.

We argue that taste heterogeneity is the driver of this self-selection and we find some evidence that controlling for unobserved taste heterogeneity by random coefficient models (so called 'Box-Cox mixed logit models' (Orro et al. 2005)) can retrieve marginal utility function for the cost attribute more in line with theoretical expectation. In this sense, the paper also adds to the literature on the interdependence between taste heterogeneity and non-linearity (Orro et al. 2005, Pinjari and Bhat 2006).

The paper contributes also by discussing the severity of SSAV for different types of choice data and we find empirical support for the severity of SSAV being higher when choice models are based on personal specific attributes (as typically in pivoted SC studies) compared to derived values from zonal data (as in our RC data).

Finally, acknowledging that we cannot prove (just indicate) the effect of SSAV with the available data, we suggest venues for new studies on this topic.

We hope that this essay has an impact on the transportation-modelling field, raising the awareness of self-selection issues in experimental design, estimation and forecasting.

Contributions of Essay 2

The title of *Essay 2* is: "Methodological challenges in modelling the choice of mode for a new travel alternative using binary stated choice data - the case of high speed rail in Norway." The essay is co-authored with Askill H. Halse, Juan de Dios Ortúzar and Luis I. Rizzi and is submitted to *Transportation Research Part A* (19.12.2014).[28]

The overall contribution of *Essay 2* is an in-depth discussion of the identification of the most appropriate inter-alternative error structure of the mode choice forecasting model including a new travel alternative when estimation is based on binary stated choice data of respondent's current mode and the new option (which in our case is HSR).

The discussion departs from the pragmatic approach chosen in the official HSR assessment study in Norway (Atkins 2012) that directly used estimated group scale parameters from the estimation model as nest parameters in the forecasting model. We show that this approach is mathematically wrong without making very strong assumptions about the relative scale between upper and lower level in nested logit models and about the correlation between current travel modes, i.e. information that is not available in these binary choice data.

We also argue that it is important to control for taste heterogeneity in the estimation model such that group scale parameter can readily be interpreted as representing unobserved attributes of alternatives. Only then the translation of variance in binary choices (as measured by group scale parameters in the estimation model) into covariance of alternatives in multinomial choices (as captured by nest parameter in the forecasting model) is reasonable.

The paper also contributes in terms of empirical illustrations using our own HSR data set. We showed that revealed choices between current modes may provide some of the missing information. We also estimate a cross-nested logit model on the pooled data illustrating that

[28] An earlier version of this paper was presented at the European Symposium on Quantitative Methods in Transportation Systems 2012 in Lausanne. The paper was also presented at the Nordic Meeting on Transport Economics (2014) in Oslo.

the correlation structure may be more complex then what a nested logit model with non-overlapping nests is able to represent.

The paper makes a strong case for collecting multinomial SC (as done in Yáñez *et al*. 2010) for the objective of building a (multinomial) forecasting model that includes a new travel alternative. We therefore believe that the paper is of relevance for the design and analysis of similar future travel mode choice studies.

Contributions of Essay 3

The title of *Essay 3* is: "Valuation of Cycling Facilities with and without Controlling for Casualty Risk." The essay is co-authored with Farideh Ramjerdi, Knut Veisten, Marit Killi and Rune Elvik and is published in *the International Journal of Sustainable Transportation*, Volume 9**,** page 364-376. It was accepted for publication 20. April 2013.

The overall contribution of *Essay 3* is a novel comparison of estimated WTP for cycling facilities when the user benefits related to reduced casualty risk is controlled for by an additional variable, and when it is not.

In contrast to the other essays in this thesis, *Essay 3* is a rather straightforward empirical paper that in its exposition puts a lot of weight on the description of the experimental design and data collection as well the statistical analysis that are used to obtain the empirical results. The work is strongly related to the two rounds of data collection (wave 1 and wave 2) in “Norwegian Valuation Study” for which results had been previously reported separately (Ramjerdi *et al* 2010, Veisten *et al* 2010). The overall goal of the project with regards to the cycling subsample was elicited representative individual WTP for cycling time (wave1), different cycling facilities (both wave 1 and wave 2) and risk reduction in cycling (wave 2). The particular contribution of *Essay 3* is the estimation of the joint data of wave 1 and wave 2; making it possible to test rigorously (using likelihood ratio tests) if the valuation of cycling facilities differs between wave 1 and wave 2, where the latter includes the additional casualty risk attribute. We find significantly different WTP, and the value of cycling facilities is almost halved when the casualty risk is controlled for.

A challenge in eliciting WTP related to cycling is that this mode does not involve monetary travel costs as a natural attribute. We chose a two-step approach similar to Börjesson and Eliasson (2012), in which respondents performed a series of mode choice experiments involving cycling and a paid mode (either bus or car), before being asked a sequence of cycling route choice experiments. The estimated VTTS from the mode choice model was then used to transfer the estimated MRS between cycling facilities and travel time to a monetary scale.

The paper provides representative values applicable to CBAs both in scenarios where one decides to incorporate VSL as a separate element, and in scenarios where one decides not to.

Contributions of Essay 4

The title of *Essay 4* is: "Accounting for user type and mode effects on the value of travel time savings in project appraisal: Opportunities and challenges." The essay is published in *Research in Transportation Economics*, Volume 47, page 50-60. It was accepted for publication 16 July 2014.[29]

The overall contribution of Essay *4* is an elaboration on how the quantification of user type and mode effects on between-mode differences in the VTTS (Fosgerau *et al* 2010, Ramjerdi *et al* 2010, Flügel *et al* 2011) can be used in economic appraisals.

Doing so I introduce a new concept called “mode effect dependent equity value” which takes into account mode effects but avoids that user-group differences due to self-selection are accounted for. I also discuss the concept of VTTS of switching modes presented in Ramjerdi *et al* (2010), that does take into account user type effects between people who switch travel modes and those how do not. It is argued that these two approaches can in different ways

[29] The paper was presented at the Annual Meeting of the Norwegian Association of Economists ("Foskermøtet") 2014 in Oslo.

improve over existing approaches of VTTS. Compared to the standard approach of "mode specific VTTS" (which is the current practise in Norway), the "mode-effect dependent equity value approach" improves on the equity dimension while the "VTTS of switching modes approach" can improve on the precision dimension.

The essay introduces formal definitions and interprets user type and mode effects both theoretically and empirically (using results obtained in Flügel *et al* (2011)). It is stressed that there are several limitations in the current data that mainly relate to that the first-best-alternative mode was used to derive the two effects in recent studies.

Finally, the essay presents a stylized case study illustrating that the applied VTTS approach may indeed affect the ranking of transport projects in economic appraisals.

I hope that this essay has some impact on the ongoing debate about which VTTS approach is the most desirable for economic appraisal in Norway as it provides some strong argument for reconsidering the current “standard mode-specific VTTS approach”. Furthermore, realising the limitations of the existing data is important input to the design of the next overall “Valuation Study” in transportation in Norway. As pointed out in the essay, framing the alternative mode choice experiment around a randomly distributed travel mode (instead of the first-best alternative mode) will enable us to better control for self-section issues.

1.8 References of Introduction Chapter

Atkinson, A. C. & Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford: Clarendon Press.

Becker, G. (1965). A Theory of the Allocation of Time. *The Economic Journal*, 75: 493-517.

Ben-Akiva, M. & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, Mass.: MIT Press.

Ben-Akiva, M. & Morikawa, T. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General*, 24 (6): 485-495.

Bergland, O. & Flügel, S. (2013). *A Bayesian Approach to Latent Heterogeneity of Preferences in Travel Model Choice Models*. International Choice Model Conference, Sydney.

Bhat, C. R. (1997). Covariance heterogeneity in nested logit models: Econometric structure and application to intercity travel. *Transportation Research Part B: Methodological*, 31 (1): 11-21.

Bierlaire, M. (2003). *BIOGEME: A free package for the estimation of discrete choice models*. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model. *Annals of Operations Research*, 144 (1): 287-300.

Bierlaire, M. (2008). An introduction to BIOGEME Version 1.8: http://biogeme.epfl.ch/v18/tutorialv18.pdf.

Bierlaire, M., Bolduc, D. & McFadden, D. (2008). The estimation of generalized extreme value models from choice-based samples. *Transportation Research Part B: Methodological*, 42 (4): 381-394.

Black, W. R. (1996). Sustainable transportation: a US perspective. *Journal of Transport Geography*, 4 (3): 151-159.

Blayac, T. & Causse, A. (2001). Value of travel time: a theoretical legitimization of some nonlinear representative utility in discrete choice models. *Transportation Research Part B: Methodological*, 35 (4): 391-400.

Box, G. E. P. & Cox, D. R. (1964). An Analysis of Transformation. *Journal of Royal Statstical Society. Series B (Methodological)*, 26 (2): 211-252.

Börjesson, M. & Eliasson, J. (2012). The value of time and external benefits in bicycle appraisal. *Transportation Research Part A: Policy and Practice*, 46 (4): 673-683.

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

Chorus, C. G., Arentze, T. A. & Timmermans, H. J. P. (2008). A Random Regret-Minimization model of travel choice. *Transportation Research Part B: Methodological*, 42 (1): 1-18.

Daly, A. & Zachary, S. (1978). Improved multiple choice models. In Hensher, D. A. & Dalvi, M. Q. (eds) *Determinants of Travel Choice*. Sussex: Saxon House.

Denstadli, J. M. & Gjerdåker, A. (2011). Transportmiddelbruk og konkurranseflater i tre hovedkorridorer: Institute of Transport Economics, Oslo. (in Norwegian)

DeSerpa, A. (1971). A Theory of the Economics of Time. *The Economic Journal*, 81: 828-846.

Flügel, S., Veisten, K. & Ramjerdi, F. (2010). Den norske verdsettingsstudien. Verdien av tid, sikkerhet og miljø. Utrygghet - Verdien av redusert rasfare og bedre tilrettelegging for syklende og gående. Oslo: Institute of Transport Economics (TØI). (in Norwegian)

Flügel, S., Hjorth, K. & Ramjerdi, F. (2011). *Between-mode differences in value of travel time estimates: Evidence from Norway*. Kuhmo-Nectar Conferance Stockholm.

Flügel, S. & Halse, A. H. (2012). *High-speed rail ridership forecasts for the corridors Oslo-Bergen and Oslo-Trondheim - Preliminary analysis* Oslo: Institute of Transport Economics. Unpublished manuscript.

Flügel, S. & Halse, A. H. (2012). *Non-Linear Utility Specifications in Mode Choice Models for High-Speed Rail*. Kuhmo Nectar Conference 2012 Berlin.

Flügel, S., Flötteröd, G., Kwong, C. K. & Steinsland, C. (2014). Evaluation of methods for calculating traffic assignment and travel times in congested urban areas with strategic transport models. Oslo: Institute of Transport Economics

Fosgerau, M., Hjorth, K. & Lyk-Jensen, S. V. (2006). *An integrated approach to the estimation of the value of travel time*. Association for European Transport and Contributors.

Fosgerau, M., Hjorth, K. & Lyk-Jensen, S. V. (2010). Between-mode-differences in the value of travel time: Self-selection or strategic behaviour? *Transportation Research Part D: Transport and Environment*, 15 (7): 370-381.

Fridstrøm, L. (2013). Norwegian Transport Towards the Two-Degree Target: Two scenarios. Oslo: Institute of Transport Economics (TØI).

Fridstrøm, L. & Alfsen, K. H. (2014). Vegen mot klimavennlig transport Oslo: Institute of Transport Economics (TØI). (in Norwegian)

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: Université de Montréal.

Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In Government, N. T. P. S. o. t. F. (ed.). *General Series*. New York. 173-196 pp.

Halse, A. H. (2011). *Demand for high speed rail in Norway: Survey design and data collection*: Institute of Transport Economics, Oslo. Unpublished manuscript.

Halton, J. (1960). On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2: 84–90.

Hamre, T. N., Rekdal, J. & Larsen, O. I. (2002). Utvikling av den nasjonale persontransportmodellen i fase 5. Del B: Estimering av modeller. *TØI rapport*. Oslo: Institute of Transport Economics.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47 (1): 153-161.

Hicks, J. R. (1956). *A Revision of Denad Theory*. Oxford: Oxford University Press.

Hotelling, H. (1938). The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates. *Econometrica*, 6 (3): 242-269.

Jara-Diaz, S. R. (1990). Consumer's surplus and the value of travel time savings. *Transportation Research Part B: Methodological*, 24 (1): 73-77.

Jara-Díaz, S. R. & Guevara, C. A. (2003). Behind the subjective value of travel time - The perception of work, leisure, and travel from a joint mode choice activity model. *Journal of Transport Economics and Policy*, 37, part 1: 29-46.

Jara-Díaz, S. R. (2007). *Transport Economic Theory*: Emerald Group Publishing Limited.

Koppelman, F. S. (1981). Non-linear utility functions in models of travel choice behavior. *Transportation*, 10 (2): 127-146.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74: 132-157.

Lea, R., Haug, E., Selvig, E., Hamre, T. N., Embry, M. & Colville-Andersen, M. (2012). Klimaeffekter av økt sykling og gåing, og suksesskriterier for økt sykling, Civitas: Copenhagenize Consulting og Numerika. (in Norwegian)

Luce, R. D. (1959). *Individual choice behaviour* New York: Wiley.

Madslien, A., Rekdal, J. & Larsen, O. I. (2005). Utvikling av den regionale modeler for persontransport i Norge. Oslo: Institute of Transport Economics. (in Norwegian)

Manski, C. (1973). *The Analysis of Qualitative Choice*. Cambridge, Mass: MIT, Department of Economics, MIT.

Marschak, J. (ed.) (1960). *Binary Choice Contrains on Random Utility* Stanford symposium on mathematical methods in the social sciences. Stanford: Stanford Univ. Press.

Marschall, A. (1920). *Principles of Economics*. 8 ed. London: Macmillan.
Mas-Colell, A., Whinston, M. D. & Green, J. G. (1995). *Microeconomic theory* Oxford, New York: Oxford University Press, Inc.
McFadden, D. (ed.) (1974). *Conditional logit analysis of qualitative choice behavior*. Frontiers in Econometrics. New York: Academic Press.
McFadden, D. (1978). Modelling the Choice of Residential Location. In Karlqvist, A., Lundqvist, L., Snickars, F. & Weibull, J. (eds) *SPATIAL INTERACTION THEORY AND PLANNING MODELS*, pp. 75-96. Amsterdam.
McFadden, D. (2000, Jul 02-07). *Disaggregate behavioural travel demand's RUM side: A 30-year retrospective*. 9th Meeting of the International-Association-for-Travel-Behaviour-and-Research, Gold Coast, Australia: Pergamon-Elsevier Science Ltd. 17-63 pp.
McFadden, D. & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15 (5): 447-470.
Mokhtarian, P. L. & Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B: Methodological*, 42 (3): 204-228.
Morikawa, T. (1989). *Incorporating Stated Preference Data in Travel Demand Analysis*: MIT, Department of Civil Engineering.
Nagel, K. & Flötteröd, G. (2012). Agent-based traffic assignment: going from trips to behavioral travelers. In Pendyala, R. & Bhat, C. (eds) *Travel Behaviour Research in an Evolving World*, pp. 261–293. Bingley, United Kingdom: Emerald Group Publishing.
Norwegian National Rail Administration (2012). Høyhastighetsutredningen 2010-2012 Konklusjoner og oppsummering av arbeidet i Fase 3- Del1 (In Norwegian)
Oort, C. (1969). The evaluation of travelling time. *Journal of Economics and Policy*, 3 (3): 279-286.
Orro, A., Novales, M. & Benitez, F. G. (2005). *Nonlinearity and Taste Heterogeneity Influence on Discrete Choice Model Forecast.*
Ortúzar, J. d. D. & Willumsen, L. G. (2011). *Modelling Transport*, vol. 4th edition. Chichester: John Wiley and Sons.
Pinjari, A. & Bhat, C. (2006). Nonlinearity of Response to Level-of-Service Variables in Travel Mode Choice Models. *Transportation Research Record: Journal of the Transportation Research Board*, 1977 (1): 67-74.
Ramjerdi, F., Flügel, S., Samstad, H. & Killi, M. (2010). Den norske verdsettingsstudien - Tid. Oslo: Institute of Transport Economics. (in Norwegian)
Raney, B. & Nagel, K. (2006). An improved framework for large-scale multi-agent simulations of travel behavior. In Rietveld, P., Jourquin, B. & Westin, K. (eds) *Towards better performing European Transportation Systems*, pp. 305–347: Routledge.
Rose, J. M., Bliemer, M. C. J., Hensher, D. A. & Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, 42 (4): 395-406.
Samstad, H., Killi, M., Flügel, S., Veisten, K. & Ramjerdi, F. (2010). Den norske verdsettingsstudien - Databeskrivelse. (in Norwegian)
Samstad, H., Ramjerdi, F., Veisten, K., Navrud, S., Magnussen, K., Flügel, S., Killi, M., Halse, A. H., Elvik, R. & San Martin, O. (2010). Den norske verdsettingsstudien - Sammendragsrapport. Oslo: Institute of Transport Economics. (in Norwegian)
Sillano, M. & Ortúzar, J. d. D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A*, 37 (3): 525-550.
Small, K. A. & Rosen, H. S. (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica*, 49 (1): 105-130.
Small, K. A. (1982). The Scheduling of Consumer Activities: Work Trips. *The American Economic Review*, 72 (3): 467-479.
Norwegian Public Roads Administration (2012). Nasjonal sykkelstrategi - sats på sykkel - Grunnlagsdokument for NTP 2014-2023. Oslo: Vegdirektoratet. (in Norwegian)

Thune-Larsen, H., Hagman, R., Hovi, I. B. & Sandberg Eriksen, K. (2009). Energieffektivisering og CO2-utslipp for innenlands transport 1994-2050.
Thurstone, L. L. (1927). A law of comparative judgment. *Psycho. Rev.*, 34.
Train, K. & Wilson, W. W. (2008). Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research Part B: Methodological*, 42 (3): 191-203.
Train, K. (2009). *Discrete Choice Methods with Simulation* Published by Cambridge University Press. Available at: http://elsa.berkeley.edu/books/choice2.html.
Veisten, K., Flügel, S. & Elvik, R. (2010). Den norske verdsettingsstudien. Ulykker - Verdien av statistiske liv og beregning av ulykkenes samfunnskostnader. Oslo: Institute of Transport Economics (TØI). (in Norwegian)
Vovsha, P. (1997). Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area. *Transportation Research Record: Journal of the Transportation Research Board*, 1607 (-1): 6-15.
Wardman, M. (2004). Public transport values of time. *Transport Policy*, 11 (4): 363-377.
WCED. (1987). *Our common future*. Oxford: Oxford University Press.
Wiederkehr, P., Gilbert, R., Crist, P. & Nadia, C. (2004). Environmentally Sustainable Transport (EST): Concept, Goal, and Strategy – The OECD's EST Project. *European Journal of Transport and Infrastructure Research*, 4 (1): 11-25
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): 285-344.
Yáñez, M. F., Cherchi, E. & Ortúzar, J. d. D. (2010). *Defining inter-alternative error structures for joint RP-SP modeling: some new evidence*. 89th Annual TRB Meeting, Washington, D.C, USA.

Essay 1

How to explain decreasing marginal dis-utility of travel time and cost? – Self-selection to attribute values in travel mode choice models

Authors:

Stefan Flügel[a] (stfl@nmbu.no),

Askill H. Halse[b] (ash@toi.no),

Ståle Navrud[a] (stale.navrud@nmbu.no)

[a] *School of Economics and Business, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway*

[b] *Institute of Transport Economics (TOI), Gaustadalleen 21, NO-0349 Oslo, Norway*

Corresponding authors: Stefan Flügel (stfl@nmbu.no, 0047 41520576)

Keywords: self-selection; attribute values; Box-Cox-Transformations; choice experiments; taste heterogeneity; high-speed rail

Abstract

Decreasing marginal dis-utility of travel time and cost are often found in travel mode choice models that are estimated on cross-sectional data. From economic theory, however, we should expect that households have *increasing* marginal dis-utilities from money and time spent on transportation. Departing from this contradiction, which is widely neglected in the literature, we argue that taste heterogeneity combined with the self-selection of travellers to attribute values may be a likely contributor to these counter-theoretical empirical results. This *self-selection to attribute values* is potentially present when utility is estimated based on cross-sectional revealed preference data, or pivoted stated choice data where attribute values are based on respondents´ actual trips. We present an empirical example of modelling the choice of high-speed rail in Norway based on pivoted stated choice data. Applying a Box-Cox mixed logit model with log-normal distributed beta-coefficients, we find that controlling for unobserved taste heterogeneity retrieves a Box-Cox parameter value for the travel cost attribute which is in line with the theoretical expectations.

1. Introduction

Conventional travel mode choice models express a traveler's utility for a given mode as a function of (monetary) travel costs, travel time and other level-of-service (LoS) variables. However, there is mixed advice regarding the correct functional form of the marginal impact of these variables on utility. Ad-hoc travel mode choice models usually include a linear-in-variable specification of the deterministic utility function implying that (dis-)utility is increasing proportionally (with constant returns) with increasing LoS-variables. As travel cost and travel time are reducing the household's disposable income and leisure time, respectively, the assumption of constant marginal dis-utility implies that constant returns are assumed for disposable income (money) and leisure time. However, as Koppelman (1981; p.131) notes: "*Economic theory suggests that there is diminishing marginal utility associated with commodities such as time and money which are available to the household (McFadden, 1976). That is, the value of time or money saved in travel depends on how much time or money remains for other activities after satisfaction of basic needs. Implicitly, this argument suggests the existence of increasing marginal dis-utility of time or money expended in travel"*. For travel mode choice models, this result from economic theory, which is shown formally later in the paper, implies *increasing* marginal dis-utility from the travel cost and travel time attribute.

Box-Cox transformation (BCT) models (Box and Cox 1964, Gaudry and Wills 1978) make it possible to test the linear-in-variable assumption in a neat fashion. As can be expected, BCT models regularly find that the assumption of constant marginal utility in linear models is restrictive. However, as summarized in the next paragraphs, a majority of BCT studies find *decreasing* marginal dis-utility contradicting expectation from micro-economic theory. Such findings are by no means a particularity of BCT models. For example, many researchers have found that a simple log or square-root transformation (implying decreasing returns) fits the data better than a linear (non-transformed) specification, while e.g. square functions are seldom applied for travel cost and time.

The literature review includes mode choice models, and has some focus on models including High-Speed-Rail (HSR) options. HSR implementation often implies massive improvements in travel time and frequency compared to traditional train. Testing linear functions against BCT models has therefore received particular attention in models including HSR options.

Gaudry and Wills (1978) were probably the first to estimate BCT models for mode choice analysis. Using aggregated data for city-pairs in Canada they show that Box-Cox models clearly outperform their linear counterparts. They find decreasing negative marginal effects of travel cost (fare) and travel time, and a decreasing positive effect of service frequency. To our knowledge, the first BCT analysis of the prospects of HSR is Gaudry and Le Leyzour (1994). Using disaggregate RP-data for the Quebec-Windsor corridor in Canada, they estimate BCT-parameters in a discrete choice framework. For travel time they find increasing marginal dis-utility for business trips, and decreasing marginal dis-utility for non-business trips. For travel cost they find decreasing marginal dis-utility for both trip purposes. An early study based on SP-data is Laferrière (1993), cited in Gaudry (2010). For the same corridor, Quebec-Windsor, Laferrière op. cit. finds decreasing marginal dis-utility for cost, travel time and access time; both for business and non-business trips. Besides Canada, other studies on non-linear impacts of LoS-variables in intercity mode choice models were conducted in Sweden (Algers and Gaudry 1994), Germany (Mandel *et al.* 1997) and France (Gaudry *et al.* 1998). All these studies find decreasing marginal dis-utility for cost and travel time. Gaudry (2010, table 7) reviews the results from these above and a few other studies, and points out the consistency in which decreasing marginal dis-utility is estimated.

Gaudry (2010, table 8) also reviews the results from different *urban* travel mode choice models (all RP-studies). In the 13 studies listed, travel cost exhibits decreasing marginal dis-utility as well. For travel time, the results are more ambiguous, and a majority of urban travel studies listed there seems to find *increasing* marginal disutility from travel time.

In Norway, BCT models have previously not been used to assess the mode choice of HSR.[1] Atkins (2012) uses a linear-in-variable model, but adds a log-term for travel cost and finds diminishing sensitivity to increasing fares. For the time attribute Atkins (2012) adds a dummy variable for travel times under 6 hours which is estimated to be positive. This makes the impact of travel time a stepwise linear function with a jump in dis-utility at 6 hours.[2]

In summary, all reviewed studies find support for decreasing marginal dis-utility from monetary travel cost. For travel time the results are ambiguous, but the majority of *intercity* mode choice models also finds decreasin*g* marginal dis-utility for travel time. Surprisingly, this apparent contradiction with economic theory is hardly discussed in the literature.

In this paper we argue that self-selection, and in particular what we refer to as *self-selection to attribute values*, can be an explanation for this counter-theoretical empirical finding. As will be discussed, this type of self-selection is most likely a challenge for all cross-section data sets where attributes values are inferred from real world data, i.e. revealed preference (RP) studies, and stated preference (SP) studies where attribute values are pivoted on RP-data.

In the empirical part of the paper we discuss estimation results of a 'pivoted' SP-study on HSR in Norway (Flügel and Halse 2012). The self-selection problem applies here because respondents' preferences for LoS-variables (e.g. their price sensitivity) will influence their real-world decision (e.g. the price of the ticket they purchase), and – by the design of 'pivoted' SP-studies (Train and Wilson 2008) – also the range of attribute values in the SP choice experiment (CE). Because of the resulting endogenous distribution of respondents to

[1] BCT-models in Norway seem only to have been applied to freight analysis (Fridstrøm and Madslien 2002). Based on SP-data they find decreasing marginal dis-utility from both travel time travel cost..

[2] Recently, two M.Sc. theses on BCT in mode choice models for HSR have been conducted in Sweden (Jiang 2010, Wang 2011). Wang (2011) finds decreasing marginal dis-utility both for the cost and time variable in consistency with most other intercity (intercity) mode choice models. The estimation results in Jiang (2010) seem somewhat limited but indicate decreasing (or constant) marginal dis-utility from cost and time as well.

attribute values, the estimated BCT parameters might indicate *decreasing* marginal dis-utilities (e.g. for the travel cost attribute) over the entire *cross-sectional* sample even though *individual* utility functions are characterized by *increasing* marginal dis-utilities.

The paper also contributes to the existing literature by suggesting 'Box-Cox mixed logit models' (Orro *et al.* 2005) as a way to retrieve expected BCT-parameters.

2. Non-linear specifications of the indirect utility function

In a standard discrete choice set up, a decision maker (traveller) *n* chooses alternative (transport mode) *i* whenever the utility Z_{ni}, exceeds the utility Z_{nj} of every other alternative *j* available in *n's* choice set, C_n:

$$(1) \qquad Z_{ni} > Z_{nj} \qquad for\ all\ j \in C_n, j \neq i$$

As researchers cannot observe all factors which affect the choice, one decomposes Z_{ni} into a deterministic part V_{ni} (a function of parameters $\boldsymbol{\theta}$ and explanatory variables X , and which is also referred to as the indirect utility function) and a random part ε_{ni}.

$$(2) \qquad Z_{ni} = V_{ni}(\boldsymbol{\theta_{ni}}, \boldsymbol{X_{ni}}) + \varepsilon_{ni}$$

For inference, the explanatory variables, X_{ni}, are assumed to be exogenous, i.e. uncorrelated with ε_{ni}.

We focus on three basic types of parametric specification of V_{ni} :

1. Linear-in-parameter and linear-in-variable specifications (in our case, the β-coefficients, except the alternative specific constant $\beta_{0,i}$, differ over attributes *k* but not over alternatives - and decision makers):

$$(3) \qquad V_{ni} = \beta_{0,i} + \sum_k (\beta_k X_{kni})$$

A main characteristic of linear models are constant marginal effects of explanatory variables X_k. Eq. 3 also implies that only differences in attribute values between alternatives, and not absolute values, matter for the choice. By subdividing X_k into value ranges, one can estimate different marginal effects for different absolute values of X_k while keeping a linear-in-variable specification. However, such piecewise-linear specifications increase the number of parameters to be estimated, and it is often difficult to justify the choice of value ranges.

2. Linear-in-parameter but non-linear-in-variable specifications, whereby T_k are fixed transformation parameters:

(4) $$V_{ni} = \beta_{0,i} + \sum_k (\beta_k X_{kni}^{T_k})$$

The assumption of constant marginal effects, and the implication that only differences in attribute values matter, is relaxed by this specification for all values of T_k other than 1. However, the true value of T_k is generally not known, and the researcher has to assume a value of T_k prior to the estimation of β_k. For example if a square-root transformation is specified, T_k is implicitly set to be 0.5. Note, that the choice of T_k will decide whether the model implies decreasing or increasing marginal effects of attribute *k.* It is, however, possible to combine transformations of the same attribute; e.g. combining a logarithmic transformation with a linear one. Then a more flexible functional form can be achieved.

3. Non-linear-in-parameter specifications, where $X_{kni}^{(\lambda_k)}$ indicates some function of explanatory variables X_k and parameters λ_k (to-be-estimated):

(5) $$V_{ni} = \beta_{0,i} + \sum_k (\beta_k X_{kni}^{(\lambda_k)})$$

A prominent specification, the Box-Cox transformation (BCT), was introduced in the seminal paper on variable transformation by Box and Cox (1964).

$X_{kni}^{(\lambda_k)}$ in (5) is specified for BCT models as:

(6) $$X_{kni}^{(\lambda_k)} = \begin{cases} \frac{X_{kni}^{\lambda_k} - 1}{\lambda_k} & \lambda_k \neq 0 \\ log X_{kni} & \lambda_k = 0 \end{cases}$$

The transformation is defined for X_{kni}> 0. The λ_k's are generally referred to as power parameters. By inspecting the second partial derivatives of Eq. 6, it is easily recognizable that λ_k > 1 (λ_k < 1) means increasing (decreasing) marginal effects. In the special case of λ_k = 1 we are back to the linear model. Figure 1 illustrates the functional form of *V* with respect to a single attribute when the attribute is an economic good X_k (i.e. whenever $\beta_k > 0$), and when it is an economic 'bad' ($\beta_k < 0$).

Figure 1: Illustration of functional forms for Box-Cox transformation

Illustration of V(β_k, λ_k, X_{kni}) holding other X_{kni} constant	$\beta_k < 0$ (X_k is a "bad", e.g. travel time t_i)	$\beta_k > 0$ (X_k is a good, e.g. leisure time L=$\tau - t_i$ with τ being the total time endowment)
$\lambda_k < 1$	 $\frac{\partial v}{\partial x_{kni}} < 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} > 0$	 $\frac{\partial v}{\partial x_{kni}} > 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} < 0$
$\lambda_k = 1$	 $\frac{\partial v}{\partial x_{kni}} < 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} = 0$	 $\frac{\partial v}{\partial x_{kni}} > 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} = 0$
$\lambda_k > 1$	 $\frac{\partial v}{\partial x_{kni}} < 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} < 0$	 $\frac{\partial v}{\partial x_{kni}} > 0; \frac{\partial^2 V}{\partial x_{kni}{}^2} > 0$

While a λ_k different from unity implies that the marginal utility of the LoS-variable k varies systematically with the size of the LoS-variable for every respondent, it is also possible to model the marginal impact as varying over respondents. This can for example be achieved by segmenting β_k by observable variables (e.g. income), or by assuming β_k to be randomly distributed. Random variable specifications are most often applied for linear-in-variable models as in Eq. 3 such that the marginal utility for a given decision maker is constant over the range of attributes. Then the taste heterogeneity of decision makers is accounted for, but not a (possible) non-linear impact of the value of the variable. Orro *et al.* (2005) seem to be the first who combine the two effects in a so-called 'Box-Cox mixed logit model'.

3. Suggestions by theory

Previously, Blayac and Causse (2001) have derived expressions of the indirect utility function *V* from a general time allocation model, and shown that the functional form of *V* supports non-linearity with respect to travel time and cost.[3] They do, however, not discuss which sign the parameters that represent the second partial derivates should have according to micro-economic theory.

The purpose of this section is not to find the most general expressions of *V*. Rather we want to show what functional forms of *V* (as exemplified by the values of BCT-parameters) can be expected when *V* is theoretically derived from a time allocation model that complies with standard assumptions of micro-economic theory.

We depart from the fixed income version of the famous goods/leisure model by K. Train and D. McFadden from 1978 (Jara-Díaz and Farah (1987), Jara-Díaz (2007)). The representative agent maximizes utility *U*, a function of consumer goods *G* and leisure time *L* given fixed income *I*, working time W, total time endowment τ as well as travel cost of the chosen alternative, c_i, and travel time of the chosen alternative, t_i, with alternative *i* being in choice set *M*. Formally:

(7) $$Max\ U(G, L)\ s.t.$$

$$G + c_i = I$$

$$L + W + t_i = \tau$$

and $i \in M$

Microeconomic theory imposes some restrictions on the utility function U with respect to the economics goods G and L. One central assumption is the *desirability assumption* that implies

[3] Their derivation is technically sophisticated (using Taylor approximations) and they focus on measures for the Value of Time. According to Orro *et al.* (2005), a working paper by Blayac and Causse from 1998 (and a lecture by Blayac from 2003) also look at the derivation related to BCT parameters.

monotony, i.e. the more of a good (bad), the better (worse). A second central assumption is the *convexity assumption* (Mas-Colell *et al.* 1995), which implies convex indifference curves between two economic goods, (here G and L). The economic intuition is that one has to compensate successive losses in leisure time with increasing income (and vice versa). This concept is well understood, and motivates e.g. overtime wages rates being higher than the usual wage rate,

Convex indifference curves imply a quasi-concave utility function (ibid, page 49)), such that utility increases at a lower or proportional rate with increasing levels of the economic good. Cobb-Douglas functions that restrict the elasticities to be between zero and one are among the functions that have this characteristic. In the theoretical transport economics literature, a Cobb-Douglas function is often chosen as the utility function for disposable money and leisure time; see e.g. Jara-Díaz (2007, page 60).

One might claim that micro-economic theory is restrictive in the sense that it does not allow for oversaturation (following from the desirability assumption). By the convexity assumption it does, however, allow for (close to) satiation, e.g. that marginal increments provide essentially no additional utility. The latter is regarded as sufficient in our case of general consumption and leisure time.

From Eq. 7, the indirect utility function (conditioned on c_i and t_i) of travel mode *i* is defined as:

(8) $$V_i \equiv U(I - c_i, \tau - W - t_i)$$

Thus, the indirect utility V_i depends on the income level after having paid for travel $(I - c_i)$, and the remaining (leisure) time after spent time traveling and working $(\tau - W - t_i)$. [4]

[4] The typical assumption in time allocation models is that leisure activities include all those activities for which one assigns more time than the minimum required. In this theoretical exercise, we thus assume that the "time spend travelling" does not include additional travel time that arises from deliberately

The two standard assumptions of economic theory imply that $\frac{\partial U}{\partial \mathrm{G}}, \frac{\partial U}{\partial \mathrm{L}} > 0$ and $\frac{\partial^2 U}{\partial \mathrm{G}^2}, \frac{\partial^2 U}{\partial \mathrm{L}^2} \leq 0$.

In combination with Eq.7 and Eq. 8 we therefore obtain:

(9) $$\frac{\partial V}{\partial c_i} = -\frac{\partial U}{\partial \mathrm{G}} < 0 \,, \quad \frac{\partial U}{\partial(\tau - W)} = -\frac{\partial V}{\partial t_i} < 0$$

and

(10) $$\frac{\partial^2 V}{\partial {c_i}^2} = \frac{\partial^2 U}{\partial \mathrm{G}^2} \leq 0 \,, \quad \frac{\partial^2 V}{\partial {t_i}^2} = \frac{\partial^2 U}{\partial L^2} \leq 0$$

What do (9) and (10) imply for the fundamental curvatures of $c_i$and t_i? Considering a standard BCT-model; i.e. a BCT-model that does not apply interaction effects with income or other covariates that might affect marginal dis-utility:

(11) $$\bar{V}_i = \beta_{0,i} + \beta_{cost} {c_i}^{(\lambda_{cost})} + \beta_{time} {t_i}^{(\lambda_{Time})} + \ldots.. \,,$$

Conditions (9) and (10) postulate that we should obtain β_{cost}, β_{time} < 0 and λ_{cost}, λ_{Time} ≥1 (compare Eq. 6 and Figure 1). Thus, in order to comply with microeconomic theory the curvature of LoS-variables must imply *increasing or constant* marginal dis-utility of travel time and travel cost.

Theory does not suggest any particular value (of unity or higher) for BCT parameters of travel time and cost. For travel cost, values close to unity might be expected especially when travel costs only make up a minor part of disposable income (which is the case in Norway for many types of trips). Indeed, when the travel cost is small, a linear approximation should suffice to explain the marginal effect (even if the marginal dis-utility of travel cost is indeed increasing over the whole range of possible travel cost values).

---

choosing a slower, but otherwise equal, travel mode, e.g., if some travellers enjoy taking slower trains because travelling is perceived pleasurable in itself. In this sense, “deliberately slow” travel would not reduce the budget for leisure time.

4. Self-selection to attribute values (SSAV)

4.1. Introduction

The question arises why so many empirical studies find decreasing marginal dis-utility for travel time and cost. One answer may lie in the data used in most larger-scale empirical studies, namely cross-sectional data. A persistent challenge in empirical studies on cross-sectional data is *self-selection*.

By *self-selection to attribute values* (SSAV)[5] we refer to a situation where travellers self-select to continuous values of explanatory variables (i.e. attribute values), according to their preferences.

For a systematic allocation of travellers to attribute values, taste heterogeneity has to be present. In order for taste heterogeneity to manifest itself in the attribute values, the travel market must work such that 1) LoS variables differ between and within travel options 2) incentives for self-selection to attribute values are provided, i.e. wealthy persons must get some incentive (extra benefit) from buying higher price tickets.[6] In this sense, taste heterogeneity can be seen as a necessary but not sufficient condition for SSAV.

Typical attribute values that are subject to self-selection are for examples: minutes used to drive from A to B (in-vehicle time), minutes arrived at the station/airport prior to departure

[5] We prefer to say "self selection to" instead of "self selection on" in order to stress that we describe an active process where travellers decide which travel options and which attribute values to choose, and thus self-select to. "Self-selection on attribute values" might give the misleading impression that there is a passive/latent self-selection on the attribute values.

[6] In marketing, "price discrimination" is a well-understood concept for increasing revenues. The idea is to utilize differences in willingness-to-pay of customers by offering slightly different products at different prices (e.g. first-class tickets versus economy-class tickets).The underlying idea is that heterogeneous travellers self-select to the price ranges that best fit their preferences and budget.

(waiting time), or dollars paid for fares (travel costs). This type of self-selection is similar, but not identical to sample selection bias in the "classical" sense of Heckman (1979), where persons due to their self-selection into distinct groups (employed/unemployed) influence estimation results at the second stage of estimation (i.e. wage as a function of explanatory variables).

That self-selection described here goes to continuous values is distinguishing it also from *self-selection to travel modes* (SSTM) where travellers self-select to distinct travel modes. When the choice sets in the mode choice models depend on the chosen alternative (which they typically do in pivoted SP-study) SSTM is likely to contribute to SSAV, as transport modes have in general different attribute value ranges. For instance, airplane fares are in general more expensive than bus fares, which might have the effect that original airplane users are faced with SP choice sets including more expensive alternatives. However, it is important to realize that there is also SSAV *within* a given travel mode. For instance, plane tickets come in price segments, and poorer travellers have a stronger incentive to book plane tickets early when they are cheaper. In the remainder of the paper we will discuss SSAV, and not further elaborate on how much of this self-selection can be associated with SSTM.

4.2. Resulting endogeneity in mode choice models

When modelling utility in a standard discrete choice set up (as in Eq. 2), the existence of SSAV makes it difficult to maintain the assumption of an exogenous explanatory variable *X*. Whenever the indirect utility functions do not perfectly account for personal-specific preferences, the random term, ε, will contain (some) taste heterogeneity and other unobserved factors influencing utility. Assuming that these unobserved factors ξ influence the value of attributes (by SSAV) in a systematic manner, *X* and ε become correlated, and *X* will be endogenous. Formally, this may be represented as:

(12) $$Z_{ni} = V_{ni}(\boldsymbol{\theta_{ni}}, \boldsymbol{X_{ni}}(\xi_{\boldsymbol{ni}})) + \varepsilon_{ni}(\xi_{\boldsymbol{ni}})$$

In general, endogeneity will lead to inconsistent estimators of the coefficients θ.

In some instances, the unobserved factors ξ may mathematically cancel out in discrete choice models. For example for pivoted SP-experiments, Train and Wilson (2008) show that estimating the SP-choices with a fixed coefficient logit model is consistent when the unobserved factors influencing the RP-choices can be assumed to apply additively and with same size to all alternatives in SP.[7] Then "[s]ince only differences in utility matter .... [t]he endogeneity of SP attributes takes a form that cancels out of the behavioral model" (ibid, page 196).

However, in the case of SSAV, the unobserved factors will in general not be additively separable from X, such that they do not simply cancel out. Rather they will be embedded in the attribute values by some functional relationship. The unobserved factors may then influence the form of the marginal utility functions as outlined below.

Specifying *V* in Eq. 12 in terms of a Box-Cox model as in (5), we can write:

(13) $$V_{in} = \beta_{0i} + \sum_k \beta_k \left(\frac{(X_{kni}(\xi_{ni}))^{\lambda_k - 1}}{\lambda_k}\right).$$

Then the first and the second partial derivative with respect to attribute X_k are given as:

(14) $$\frac{\partial V_{in}}{\partial X_{kni}} = \frac{\partial X_{kni}(\xi_{ni})}{\partial X_{kni}} \beta_k (X_{kni}(\xi_{ni}))^{\lambda_k - 1}$$

(15) $$\frac{\partial^2 V_{in}}{\partial {X_{kni}}^2} = \frac{\partial X_{kni}(\xi_{ni})}{\partial X_{kni}} \beta_k (\lambda_k - 1)\, (X_{kni}(\xi_{ni}))^{\lambda_k - 2} + \frac{\partial^2 X_{kni}(\xi_{ni})}{\partial {X_{kni}}^2} \beta_k\, (X_{kni}(\xi_{ni}))^{\lambda_k - 1}$$

For linear models ($\lambda_k = 1$), Eq. 15 is reduced to $\beta_k \frac{\partial^2 X_{kni}(\xi_{ni})}{\partial {X_{kni}}^2}$. Furthermore this expression for the second partial derivative will be zero whenever the functional relationship between X_{kni} and ξ_{ni} is also linear, which might be a reasonable assumption. Given this, it may be argued that SSAV has minor effects for linear-in-variable models. This corresponds with the

[7] It is argued that this may be achieved in SP by instructing the respondents to think back at the actual situation, and reconsider his/her transport choice by taking into account the listed attributes with its new values while assuming that everything which is not listed is the same for all alternatives.

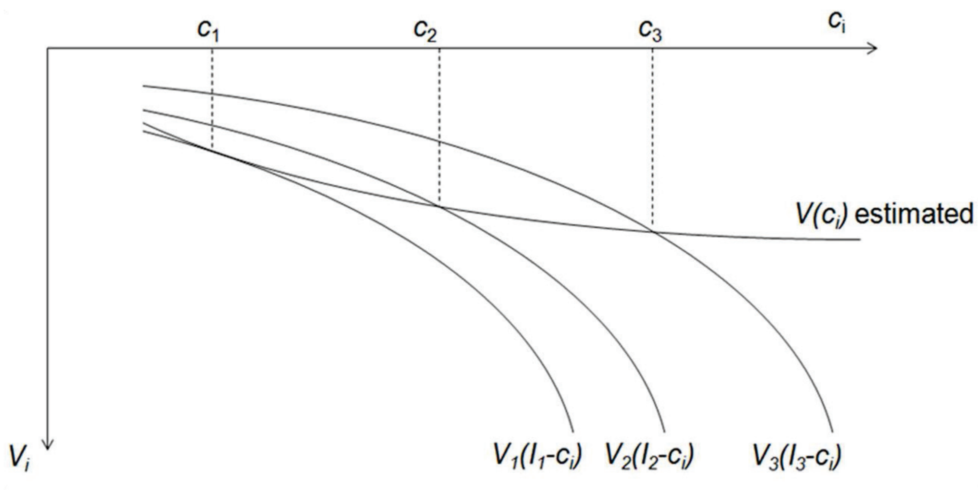
intuition that in linear models, where only differences in utility matter and marginal utilities are constant over the whole attribute range, the absolute size of attribute values (be it endogenous or not) is not important for the behavioural model. Of course, when the true functional relationship is indeed non-linear, a naive linear-in-variable model should not be used (despite that it may mask the SSAV problem).

For non-linear-in-variable models the unobserved factors are likely to have an effect on the marginal utility, and it is even possible that ignoring the impact of unobserved factors can have a decisive effect on the sign of the second partial derivates in Eq.15.

4.3. Illustration of possible effects

For an illustration of the possible effects of SSAV, suppose that there is taste heterogeneity with respect to monetary travel costs (the fare). Assume the existence of three distinct groups of homogenous persons (1: low income, 2: middle income, and 3: high income). The utility function for each group, which shows increasing dis-utility from travel costs (as suggested by theory) are depicted in Figure 2.

Figure 2: Stylized example of self-selection to attribute values



Assume that the three groups consider different alternatives involving different prices (c_1, c_2 and c_3) such that the disutility for c_1, c_2, and c_3 are only observed as the disutility of the low income, middle and high income group, respectively, and not as the disutility of the remaining groups at this price level. In more realistic settings there will be overlapping price ranges instead of distinct prices, but the argument in this stylized example applies there as well. Based on the three utility values ($V_1(I_1\text{-}c_1)$, $V_2(I_2\text{-}c_2)$, $V_3(I_3\text{-}c_3)$), the researcher would then estimate a functional form as indicated in figure 2. This utility function for the cross-sectional data implies decreasing marginal dis-utility although all single utility functions are characterized by increasing marginal dis-utility. The observation that analyses on an aggregated level can give rise to misleading conclusions compared to results drawn from analysis on the individual level is not limited to this application (see literature on ecological fallacy and the Simpson's paradox, e.g. Robinson (1950) and Blyth (1972)).

4.4. Which kind of choice data is prone to SSAV?

For longitudinal data with sufficient observations per decision maker it is in principle possible to estimate individual utility functions. In this case, self-selection does not apply, and each utility function should, according to economic theory, yield person-specific Box-Cox parameters values of one or greater. Most often, representative longitudinal data with many observations per individual (and sufficient variation in LoS) will not be available, and one has to rely on cross-sectional data to infer representative utility functions.

For studies based on cross-sectional revealed preference (RP) choice data, SSAV is always a potential challenge. It is worth noting that if aggregated LoS-data is applied, some of the person-specific self-selection will be masked in the average values. For example, when LoS-variables are aggregated over zones (as in many transport model systems), self-selection works only through the choice of residence. Then SSAV applies as persons with relative low disutility of travel time may settle (voluntarily) further away from their place of work etc. The case of SSAV is stronger when person-specific data (e.g. reported by the travellers themselves) is available. The variation in attribute values will then be much larger and more affected by the taste heterogeneity across decision makers. For example, travellers that make a lot of pauses on their car trip or that arrive much earlier than required at the airport will be likely to have lower preference for time savings in transport. The existence of such travellers (and the SSAV associated with them) would be hidden when using aggregated (average) data.

For 'pivoted' choice experiments (CEs), which have become popular and are frequently applied in transportation studies, the case of SSAV is particularly important. Here, CEs are constructed from RP choices, and at least some of the attributes vary around the values of the chosen alternative in the RP choice. This is done in order to enhance the realism of the hypothetical choice task. In pivoted choice experiments, the unobserved factors that influence the RP choice will carry over to the constructed attribute values in the SP survey. For instance, more price sensitive travellers choosing a rather cheap travel option in RP will be confronted with choice tasks in which they are offered fares from the lower range, and will not face choice tasks that involve expensive travel options.

For non-pivoted SP studies where every respondent faces attributes values within the same range (independent of the respondent's actual choice and preferences), the self-selection problem is avoided. Here the estimated utility functions can be thought of as utility functions of a representative (average) agent such that results which are consistent with theory can be expected.

4.5. Strategies to control for SSAV

SSAV can lead to serious forecasting bias. Consider again the example in Figure 2, and assume that one wants to estimate the initial ridership of a new travel mode (say HSR) for which fares are set to c_3. Applying the utility function which is estimated on the whole population for every decision maker, a moderate dis-utility from high fares is predicted. However, the actual dis-utility of group 1 and 2 at c_3 is very high, such that only a minority will choose HSR with fares at c_3. Thus the ridership will be overestimated in the forecasting model, especially if the group of rich travellers is small. Testing a more conservative price scenario (e.g. c_2) the forecasting error is likely to be considerably smaller.

Having possible forecasting biases in mind, strategies for controlling for self-selection are called for. The described self-selection problem differs somewhat from other more standard and well-understood self-selection problems in transportation, for instance the attitudinal self-selections making the causal effect between residential choices and travel behaviour difficult to disentangle (Mokhtarian and Cao 2008). Most of the methods investigated by Mokhtarian and Cao (2008) to control for (attitudinal) self-selection (e.g. direct questioning and structural equation models) do not seem applicable to our case or require data which are difficult to obtain (longitudinal designs). Also (two-stage) sample selection models seem impractical given that the self-selection is not working through distinct groups, but through overlapping attribute value ranges.

One straightforward practical attempt to (partly) account for the described self-selection problem is to statistically control for the presumed drivers behind self-selection, i.e. taste heterogeneity. By accounting for the heterogeneity (reflected in the beta-coefficient), it might be possible to retrieve expected BCT parameters. In the next section, we present a case study where we use this approach.

5. Empirical case study

5.1. Data

The empirical part of the paper is related to a study on forecasting ridership in scenarios including High-Speed Rail (HSR) in Norway (Flügel and Halse 2012). This study was conducted by the Institute of Transport Economics, and is not related to the official Norwegian assessment study of HSR (Jernbaneverket 2012, Atkins 2012). It focuses on two main long-distance corridors in Norway.[8] The respondents of the SP part of this study, which was a self-administered internet questionnaire, were recruited from an on-site RP study (a pen-and pencil questionnaire) that gathered basic information on the current choice of travel mode.[9]

The CEs in the SP part involve two travel mode alternatives: i) the respondent's current mode (air, car, train or bus as reported in the RP choice), and ii) a hypothetical HSR option. The attributes characterizing the travel modes were total travel costs, travel time (divided into in-vehicle, 'access' and 'egress' time), service frequency (number of departures per day), and the share of the ride spent in tunnels. The attribute values for the current mode are based on reported reference values. In a first sequence of eight choice tasks ('CE1'), attribute values of the current mode are fixed to the reported values. In a second sequence of six choice tasks ('CE2') attribute values are pivoted around the reported reference values using moderate percentage changes. The attribute values for HSR differ between choice tasks, and are set so as to appear realistic in a scenario where HSR exists in Norway (see Halse (2012) for details of the experimental design). Tables A1 and A2 in the appendix summarize the details of the design of the attribute values, and Table A3 provides data on sample sizes and travel mode choice behaviour observed in the RP and SP surveys.

[8] The corridors are from the capital Oslo to the two second biggest cities Bergen and Trondheim. The distance between Oslo and Bergen and Oslo and Trondheim is about the same (500 km by car). The same goes for the level of service of different modes; with the exception of bus which has a poor service between Oslo and Bergen. We therefore dismiss bus from the choice set for all connections between Oslo and Bergen.

[9] Market penetration of internet is very high in Norway (Lindhjem and Navrud 2011). The response rate for the SP-data (invited/completed questionnaire) was about 33%. The response rates of the RP-study were 74, 60 and 25 percent among rail, bus and air passengers, respectively, and 80 percent among car drivers (Denstadli and Gjerdåker 2011). Choice-based sampling in RP carries over to the SP-study. Hence, we have used external weights in our estimations in order to offset the unrepresentative sampling in SP.

5.2. RP versus SP estimation results

In this section we want to compare parameter estimation results for RP-choices (i.e. the choice between the current travel modes bus, air, train and car) with the estimates from SP-choices (i.e. the choice between one of the four current modes and HSR).

As precise LoS-variables are not accessible for all available current travel modes in RP, we had to import zonal LoS-variables from the national network model (Hamre *et al* 2002) given the reported origin-destination (OD-pair). These variables are average values for the OD-pair, and do not reflect personal-specific preferences. The average fare for an OD-pair in a given mode does for example not reflect when, and at which exact price, a ticket is bought. As discussed in 4.2, this is likely to reduce the impact of SSAV compared to the SP-choices, where attributes values are based on the reported person-specific values. We therefore expect the BCT parameters for the RP-choices to be greater than the corresponding parameters for the SP-choices.

Utility functions are specified as:

(16) $$Z_{int} = \beta_{0i} + \sum_k \beta_k * (\frac{X_{kint}^{\lambda_k} - 1}{\lambda_k}) + \eta_{ni}^{(SP)} + \varepsilon_{int}$$

where the attributes *k* are frequency, time, and cost; and alternatives *i* are car, bus, train, air, and HSR. (The latter only occurs in SP). The index *t* represents the choice task number per respondent. ε_{int} is the noise term, which is assumed to be i.i.d.-Gumbel distributed in order to yield the well-known logit model. For the SP-choices we include normally distributed person-specific error components ($\eta_{ni}^{(SP)}$) with generic standard deviation (σ) in all alternatives allowing for person-specific taste differences for travel modes. This was found to be an important feature of the econometric model in order to take into account the panel structure of the data, and to account for the choice behavior of respondents who (seemingly) made their choices without considering the particular attribute values ("Non-Traders" in Table A3). All beta (and lambda) coefficients (except the alternative specific constants β_0) are modelled as generic over travel modes. This is likely to be restrictive for the travel time attribute, as the disutility of travel time is likely to depend on the comfort level of the particular travel mode (see discussion later). The analysis is performed for leisure trips, which is the largest segment in our sample. The estimation was run with BIOGEME (Bierlaire 2003, 2008). 1000 Halton draws are used to simulate the random terms.

Table 1: Revealed Preference (RP) and Stated Preference (SP) estimation results

	RP-choice based on zonal LoS variables				**SP-choice based on pivoted design**			
Choice set	All existing travel modes (not HSR)[a]				The current mode and HSR			
Model	Multinomial logit model				Error component logit for panel data			
BCT-parameter	λ=1		λ free		λ=1		λ free	
Model Index	RP1		RP2		SP1		SP2	
	Value	robust T-value (0)	Value	robust T-value (0)	Value	robust T-value (0) [b]	Value	robust T-value (0)[b]
β_0 (Bus)	-0.278	-3.51	-0.899	-7.39	0.442	0.77	-0.471	-0.47
β_0 (Air)	-0.447	-2.06	-0.207	-0.14	-3.06	-4.65	-3.17	-2.68
β_0 (Train)	-0.533	-14.65	-1.52	-10.68	-1.57	-4.48	-2.17	-1.97
β_0 (HSR)					-0.812	-1.92	-1.39	-1.65
$\beta_{FREQUENCY}$ (Gen)	0.0253	6.36	-1.08	7.85[c]	0.0686	4.80	2.08	4.81[c]
β_{COST} (Generic)	-0.00393	-28.03	-0.00017	-27.41[c]	-0.0069	-7.69	-0.300	-11.58[c]
β_{TIME} (Generic)	-0.00697	-21.46	-0.053	-20.80[c]	-0.0118	-12.23	-0.912	-12.22[c]
σ (Generic)					2.84	10.35	3.26	8.59
	Value	robust T-value (1)	Value	robust T-value (1)	Value	robust T-value (1) [b]	Value	robust T-value (1) [b]
$\lambda_{FREQUENCY}$ (Gen)	1	fixed	-0.553	-5.41	1	fixed	-0.769	-4.96
λ_{COST} (Generic)	1	fixed	1.51	1.57	1	fixed	0.443	-7.63
λ_{TIME} (Generic)	1	fixed	0.675	-2.50	1	fixed	0.270	-4.29
# parameters		6		9		8		11
# obser-vations		8100		8100		8402		8402
# respond-ents		8100		8100		607		607
Null LL		-9268.442		-9268.442		-5822.436		-5822.436
Final LL		-6166.302		-6099.557		-2965.661		-2877.706
Adj. Rho-square		0.334		0.341		0.489		0.504

a) Car available is in choice set when household owns car; bus available only for corridor Oslo-Trondheim.
b) Robust T-values take into account the panel structure. c) Conditioned on the estimated lambda values. In

BIOGEME these values are obtained by running one additional iteration where power terms are fixed to the estimated value.

For both the RP and the SP-choices, the BCT-models (RP2, SP2) clearly fit the data better than their linear counterparts (RP1, SP1). All Box-Cox parameters are statistically different from unity with the power parameter for cost in RP as the only exception. The Box-Cox parameter for frequency is about -0.55 in RP2 and -0.77 in SP2. Decreasing marginal utility for the number of departures per day is consistent with intuition and theory (seeing frequency as an economic good). For travel time, decreasing marginal dis-utility is estimated, contradicting suggestions from theory while being consistent with most empirical intercity mode choice studies. The power parameter in RP (0.675) is higher than in SP (0.270), which is line with our previous discussion. While the power parameter for cost in RP is 1.51, and thus consistent with theoretical expectations, a value of 0.443 is obtained in SP. Following the reasoning of this paper, this value is likely to be affected by SSAV driven by the taste heterogeneity across respondents. In the next section we show that this value can in our case be increased to 1.17 when controlling for taste heterogeneity.

5.3. Controlling for taste heterogeneity in SP

Before performing Box-Cox mixed logit models (Orro *et al.* 2005) to statistically control for *unobserved* taste heterogeneity, we attempted to account for *observed* taste heterogeneity by segmenting the cost coefficient by (1) income groups and (2) groups constructed based on the "stated price sensitivity"[10]. Table A4 (in the appendix) shows that there is indeed

[10] Respondents stating, after the choice experiments, that they considered mainly or only the price attribute when making their SP-choice, and in addition had either low income or a reported total cost of the reference trip of less than 400 NOK, are labelled "most price sensitive". Respondents labelled "less price sensitive" were those stating that they considered the price attribute a "little" or "not at all"; and in addition had either high income, a reported total cost above 800 NOK or did not pay for the reference trip themselves. All remaining respondents are allocated to a group labelled "moderately price sensitive".

SSAV as rich (and less price sensitive) respondents tend to choose more expensive travel modes, and for a given travel mode the more expensive options. The estimated value of λ_{COST} (using the same model as in SP2, just adding the segmentation of β_{COST}) increases from 0.443 (in SP2) to 0.502 with the income segmentation and to 0.512 with the "stated price sensitivity" segmentation (Table A5 in the appendix). An increase in the estimates is expected as some of the taste heterogeneity is controlled for. However, the values are still significantly lower than unity, indicating that there might be considerable taste heterogeneity within the segmented groups. Hence, the simple segmentation (here in three groups) was not sufficient to retrieve the expected value of the BCT parameter for the travel cost attribute.

Next, we specify a Box-Cox mixed logit model where marginal utilities are assumed to be log-normally distributed over respondents.

$$(17) \qquad Z_{int} = \beta_{0i} - e^{\beta_{time,n}}\left(\frac{X_{time,int}^{\lambda_{time}}-1}{\lambda_{time}}\right) - e^{\beta_{cost,n}}\left(\frac{X_{cost,int}^{\lambda_{cost}}-1}{\lambda_{cost}}\right)$$

$$+e^{\beta_{freq,n}}\left(\frac{X_{freq,int}^{\lambda_{freq}}-1}{\lambda_{freq}}\right) + \eta_n^{(SP)} + \varepsilon_{int}$$

with

$$\beta_{time,n} \sim N(\mu_{time}, \sigma_{time}), \beta_{cost,n} \sim N(\mu_{cost}, \sigma_{cost}), \beta_{freq,n} \sim N(\mu_{freq}, \sigma_{freq}), \eta_n^{(SP)} \sim N(0, \sigma)$$

This implies that we restrict the marginal impact of LoS-variables to have their theoretically expected sign. The average value of $e^{\beta_{k,n}}$ is obtained as $e^{\mu_k+\sigma_k^2/2}$.

Table 2 reports the estimation results for a model version with λ=1 (SP3), and a model version where we allow for non-linearity-in variables (SP4).

Table 2: SP estimation results with log-normal distributed β-coefficients

BCT-parameter	λ=1		λ free	
Model Index	SP3		SP4	
	Value	robust T-value (0) [a]	Value	robust T-value (0)[a]
β_0 (Bus)	-2.65	-1.51	-4.19	-4.13
β_0 (Air)	-3.42	-3.25	-5.53	-4.96
β_0 (Train)	-1.45	-1.63	-2.93	-4.21
β_0 (HSR)	0.444	0.51	-1.39	-2.26
$\mu_{FREQUENCY}$ (Gen)	-5.74	-1.46	-0.556	-2.13[b]
$\sigma_{FREQUENCY}$ (Gen)	2.9	1.93	1.89	16.15[b]
μ_{COST} (Generic)	-4.52	-34.37	-5.58	-51.81[b]
σ_{COST} (Gen)	0.997	11.11	1.04	14.51[b]
μ_{TIME} (Generic)	-4.29	-13.89	0.466	4.05[b]
σ_{TIME} (Gen)	1.13	9.64	1.07	24.61[b]
σ (Generic)	2.61	7.93	2.54	10.76[b]
	Value	robust T-value (1)	Value	robust T-value (1)
$\lambda_{FREQUENCY}$ (Gen)	1	fixed	-0.650	-8.73
λ_{COST} (Generic)	1	fixed	1.170	1.24
λ_{TIME} (Generic)	1	fixed	0.190	-5.00
# parameters		11		14
# observation		8402		8402
# respondents		607		607
Null LL		-5822.436		-5822.436
Final LL		-2605.147		-2569.757
Adj. Rho-square		0.551		0.556

a) Robust T-values take into account the panel structure. b) Conditioned on the estimated lambda values

In SP4, the value of λ_{COST} is estimated at 1.17, and thereby indicates increasing marginal disutilities from travel cost as suggested by economic theory. A value close to (but not below) unity can be expected, considering that travel costs in Norway constitute a relatively small fraction of disposable income. The estimated value of 1.17 is not significantly different from unity (its 95% confidence interval is [0.90; 1.44]), but significantly different from 0.443, the value estimated in model SP2 (Table 1). This indicates that we have offset some or all of the impact of SSAV for that attribute.

While expected power parameters are retrieved for travel cost, the power parameter for travel time has not changed much. Travel time preferences seem much more complex than travel cost preferences partly because the former are likely to be travel mode specific. We tried to model beta- and lambda parameters as alternative-specific, but this could not be utilized in this paper because of identification problems. Note, that such BCT-mixed logit models are highly non-linear, which seems to put practical limits to the degree of coefficient segmentation.

6. Conclusions and directions of further research

The contribution of the paper to the literature is the introduction of *self-selection to attribute values* (SSAV) – i.e. the endogenous allocation of individuals to attribute values – as one explanation for decreasing marginal dis-utility of travel time and cost in mode choice models. These results contradict micro-economic theory, but this seems largely ignored in the transportation literature. To our knowledge, this is the first paper that explicitly discusses and warns against the possible effects of SSAV on the estimation of functional forms. We also suggest that controlling for unobserved taste heterogeneity in the model can retrieve marginal utility function more in line with theoretical expectation. Thus, the paper also adds to the literature on the interdependence between taste heterogeneity and non-linearity (Orro et al. 2005, Pinjari and Bhat 2006).

SSAV is a possible challenge not only for travel mode choice modelling, but for many types of choice modelling (with underlying non-linear utility functions) independent of the research area.

SSAV is likely to be more severe in cross-sectional data with person-specific input data. This is supported by our analyses comparing estimation results of RP-choices based on zonal data with SP-choices based on reported (person-specific) data.

Controlling for taste differences by simple sample segmentations was found to be insufficient to account for SSAV, but we obtained some evidence that random parameter models controlling for unobserved taste heterogeneity might be capable of retrieving Box-Cox Transformation (BCT)-parameters which are more consistent with suggestions from economic theory. However, our empirical case study only gives some indications of the effect of SSAV, and the capability of controlling for SSAV by random coefficient models. Thus, more empirical tests on different data sets, and more research on the interdependency of taste heterogeneity and non-linearity (initiated by Orro *et al*. 2005) are called for.

We warn against relying on non-linear utility specifications without investigating possible effects of self-selection. This concerns both BCTs and specifications that predefine decreasing marginal utility such as log- or square-transformations. Such specifications might fit cross-section data better than linear models, but they might do a poor job in predicting individual choice behaviour given that utility functions are indeed as predicted by theory. This is likely to be a particular problem when testing scenarios that involve large changes, e.g. scenarios of very expensive fares or substantial time reductions, which is often the case in predictions of future HSR-ridership. A favourable model accounts for taste heterogeneity and applies linear or increasing (dis-)utilities for LoS-variables. We note that many mode choice models implicitly address this issue by the use of interaction effects. For instance, dividing the cost attribute by income accounts for (some of) the systematic taste heterogeneity, and may be seen as a good practical approach to linearizing the effect of explanatory variables.

Given the theoretical considerations, and the evidences from our study, we *hypothesize* that:

(1) BCT-parameters for LoS-variables estimated for a complete cross-sectional data set using a model without interaction effects will be lower the greater SSAV is. And SSAV will be greater (i) the higher the taste heterogeneity is in the sample (ii) the greater the variation in characteristics (attribute values) of the travel modes respondents can choose between (self-select to)[11], and (iii) for pivoted SP-studies, the less the designed attributes values vary around the RP-values.[12]

[11] Obviously, when all transport modes offer identical LoS variables to all travelers, SSAV is not possible (even if travelers have heterogeneous preferences). To the extent that urban transport systems are more homogenous than intercity transport systems, the reduced possibility of SSAV might explain why many urban studies find increasing disutility of travel time.
[12] If this hypothesis is confirmed, there seems to be a trade-off between increased realism from pivoted designs (letting design attribute value remain close to actual values) and the severity of the self-selection problems.

(2) BCT-parameters will increase and approximate their true values the better taste heterogeneity is controlled for.

(3) The forecasting bias (when applying BCT-parameters without controlling for taste heterogeneity) will increase with: (i) increasing SSAV, and (ii) increases in the assumed attribute values of future transport options.

To test these hypotheses empirically, one needs data in which these features vary in a controlled manner. One possibility would be to use synthetic choice data.

To assess the effect of SSAV on real-data, studies including choice tasks which are both pivoted and non-pivoted are called for. SP-studies with more observations per traveller, making it possible to estimate individual utility functions, would also be very interesting in this context.

Acknowledgement

The paper uses data of the HSR-related part of the TEMPO-project. We would like to thank all persons that have contributed to the TEMPO project, and the Research Council of Norway for financing it. The usual disclaimer applies.

References:

Algers, S., Gaudry, M., 1994. Are Nonlinearity and Segmentation Substitutes, Complements, or Unrelated in Logit Models? The Swedish Intercity Passenger Case, Publication AJD-28 Agora Jules Dupuit Université de Montréal.
Atkins, 2012. Model Development Report, Norway High Speed Rail Assessment Study: Phase III. http://www.jernbaneverket.no/PageFiles/17376/Norway%20HSR%20Phase%20III_Model%20Development_Final%20Report_Atkins_25-01-12.pdf (27.3.2013)
Bierlaire, M., 2003. BIOGEME: A free package for the estimation of discrete choice models, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.
Bierlaire, M., 2008. An introduction to BIOGEME Version 1.8. http://biogeme.epfl.ch/v18/tutorialv18.pdf (27.3.2013)
Blayac, T., Causse, A., 1998. Vers une meilleure appréciation de la valeur subjective du temps: légitimité des transformations Box-Cox, Working Paper LAMETA. University of Montpellier, France.
Blayac, T., Causse, A., 2001. Value of travel time: a theoretical legitimization of some nonlinear representative utility in discrete choice models. Transportation Research Part B 35, 391-400.

Blayac, T., Causse, A., 2003. Value of travel time: a theoretical legitimization of some Box-Cox transformations in discrete choice models, 20èmes Journées de Microéconomie Appliquée 5 et 6 juin 2003 à Montpellier.
Blyth, C. R., 1972 On Simpson's Paradox and the Sure-Thing Principle. *Journal of the American Statistical Association,* 67, 364-66.

Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformation. Journal of Royal Statstical Society. Series B (Methodological) 26, 211-252.
Denstadli, J.M., Gjerdåker, A., 2011. Transportmiddelbruk og konkurranseflater i tre hovedkorridorer. Institute of Transport Economics, Oslo.
Flügel, S., Halse, A.H., 2012. Non-Linear Utility Specifications in Mode Choice Models for High-Speed Rail, Kuhmo Nectar Conference 2012 Berlin.
Flügel, S., Hjorth, K., Ramjerdi, F., 2011. Between-mode differences in value of travel time estimates: Evidence from Norway, Kuhmo-Nectar Conferance Stockholm.
Fosgerau, M., Hjorth, K., Lyk-Jensen, S.V., 2010. Between-mode-differences in the value of travel time: Self-selection or strategic behaviour? Transportation Research Part D 15, 370-381.
Fridstrøm, L., Madslien, A., 2002. A Stated Preference Analysis of Wholesalers' Freight Choice, in: R, D. (Ed.), Demanda di Trasporto Merci e Preferenze Dichiarate – Freight Transport Demand and Stated Preference Experiments. Franco Angeli, Milano, pp. 223-250.
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. Université de Montréal.
Gaudry, M., Heinitz, F., Last, J., Mandel, B., 1998. Methodological Developments within the Quasi-Direct Format Demand Structure: the Multicountry Application for Passengers MAP-1, Publication CRT-98-46. Centre de recherche sur les transports, Université de Montréal.
Gaudry, M., Le Leyzour, A., 1994. Improving a Fragile Linear Logit Model Specified for High Speed Rail Demand Analysis in the Quebec-Windsor Corridor of Canada. Université de Montréal.
Gaudry, M.J.I., Wills, M.J., 1978. Estimating the functional form of travel demand models. Transportation Research 12, 257-289.
Halse, A.H., 2012. Demand for high speed rail in Norway: Survey design and data collection, working paper. Institute of Transport Economics, Oslo.
Hamre, T.N., Rekdal, J., Larsen, O.I., 2002. Utvikling av den nasjonale persontransportmodellen i fase 5. Del B: Estimering av modeller, Institute of Transport Economics, Oslo.
Heckman, J.J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153-161.
Jara-Díaz, S.R., 2007. Transport Economic Theory. Emerald Group Publishing Limited.
Jara-Díaz, S.R., Farah, M., 1987. Transport demand and users' benefits with fixed income: The goods/leisure trade off revisited. Transportation Research Part B 21, 165-170.
Jernbaneverket, 2012. Høyhastighetsutredningen 2010-2012 Konklusjoner og oppsummering av arbeidet i Fase 3- Del1. http://www.jernbaneverket.no/PageFiles/17299/Rapport_Del_1.pdf (27.03.2013)
Jiang, S., 2010. Modelling demand for High Speed Rail in Sweden - Business Trips. KTH, Stockholm.
Koppelman, F.S., 1981. Non-linear utility functions in models of travel choice behavior. Transportation 10, 127-146.
Laferrière, R., 1993. Les effets de seuil un train à haute vitesse dans le Corridor Québec-Windsor, Sixièmes Entretiens du Centre Jacques Cartier. Publication AJD-20, Agora Jules Dupuit, Université de Montréal, Lyon, p. 14
Lindhjem, H., S. Navrud 2011: Using Internet in Stated Preference Surveys: A review and

comparison of survey modes. *International Review of Environmental and Resource Economics*, 5 (4); 309–351

Mabit, S.L., 2008. Sample selection and taste correlation in discrete choice transport modelling, DTU Transport. DTU, Copenhagen.

Mandel, B., Gaudry, M., Rothengatter, W., 1997. A disaggregate Box-Cox Logit mode choice model of intercity passenger travel in Germany and its implications for high-speed rail demand forecasts. The Annals of Regional Science 31, 99-120.

Mas-Colell, A., Whinston, M.D., Green, J.G., 1995. Microeconomic theory Oxford University Press, Inc, Oxford, New York.

McFadden, D., United States. Dept. of, T., Seminar on Emerging Transportation Planning, M., 1976. The theory and practice of disaggregate demand forecasting for various modes of urban transportation. Dept. of Transportation, Office of the Secretary, Washington.

Mokhtarian, P.L., Cao, X., 2008. Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. Transportation Research Part B 42, 204-228.

Orro, A., Novales, M., Benitez, F.G., 2005. Nonlinearity and Taste Heterogeneity Influence on Discrete Choice Model Forecast., Association for European Transport and contributors.

Pinjari, A. Bhat C. 2006. Nonlinearity of Response to Level-of-Service Variables in Travel Mode Choice Models. Transportation Research Record: Journal of the Transportation Research Board. 1977, 67-74.

Robinson, W. S. 1950. Ecological Correlations and the Behavior of Individuals. *American Sociological Review,* 15**,** 351-357.

Train, K., McFadden, D., 1978. The goods/leisure tradeoff and disaggregate work trip mode choice models. Transportation Research 12, 349-353.

Train, K., Wilson, W.W., 2008. Estimation on stated-preference experiments constructed from revealed-preference choices. Transportation Research Part B 42, 191-203.

Wang, Q., 2011. Travel demand forecasting with stated choice data: Swedish domestic long-distance trips. KTH, Stockholm.

Appendix

Table A1. Attribute levels in CE1

Level / Attribute		1	2	3
Total costs, reference alternative		Reported Base value	--	--
In-vehicle time, reference alternative		Reported Base value	--	--
Access time, reference alternative		Reported Base value	--	--
Egress time, reference alternative		Reported Base value	--	--
Frequency, reference alternative		Reported Base value	--	--
Tunnel share, reference alternative		Constructed Base value	--	--
Costs, HSR	High cost	-30 %	Constructed Base value	+ 30 %
	Medium cost	-30 %	Constructed Base value	+ 30 %
	Low cost	-30 %	Constructed Base value	+ 30 %
In-vehicle time, HSR		- 50 min.	Constructed Base value	+ 50 min.
Access time, HSR		Reported Base value	--	--
Egress time, HSR		Reported Base value	--	--
Frequency, HSR (departures/day)		6	10	14
Tunnel share, HSR		-33 %	Constructed Base value	+ 33 %

Table A2. Attribute levels in CE2

Level / Attribute		1	2	3
Total costs, reference alternative		-30 %	Reported Base value	+ 30 %
In-vehicle time, reference alternative	Air	Reported Base value	+ 10 %	--
	Car, train, bus	-25 %	Reported Base value	--
Access time, reference alternative[a]	Short access trip	-33 %	Reported Base value	--
	Long access trip	-25 %	Reported Base value	--
Egress time, reference alternative[a]	Short egress trip	-33 %	Reported Base value	--
	Long access trip	-25 %	Reported Base value	--
Frequency, reference alternative	Air	-50 %	Reported Base value	--
	Train, bus	Reported Base value	+ 50 %	--
Tunnel share, reference alternative		Constructed Base value	+ 50 %	--
Costs, HSR	High cost	-30 %	Constructed Base value	+ 30 %
	Medium cost	-30 %	Constructed Base value	+ 30 %
	Low cost	-30 %	Constructed Base value	+ 30 %
In-vehicle time, HSR (compared to constructed base level in CE1)		- 40 min.	+ 40 min.	--
Access time, HSR[b]	Short access trip	-33 %	+ 33 %	--
	Long access trip	-25 %	+ 25 %	--
Egress time, HSR[b]	Short egress trip	-33 %	+ 33 %	--
	Long access trip	-25 %	+ 25 %	--
Frequency, HSR (departures/day)		6	14	--
Tunnel share, HSR	Bus passengers	-33 %	+ 11 %	--
	Other segments	-25 %	+ 25 %	--

[a] The respondent was placed in the “long access egress trip” segment if he/she travelled by air and in the “short access egress trip” segment if he/she travelled by train or bus. [b] The respondent was placed in the “long access egress trip” segment if he/she travelled by car or if he/she travelled by air and the trip was not work-related. Train and bus passengers and those travelling by air for work-related purposes were placed in the “short access/egress trip” segment.

Table A3: sample size and general choice behavior

Sample Size and General choice behaviour		Group size RP	Group size SP	Percent of SP choices (%)			Percent of respondents always choosing in SP ("Non-Traders") (%)		
Purpose	**RP-choice**	**Final sample**	**Final sample**	**Current mode**	**HSR**	**Opt-out**	**Current mode**	**HSR**	**Switch between**
Leisure	**car**	3833	320	61.6	37.7	0.6	31.3	13.8	55.0
	air	920	76	34.5	64.7	0.8	1.3	31.6	67.1
	train	2867	176	40.8	57.2	2.0	6.3	18.2	75.6
	bus	480	35	41.0	58.5	0.6	0.0	17.1	82.9
Work-related travel	**car**	698	46	56.7	41.7	1.6	23.9	15.2	60.9
	air	1051	132	41.6	57.2	1.2	4.5	22.0	73.5
	train	482	31	44.9	54.4	0.6	9.7	19.4	71.0

Table A4: Self-selection to travel cost of reference trip in SP-data

Income group	chosen mode	Reported total cost of reference trips Count	Mean (NOK)	S.E.M. (NOK)	Assigned price sensitivt group	chosen mode	Reported total cost of reference trips Count	Mean (NOK)	S.E.M. (NOK)
low income (under 350.000 NOK)	car	85	515.40	20.08	**high price sensitivity**	car	65	462.12	22.35
	air	27	819.59	73.04		air	18	867.83	105.78
	train	86	440.26	19.53		train	76	399.57	22.19
	bus	26	336.38	18.58		bus	24	348.79	18.88
	all modes	224	**502.44**	**16.39**		**all modes**	183	**461.19**	**19.04**
middel income	car	110	577.45	22.10	**moderate price sensitivity**	car	173	571.22	17.90
	air	21	875.05	49.51		air	31	821.87	42.77
	train	53	470.89	28.28		train	71	515.82	23.85
	bus	7	426.71	34.59		bus	11	403.73	38.61
	all modes	191	**575.08**	**18.00**		**all modes**	286	**578.19**	**14.31**
high income (over 500.000 NOK)	car	123	647.36	29.48	**low price sensitivity**	car	82	717.32	36.30
	air	27	835.15	56.61		air	27	839.26	52.45
	train	36	507.14	52.81		train	29	509.86	55.26
	bus	2	539.50	0.50		bus	0	.	.
	all modes	188	**646.33**	**24.14**		**all modes**	138	**697.58**	**27.94**

Table A5: SP estimation results with segmentation after income groups and price sensitivity

segmentation	**Simple income groups** G1=low income (36.8%) G2= middle income (32.1%)[a] G3= high income (31.1%)		**Price sensitivity groups** G1=high price sensitivity (31.2%) G2= moderate price sensitivity (45.8%) G3= low price sensitivity (23.0%)	
	Value	robust T-value (0)[b]	Value	robust T-value (0)[b]
β_0 (Bus)	-0.832	-0.66	-1.79	-1.55
β_0 (Air)	-3.37	-0.31	-4.12	-3.96
β_0 (Train)	-2.12	-1.73	-3.12	-3.23
β_0 (HSR)	-1.53	-0.89	-1.77	-2.03
$\beta_{FREQUENCY}$ (Gen)	2.18	4.20[c]	2.33	4.83[c]
β_{COST} (G1)	-0.247	-7.59[c]	-0.265	-8.99[c]
β_{COST} (G2)	-0.171	-7.72[c]	-0.166	-7.78[c]
β_{COST} (G3)	-0.167	-8.88[c]	-0.0607	-4.76[c]
β_{TIME} (Generic)	-1.02	-8.88[c]	-1.58	-12.49[c]
σ (Generic)	3.23	1.07	3.01	16.07
	Value	robust T-value (1)[b]	Value	robust T-value (1)[b]
$\lambda_{FREQUENCY}$ (Gen)	-0.795	-2.01	-0.823	-5.21
λ_{COST} (Generic)	0.502	-1.37	0.512	-3.71
λ_{TIME} (Generic)	0.256	-4.48	0.186	-5.12
# parameters		13		13
# observation		8402		8402
# respondents		607		607
Null LL		-5822.436		-5822.436
Final LL		-2851.001		-2738.419
Adj. Rho-square		0.508		0.527

a) Includes 4 respondents with missing income information b) Robust T-values take into account the panel structure. c) Conditioned on lambda values

Essay 2

Methodological challenges in modelling the choice of mode for a new travel alternative using binary stated choice data - the case of high speed rail in Norway

Stefan Flügel[a], Askill H. Halse[b], Juan de Dios Ortúzar[c] and Luis I. Rizzi[c]

[a] *Department of Economics and Resource Management, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway*

[b] *Institute of Transport Economics (TOI), Gaustadalleen 21, NO-0349 Oslo, Norway*

[c] *Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Casilla 306, Cod. 105, Correo 22, Santiago, Chile*

Keywords: High-speed rail, stated choice, group scale parameters, revealed choice, cross-nested logit model

Abstract:

Binary stated choices between traveller's current travel mode and a not-yet-existing mode might be used to build a forecasting model with all (current and future) travel alternatives. One challenge with this approach is the identification of the most appropriate inter-alternative error structure of the forecasting model.

By critically assessing the practise of translating estimated group scale parameters into nest parameters, we illustrate the inherent limitations of such binary choice data. To overcome some of the problems, we use information from both stated and revealed choice data and propose a model with a cross-nested logit specification, which is estimated on the pooled data set.

1. Introduction

A large-scale study on the feasibility and social benefits of high-speed rail (HSR) in Norway was recently carried out (Jernbaneverket 2012). The estimated market potential of HSR is naturally a crucial element in this quest, as the predicted ridership has a direct effect on expected revenues, user benefits and greenhouse gas reductions. The demand forecasting model (Atkins 2012) was based on a stated choice (SC) study where respondents faced customized surveys based on their current mode choice (revealed choice, RC). The survey included binary choice experiments (CE) between the respondents' current modes and a new HSR alternative (Figure 1 shows a schematic illustration). A similar approach was used in an independent market study conducted by the Institute of Transport Economics, TØI (Flügel and Halse 2012).

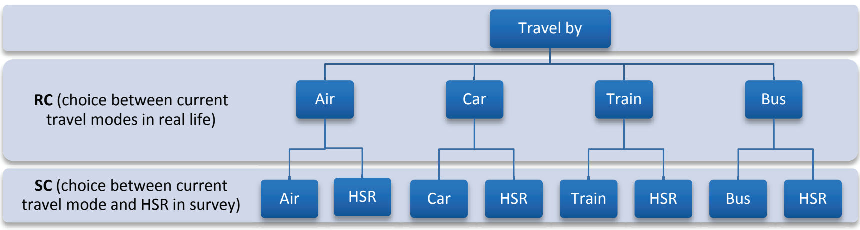


Figure 1. Decision structure in recent Norwegian HSR-studies

The main advantage of binary CE (instead of CE with a full choice set) is the simplification of the respondent's choice task. In a travel mode choice context, CE often entail a rather high degree of complexity because of the large number of attributes typically required to characterise each alternative. Lowering the overall number of attributes is likely to increase respondents' ability to choose between alternatives (Caussade *et al* 2005). In a pivot design, where respondents are typically instructed to recall the last trip they made, it is quite natural to discard the rejected travel alternatives letting the respondent focus on the current travel mode and the hypothetical new alternative.

However, while it is desirable to reduce the respondent's choice set from an experimental design point of view (in our case: providing personal specific choice sets consisting of respondent's current mode and HSR), one would like to build a forecasting model that allows considering the whole future choice set and which applies to all future decision makers, independent of their chosen mode at the time the CE were conducted. This applies, in particular, to HSR implementation scenarios that usually involve long-term predictions. Changes in many level-of-service (LoS) variables of, potentially, all travel modes are possible not only because of the long time horizon, but also because a HSR implementation is likely to affect the competitive structure of the whole travel market. Therefore, it seems unduly restrictive to limit choice sets and to condition model parameters for choice predictions in the forecasting year (e.g. in 2024, the earliest possible year for a HSR-implementation in Norway) on current RC choices (data from year 2010 in our case). Consequently, a model with a generic choice set and utility functions, independent of the original self-selection of travellers to travel modes is necessary.

Of course, aiming for a generic forecasting model based on binary stated choices (with only one alternative, HSR, being part of every respondent's choice set) is not optimal, as it does not allow considering directly how current car users, say, react to the LoS of other current modes (air, bus and traditional train). When specifying transport specific coefficients in the utility function, one needs to assume that, for example, the current car user's marginal utility (MU) of in-vehicle-time (IVT) by car is representative of everyone's MU for IVT by car.

However, possible challenges in finding an appropriate deterministic utility function are not the focus of this paper; moreover, we will assume - unless specified differently - that we can find deterministic utility functions (up to a scale parameter) that fit all *user groups* (defined on the basis current mode choice) "equally well".

For estimation, the different binary choice datasets are typically merged and a mode choice model with a common set of coefficients for HSR is estimated. In this procedure, different scale parameters (so called *group scale* parameters), that are inversely proportional to the error variances associated with each experiment, ought to be estimated to account for the fact that they might actually differ (Louviere *et al* 2000).

While the group scale parameters facilitate the estimation of a common deterministic utility function based on user-specific binary choices, it is not obvious how these parameters may be carried over to a forecasting model with a full choice set. In particular, setting up a nested logit (NL) model by naively treating group scale parameters as structural (nest) parameters, as done by Atkins (2012) in the official assessment study for HSR in Norway, involves several pitfalls:

(i) The group scale parameters only reflect the relative utility scale in choices between the different binary choice tasks (i.e. HSR versus one of the current modes) but not the utility scale difference between existing travel modes. In most cases, this means that the scale at the upper level of the nesting structure and the correlation structure among current modes has to be assumed implicitly (see sections 3.1. and 3.3); we will discuss how RC data between current modes might be utilized here (see section 4).

(ii) In many instances a NL model might not be flexible enough to account for the correlation structure suggested by the various group scale parameters. We propose the cross-nested logit (CNL) model as a more flexible structure for this purpose.

(iii) The group scale parameters do not only reflect “similarity” of transport modes, (i.e. the degree to which two or more alternatives share unobserved features, which is the classical interpretation of nest parameters, see Ortúzar and Willumsen, 2011, section 7.4.2). They might also include other error sources – in particular unobserved taste heterogeneity – that are associated with characteristics of the user groups rather than of the modes. We will discuss this in more detail in section 3.2., and using an empirical example, we will also show that results change after accounting for unobserved taste heterogeneity with random coefficients models (section 3.4.).

As the paper is mostly concerned with error variance differences (utility scale differences) between various user groups, travel modes and datasets, it is important to stress that the error term is, as usual, conditioned by the specification of the deterministic part of utility (i.e. the selection of explanatory variables and their functional form). For instance, when talking about correlation (or “similarity”) of travel modes, we are always relating to those parts of the utility function that are not accounted for by the explanatory variables. Indeed, correlation patterns in the error term are nothing desirable in itself and one would ideally strive for a multinomial logit (MNL) model by including all the variables that might explain correlation among travel alternatives. However, this is often not possible in practise (some variables are unobservable, others are just too expensive to collect). Thus, a question often asked to the researcher refers to the most appropriate correlation (nesting) structure in the forecasting model.

The particular contribution of this paper is an in-depth discussion about the most appropriate (alternative specific) correlation structure in a case where the deterministic component of the utility functions and the scale parameters are estimated from (user group specific) binary SC data. While obviously the (relative) size of the scale parameters will depend on the chosen deterministic utility function, most of the discussion in this paper can be made at a general level without being specific about the chosen deterministic utility functions.

2. Alternative Model Forms

2.1. Multinomial logit (MNL) model

We describe a standard discrete choice set up where traveller *n* chooses between different transport modes *i* belonging to a (personal specific) choice set C_n, according to the following choice rule:

$$(1) \qquad U_{in} = V_{in} + \varepsilon_{in} > U_{jn} = V_{jn} + \varepsilon_{jn} \qquad for\ all\ j \in C_n .$$

where the deterministic component of utility V_{in} is a function of attributes $\underline{X}$ and a set of parameters $\underline{\beta}$ to be estimated; in a MNL model, the random terms ε_{in} are assumed to distribute IID-Gumbel with mean zero and variance given by: $\sigma^2 = \frac{\pi^2}{\lambda^2 6}$ where $\lambda > 0$ is the scale parameter of the distribution. With this, the MNL model choice probabilities are given by (McFadden 1974, Ortúzar and Willumsen, 2011, Chapter 7):

$$P_{in} = Pr\left(V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn}\right) = \frac{e^{\lambda V_{in}}}{\sum_{j \in C_n} e^{\lambda V_{jn}}} \quad (2)$$

The IID-assumption in equation (2) implies proportional substitution patterns across alternatives as the utility of any two alternatives is uncorrelated. Note that the scale factor λ cannot be estimated separately from the parameters $\underline{\beta}$ in the deterministic component of utility, so it has to be normalized (see the discussion on identifiability by Walker 2001). Also note that the overall scale of utility *U,* normally abbreviated with *μ*, is arbitrary, i.e., when all utilities U_{in} in (1) are multiplied by a positive scalar *μ* the choice probabilities in any discrete choice model, including MNL, do not change.

2.2. Heteroskedastic logit (HL) model

The analyst might wish to allow for different error variances for different subgroups in the data. For this, the scale factor (which is inversely proportional to the error variance) can be assumed to be non-generic allowing for different *group scale* parameters, υ_{g_n}; in this case the choice probabilities of this HL[1] model become:

$$P_{in} = \frac{e^{\upsilon_{g_n} V_{in}}}{\sum_{j \in C_{g_n}} e^{\upsilon_{g_n} V_{jn}}} \; for\ all\ i \in C_{g_n} \quad (3)$$

Note that not all group scale parameters can be estimated simultaneously. For identification, one of them has to be fixed (typically at the value 1). The group scale parameters affect the resulting choice probabilities as illustrated in Figure 2.

[1] Note that this model is not identical to the "heteroscedastic extreme value model" (Bhat 1995), which is called "Heteroskedastic Logit" by Train (2009, page 92), as there the error variance of each alternative varies. In the HL model described here, the error variance varies for every subgroup but is the same for all alternatives in a subgroup.

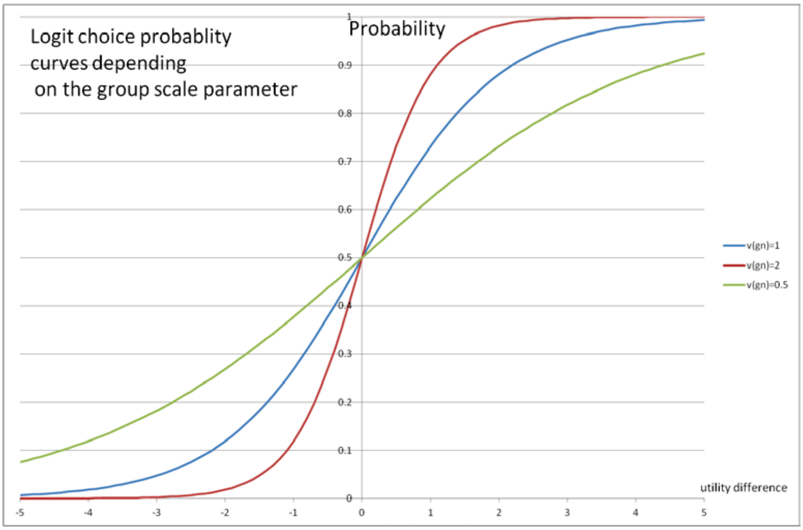


Figure 2: Illustration of effect of group scale parameter on choice probability

If the group scale parameters are different in value, the IID-assumption only applies within the subgroup but is relaxed for a joint sample that combines choice data from different user groups.

2.3. Nested logit (NL) model

In a NL model, alternatives are allocated into non-overlapping nests, *m*, that contain alternatives *i* =1, ..., J_m. NL models (Williams 1977; Daly and Zachary 1978) can also be derived from the family of GEV-models (McFadden 1978). Using GEV-notation, the choice probability for the NL model is given as:

$$(4) \qquad P_{in} = \frac{\left(\sum_{j=1}^{J_m} e^{\mu_m V_{jn}}\right)^{\frac{\mu}{\mu_m}}}{\sum_{m=1}^{M}\left(\sum_{j=1}^{J_m} e^{\mu_m V_{jn}}\right)^{\frac{\mu}{\mu_m}}} \frac{e^{\mu_m V_{in}}}{\left(\sum_{j=1}^{J_m} e^{\mu_m V_{jn}}\right)}$$

where μ_m are scale parameters applied to the alternatives in nest *m*. We refer to them as *nest* or *structural* parameters. Similar to the group scale parameters υ_{g_n}

(but arising from a different perspective), the nest parameters are inversely related to the corresponding error variance. A restriction of NL model is that nests cannot overlap, that is, each alternative can only enter one nest.

The overall scale of utility, μ, here interpreted as the scale for the choice between nests, is an arbitrary positive number and only the ratio $\frac{\mu}{\mu_m}$ has a behavioural interpretation. It can be shown (e.g. Bhat 1997) that the correlation between the utilities of two alternatives *I* and *j* is given by:

(5) $$Corr(U_i, U_j) = \left(1 - (\frac{\mu}{\mu_m})^2 \right) d_{ij}$$

where d_{ij} is one when *i* and *j* belong to nest *m* and zero otherwise.

A low error variance in nest *m* (i.e. μ_m relatively larger than μ) implies a large correlation among utilities between the nested alternatives. For GEV-conditions to hold, we need $\mu_m \geq \mu > 0$. This implies that the utility of the nested alternatives must be positively correlated. This has to be taken into account when setting up a nested structure. The choice probability in (4) has a nice two-fold interpretation as the product of the probability of choosing between nests (choice at the 'upper level') and the probability of choosing between alternatives in the chosen nest (choice at the 'lower level').

2.4. Cross-nested logit (CNL) models

A GEV-model that allows alternatives to enter several nests is the CNL (Williams 1977; Vovsha 1997; Bierlaire 2006). Choice probabilities in the CNL model are given by (Abbe *et al*. 2007, page 797)[2]:

$$(6) \qquad P_{in} = \sum_{m=1}^{M} \frac{\left(\sum_{j=1}^{J_m} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{jn}}\right)^{\frac{\mu}{\mu_m}}}{\sum_{m=1}^{M}\left(\sum_{j=1}^{J_m} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{jn}}\right)^{\frac{\mu}{\mu_m}}} \frac{\alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{in}}}{\left(\sum_{j=1}^{J_m} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{jn}}\right)}$$

Bierlaire (2006, page 293) also derives the following conditions to be met by a CNL model:

1. $\mu_m \geq \mu > 0$ for all $m = 1, \ldots, M$

2. $\alpha_{im} \geq 0$ for all *j* = 1, ... , J , *m* = 1, ..., M

3. $\sum_{m=1}^{M} \alpha_{jm} > 0$ for all *j* = 1, ..., J.

Following Train (2009), we refer to the α-parameters as *allocation parameters*[3]. The NL-model is a special case of CNL-model where all α_{jm} are zero except for the nest *m* the alternative is included in. The exact correlation structure of CNL

---

[2] Equation (6) is the resulting choice probability for the most general formulation of the CNL, but simpler formulations are available (Ben-Akiva and Bierlaire 1999; Wen and Koppelman 2001). BIOGEME 1.8 (and later versions) use the CNL model version in equation (6).

[3] For interpretation and parameter identification, the condition $\sum_{m=1}^{M} \alpha_{jm} = 1$ should be imposed. Then, the allocation parameters are readily interpreted as the portion of an alternative that enters each nest. However, the relationship between α_{im} and μ_m is not obvious. Intuitively, a high correlation between nested alternatives (high μ_m) should go along with a relative high portion of a particular alternative being associated with that nest. However, we are not aware of suggestions for possible functional relationship between α_{im} and μ_m.

models (Abbe *et al.* 2007[4]) is much more involved than that for the NL (5). The following is an approximation proposed by Papola (2004):

$$(7) \qquad Corr(U_i, U_j) \approx \sum_{m=1}^{M} \alpha_{\mathrm{im}}^{\frac{1}{2}} \alpha_{\mathrm{jm}}^{\frac{1}{2}} \left(1 - (\frac{\mu}{\mu_{\mathrm{m}}})^2 \right)$$

Equation (7) underlines the fact that the allocation parameters affect the correlation structure of the model. Equations (6) and (7) can be thought of as 'weighted averages' of (4) and (5) respectively, where averages are taken over nests and the allocation parameters represent weights.

3. Deriving a NL Model from a HC Model on Binary SC Data

In this section we will critically assess the procedure of translating group scale parameters obtained from an estimated HL model into nest parameters of a NL forecasting model; this approach was applied by Atkins (2012a) in the Norwegian HSR assessment study, so the discussion has practical relevance. In this context the group scale parameters stem from the different subsets of travellers (using various travel modes in practice) being subject to different SC experiments (asking them to choose between their current mode and HSR, see Figure 1).

We assume that no RC data is available, to replicate the situation in Atkins (2012) that only used SC data in their estimation model. In section 4 we will discuss the use of RC data as a supplement.

[4] See equation (20) in Abbe *et al* (2007, page 800) for the exact formula of correlation between two alternatives in a CNL.

3.1. Mathematical conditions

As mentioned above, group scale parameters (υ_{g_n}) and nest parameters (μ_m) are both inversely proportional to their related error variances. The difference lies in which error variance is considered. Group scale parameters relate to the error variance in choices between (non-nested) alternatives of a particular user group. Nest parameters, instead, relate to the choice between alternatives in one nest (independent of the user type). An interesting question is under which conditions the two types of scale parameters may be equal and have the same behavioural implications. To answer this, we examine under which conditions the resulting choice probabilities (P_{in}) in (4) and (3) would be equivalent; that is, under which conditions a NL model can be written as a HL model with scale parameters related to groups with (possibly) different choice sets.

This is shown formally in Box 1.

$$\frac{(\sum_{j=1}^{J_m} e^{\mu_m V_{jn}})^{\frac{\mu}{\mu_m}}}{\sum_{m=1}^{M}(\sum_{j=1}^{J_m} e^{\mu_m V_{jn}})^{\frac{\mu}{\mu_m}}} * \frac{e^{\mu_m V_{in}}}{\sum_{j=1}^{J_m} e^{\mu_m V_{jn}}} \rightarrow \frac{e^{\upsilon_{g_n} V_{in}}}{\sum_{j \in C_{g_n}} e^{\upsilon_{g_n} V_{jn}}}$$

if and only if:

I) $C_{g_n} = 1, \dots, \mathrm{J_m}$

II) $\mu_\mathrm{m} = \upsilon_{g_n}$

and III) $\frac{(\sum_{\mathrm{j=1}}^{\mathrm{J_m}} \mathrm{e}^{\mu_\mathrm{m} \mathrm{V_{jn}}})^{\frac{\mu}{\mu_\mathrm{m}}}}{\sum_{\mathrm{m=1}}^{\mathrm{M}}(\sum_{\mathrm{j=1}}^{\mathrm{J_m}} \mathrm{e}^{\mu_\mathrm{m} \mathrm{V_{jn}}})^{\frac{\mu}{\mu_\mathrm{m}}}} = \begin{cases} 1 & for\ i \in C_{g_n} \\ 0 & otherwise \end{cases}$

Box 1: Mathematical conditions for NL model equalling a HL model with user-group specific choice sets

This implies that the estimated group scale parameters υ_{g_n} could only be used as nest parameters (in a mathematical sense) if alternatives were nested according to the group-specific choice sets (condition I) and if the choice between nests was deterministic (condition III). This puts hard/impractical restrictions to the methodological correctness of a naive translation and, indeed, strong assumptions are required in practical approaches (see section 3.3.).

The fundamental reason for the immediate mathematical incompatibility between group scale and nest parameters goes back to the fact that the former are user group specific while the latter are travel mode specific. This point is essential also for the interpretation of scale parameters discussed in the next section.

3.2. Source and interpretation of scale parameters

The inverse proportionality of the scale parameters to the error variances implies that the (classical) sources of the error term in discrete choice models (Manski 1973, Ortúzar and Willumsen 2011), that is, unobserved attributes, taste heterogeneity, measurement errors and use of instrumental/proxy variables, are the possible main sources of the utility scale parameters.

We recall that the NL and HL models are both more flexible than the MNL model as they relax the IID assumption of the error terms (which is often a restrictive assumption in practise). The relaxation of the IID assumption by the NL model is based upon the fact that the error term of the utility functions of different travel modes are correlated. Travel modes with (significant) positive correlation should be candidates to be nested together. The idea is to account for non-proportional substitution patterns caused by the correlated error terms. The typical interpretation is that travel modes that are closer substitutes (those nested together) share unobserved attributes (Williams 1977). Relaxing the IID assumption and accounting for the patterns of unobserved attributes can be important, as illustrated by the well-known blue bus/red bus paradox (Mayberry 1973, Ortúzar and Willumsen, page 214).

The relaxation of the IID assumption in the HL model stems instead from different error variances associated with different subsets of the data. Various reasons for error variances (and thereby scale parameters) to differ are possible. An obvious candidate is the potentially variable impact of unobserved attributes between travel modes involved in the different choice sets. Common sense suggests, for example, that HSR should share more unobserved attributes with the traditional train than with car, in which case the binary choices between train and HSR are less affected by those unobserved factors. On the other hand, the choice between HSR and car is likely to be affected by several unobserved attributes that differ between the modes (i.e. the varying utility associated with having a car available at the point of destination), making the overall impact of the error term more important.

For example, if the unobserved, and varying, “need to have a car at destination” has greater importance, the superior ("observed") LoS of HSR might not impact the choice probabilities between car and HSR that much. Thus, a relatively low group scale parameter for the car user group should be expected. From this perspective it seems reasonable to use information about group parameter scales to derive a NL forecasting model

However, the scale parameters for the different binary choices might also be high when the taste heterogeneity of users within the subgroup is relatively low. Taste heterogeneity is, to a large degree, unobserved as it involves unobserved factors/preferences relating to the users. The difference with the unobserved attributes discussed in the previous paragraph is that the latter are travel mode specific while taste heterogeneity is person-specific (or user-group specific). Car users, for instance, might have less homogenous preferences of the (observed) LoS than train users (e.g. the subjective Value of Time might vary more among car drivers than among train users). From this perspective, the error variance in the car/HSR choices might be lower than in the train/HSR choices. If (user group specific) taste heterogeneity is the predominant source of the error variance, the estimated scale parameters are not suitable to represent (travel mode specific) nest parameters; the translation of "variance" into "correlation" would not be sound in this case.[5]

3.3. Validity of Practical Approaches

In this section we discuss the validity of (and the necessary assumptions required for) practical approaches to construct a hierarchical forecasting model based on binary SC data. Striving for such a model (instead of applying a simple MNL model ignoring the different sizes of group scale parameters), acknowledges the fact that there might be indeed non-proportional substitution patterns between travel modes worth accounting for when predicting choice probabilities.

The mathematical conditions (section 3.1.) require making some assumptions. A HL model as in equation (3), does not include the respondent's 'choice' about which user group s/he belongs to (this is predefined by the researcher based on

[5] Another potential source of group scale variability are the different degrees of measurement errors in the subsamples. Arguably this is not an issue in CE where attribute values are directly coded as they are presented in the respondent's screen. The use of proxy variables can also be a source for different scale parameters. This applies when a specified proxy variable is a precise representation of the actual variable for the binary choices of one user group, but an imprecise one for the choices of another user group. We will not discuss this further here but maintain the assumption made at the introduction, that the specification of the deterministic utility function could be done "equally well" for all user groups.

the non-modelled RC). Thus, the scale in choices between current modes is unobserved. In the absence of further information, it is necessary to assume the relative scale at the upper level of that hierarchical forecasting model. The choice is restricted by the fact that the scale at the upper level in a NL or CNL model cannot be larger than the scale at the lower level. A sensible choice, resulting in the least complex implicit structure, is to assume that the scale at the upper level equals the lowest estimated group scale parameter yielding a degenerate nest with the current travel mode characteristic for the user group with the lowest group scale parameter.

Based on the discussion in section 3.2, it is evident that we have to make sure that the group scale parameters do represent different substitution patterns of transport modes. For now we assume that taste heterogeneity could be controlled for in the deterministic utility function and that measurement errors and proxy variables are not an issue. With this assumption, different error variances in subsamples can indeed be interpreted as representing different substitution patterns (correlation) across travel modes and HSR should be nested with the current travel mode(s) associated with the highest group scale parameter(s). If two (or more) group scale parameters are different from each other, the different degrees of correlation between HSR and the corresponding transport modes should be taken into account with a CNL specification.

To make things more specific, Table 1 discusses four potential cases of estimated group scale parameters. The group scale parameter for car-users is fixed to unity in these examples. If all estimated group scale parameters were close (and insignificantly different) to unity (i.e. case 1 in Table 1), a MNL would be obtained.

Table 1: Possible nesting structures suggested by group scale parameters (SC data only)

	Case 1	**Case 2**	**Case 3**	**Case 4**
Estimated group scale parameter in SC "current mode vs. HSR"	car-users ≡1 ; train-users ≈ 1; air-user ≈1;	car-users ≡1 ; train-users ≈ 1; air-user ≈3;	car-users ≡1 ; train-users ≈ 3; air-user ≈3;	car-users ≡1 ; train-users ≈ 2; air-user ≈3;
Proposed nest 1	-	car	car	car
Proposed nest 2	-	train	train, air, HSR	train, HSR
Proposed nest 3	-	air, HSR	-	air, HSR
Proposed structure of forecasting model*	MNL	NL	NL	CNL

*Under the assumption that the overall scale (at the upper level) is one.

As much of the discussion provided here is (implicitly) about the reasonability of translating "variance" into "covariance", it is useful to take a closer look at the covariance structure associated with the estimation and forecasting model. Let **x** and **y** denote the following vectors:

$$(8) \qquad x' = \begin{pmatrix} U_{car-user,car} \\ U_{train-user,train} \\ U_{air-user,air} \\ U_{car-user,HSR} \\ U_{train-user,HSR} \\ U_{air-user,HSR} \end{pmatrix}, \quad y' = \begin{pmatrix} U_{car} \\ U_{train} \\ U_{air} \\ U_{HSR} \end{pmatrix}$$

Then, Cov $(x'x)$ is the covariance matrix consistent with the HL model, while Cov $(y'y)$ would be the covariance structure for the proposed forecasting model. For case 1 we would need to translate:

$$(9) \qquad \mathrm{Cov}(x'x) = \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ into } \mathrm{Cov}\,(y'y) = \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

While being logical, the validity of this translation rests on the aforementioned assumptions regarding the size of the upper level scale parameter, which in this case implies that the scale in choices between current modes is assumed to be equal to the scale in choices between single current modes and HSR. The correctness of this assumption is not testable (without additional data on choice between all alternatives) due to the inherent missing information in binary choice data with only one common travel mode.

In case 2, if the group scale parameter for air-users was estimated as significantly higher than those for the remaining groups, this would indicate that air and HSR are closer substitutes and a NL structure with air and HSR in one nest and two degenerate nests for car and train could be proposed. In that case the following translation would apply:

$$(10) \quad \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & (1/3)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1/3)^2 \end{pmatrix} \text{ into } \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & (1-1/3)^2 \\ 0 & 0 & (1-1/3)^2 & 1 \end{pmatrix}$$

Thus, the variance in the binary choices between air and HSR for current air users would be used to set the covariance between air and HSR for the full forecasting model (via equation 5). This is only valid if the group scale parameters can really be interpreted as accounting for different degrees of similarity (related to unobserved attributes) regarding the transport modes (see the discussion above). Given the above made assumption of μ= 1, the group scale parameter estimated as equal to three can be directly used as the structural parameter in the nest containing air and HSR.

In case 3, if the scale parameter for train-users is estimated as significantly greater than one and insignificantly different from the air-user scale parameter, train, air and HSR might be nested together. Similar to the second case, the following correlation structure in the forecasting model would be proposed:

(11)

$$\frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & (^1/_3)^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & (^1/_3)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & (^1/_3)^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & (^1/_3)^2 \end{pmatrix} \text{ into } \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & (1-1/3)^2 & (1-1/3)^2 \\ 0 & (1-1/3)^2 & 1 & (1-1/3)^2 \\ 0 & (1-1/3)^2 & (1-1/3)^2 & 1 \end{pmatrix}$$

Note, that the correlation between the train and air alternatives is derived from the correlation between train and HSR, and air and HSR (assuming that taste heterogeneity is controlled for). This is not obvious and cannot be assessed without choice data between train and air (see the discussion in section 4).

Finally, in case 4, if the train-user scale is greater than one but significantly lower than the air-users scale, the only valid option would be to allow for HSR entering one nest with train and another nest with air. In that case a CNL model would be required[6] and the following correlation structure would be desirable[7]:

(12)

$$\frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & (^1/_2)^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & (^1/_3)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & (^1/_2)^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & (^1/_3)^2 \end{pmatrix} \text{ into } \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & (1-1/2)^2 \\ 0 & 0 & 1 & (1-1/3)^2 \\ 0 & (1-1/2)^2 & (1-1/3)^2 & 1 \end{pmatrix}$$

Note that it may be difficult, in application, to find a CNL model that implies this correlation structure, as the choice of allocation parameters in conjunction with

---

[6] Apart from a CNL model, a fully general mixed logit (ML) model (Train 2009), might be an alternative and provide an even better way to handle this issue at the expense of more complex estimation, interpretation and application.

[7] As for case 3, the correlation assumed for train and air is somewhat arbitrary. It might be reasonable to allow for correlation between air and train as well, but this cannot be directly derived from the given binary data alone.

the nest parameters is a non-trivial task. This suggests estimating a CNL model from the data rather than to try to infer one such model from a HL model (see section 4).

3.4. Empirical Illustration on Own SC Data

This subsection provides some estimation results that supplement the theoretical discussion of the previous sections. We use data from an independent SC study conducted by the Institute of Transport economics (TØI) in 2010 (Halse 2012, Flügel and Halse 2012a). Similar to the official assessment study (Jernbarnverket 2012, Atkins 2012), the SC consisted of binary choices and were pivoted on observed RC data (Figure 1). In fact, the RC stem from an on-side, pen-and-pencil study that asked travellers to provide general information about their current mode choice in the main long distance corridors in Norway[8] (Denstadli and Gjerdåker 2011). In the last item, travellers were asked to leave their e-mail address to receive a web-based survey concentrating on high-speed rail.

In the SC-survey, each respondent had to make 14 choices between its current mode of transport (as observed in the on-side study) and a hypothetical HSR. The attributes characterizing the transport modes were: total travel costs, in-vehicle travel time, travel time to station/airport (‘access time’), travel time from station/airport (‘egress time’), frequency (number of departures per day) and the share of the ride spent in tunnels ('tunnel share'). In the first eight choice tasks ('CE1'), the attributes of the current mode were kept fixed to their reported values, while they varied within certain percentage changes in the last six choice tasks ('CE2') (see details in Halse 2012). CE2 included also an opt-out option ('neither

[8] Oslo-Trondheim, Oslo-Bergen and Stavanger-Bergen. For the SC-study, only the former two corridors were considered.

of the two alternatives'), which was, however, seldom chosen and not considered in the models of this paper.

A sample of 893 respondents completed the online SC-study (about 33% of the invited respondents). We focus here on the subsample of leisure trips for which 607 respondents were considered. The general choice behaviour of the subsample is summarised in Table 3.

Table 3: Sample size of user groups and general choice behaviour (leisure trips)

User group as defined by the RC choices	**Group size RC**	**Group size SC***	**Percent of SC choices (%)**			**Percent of respondents always choosing in SC ("Non-Traders") (%)**		
	Final sample	Final sample	Current mode	HSR	Opt-out	Current mode	HSR	Switch between
Car	3833	320	61.6	37.7	0.6	31.3	13.8	55.0
Air	920	76	34.5	64.7	0.8	1.3	31.6	67.1
Train	2867	176	40.8	57.2	2.0	6.3	18.2	75.6
Bus	480	35	41.0	58.5	0.6	0.0	17.1	82.9

*Compared to a representative dataset (Denstadli and Gjerdåker 2011), we have under-sampled current air users somewhat and over-sampled current car and train-users. External weights were used during estimation to offset this.

Car drivers are least likely to choose HSR in the SC data and a considerable share of car-users (31.3%) always choose car over HSR, i.e. independent of the varying attribute values in the 14 choice situations per respondent. This indicates that unobserved factors may have affected many of the choices between car and HSR.

Table 4 provides estimation results for HL models on pooled data of different binary SC. As a first benchmark, we include a model where all group scale parameters are fixed to one; in this case the HL model collapses to an MNL model. The difference between SC_HL_1 and SC_HL_2 is that the latter has random coefficients (normally distributed over decision makers) related to the most important level-of-service (LoS) attributes: in-vehicle time, access/egress time, travel cost and (the inverse) of the frequency measured as the number of departures per day. All models were estimated with BIOGEME (Bierlaire 2003, 2008).

Table 4: MNL and HL models on SC data

Model Index	**SC_MNL**		**SC_HL_1**		**SC_HL_2 (random coefficients*)**	
Coefficient	**Value**	**Rob. t-stat (0)**	**Value**	**Rob. t-stat (0)**	**Value**	**Rob. t-stat (0)**
Travel cost (NOK)	-0.00283	-10.45	-0.00189	-6.69	-0.00963	-6.08
sigma cost					0.00551	4.19
Interaction: Dummy "missing income" - travel cost	-0.00709	-3.1	-0.00476	-3.85	-0.00778	-1.27
Interaction: Dummy "did not pay" - travel cost	0.00155	2.87	0.00105	2.92	0.0051	10.32
In-vehicle* (min)	-0.0023	-2.41	-0.00176	-2.98	-0.0155	-5.64
sigma in-vehicle time					0.0119	8.95
Access + egress time** (min)	-0.00169	-0.99	-0.00235	-2.29	-0.00842	-4.82
sigma acc+eg time					0.034	2.75
Dummy (travel time <6h)	0.496	3.06	0.229	2.11	0.145	1.03
1/frequency	-1.15	-3.09	-0.474	-2.68	-2.75	-4.34
sigma 1/ frequency					4.37	5.75
Tunnel share (%)	-0.00232	-0.47	-0.00251	-0.9	-0.0175	-2.74
ASC-HSR	0.248	0.86	0.184	0.95	0.724	1.53
ASC-Car	0	fixed	0	fixed	0	fixed
ASC-Air	-0.819	-2.92	-0.511	-2.98	-1.42	-1.73
ASC-Train	-0.111	-0.5	0.00269	0.02	0.218	0.37
ASC-Bus	0.3	0.96	0.374	2.08	0.331	0.63
Group scale parameters	**Value**	**Rob. T-stat (1)**	**Value**	**Rob. T-stat (1)**	**Value**	**Rob. T-stat (1)**
Car-users	**1**	**fixed**	**1**	**fixed**	**1**	**fixed**
Air-users	**1**	**fixed**	**1.88**	**1.89**	**2.51**	**2.25**
Train-users	**1**	**fixed**	**2.74**	**3.43**	**1.75**	**1.81**
Bus-users	**1**	**fixed**	**4.35**	**2.66**	**2.42**	**1.77**
No. of parameters		12		15		19
No. of observations		8402		8402		8402
No. of respondents		607		607		607
Null-LL		-5822.44		-5822.44		-5822.44
Final-LL		-4677.25		-4572.47		-2612.68
Adjusted rho-square		0.195		0.215		0.548

*Using 500 Halton draws

Comparing first the MNL model with the HL model, we see that the latter has a considerably better final log-likelihood statistic indicating that the inclusion of the group scale parameters improved the estimation on the joint SC data set.

As none of the estimated group scale parameters was below one, the lowest group scale parameter (e.g. the highest error variance) is related to the choices between car and HSR. This seems to fit well with the intuition that car and HSR share the least unobserved attributes with each other (see the discussion above). Surprisingly, the highest group scale parameter appears to be the one corresponding to the SC choices of bus-users. A naive interpretation of this result would indicate that bus and HSR are the closest substitutes and that this should be considered in a forecasting model by nesting bus and HSR in the nest associated with the highest structural parameter. However, from the discussion in 3.2 we should recall that different degrees of unobserved taste heterogeneity in the different subsamples (user groups) should be considered as well.

The estimation results for the random coefficients model (SC_HL_2) show controlling for taste heterogeneity among decision makers lead to considerable changes in the estimated group scale parameters. For example, the group scale parameters for bus and train users are reduced while that for air users is increased. All group scale parameters are not significantly different from two. In the context of finding a plausible structure for a forecasting model, this may suggest a NL model with a (degenerate) nest for the car alternative and a single nest including all public transport options (Figure 3).

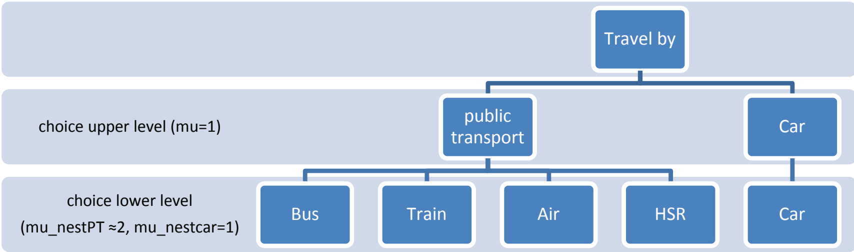


Figure 3: A possible nesting structure for a forecasting model as suggested from SC

While being intuitive and more plausible than what might have been suggested from the results that do not control for unobserved taste heterogeneity, this derivation still rests on two strong implicit assumptions required as a consequence of the missing data issue in the binary stated choices: (i) that the scale between both nests (car and "public transport") equals the scale in the binary choice between car and HSR and (ii) that the correlation in the choices between the current public transport modes (bus, train and air) is derived from the scale in the binary choice between these modes and HSR.

4. Using Additional RC Data among Current Travel Modes

4.1. Motivation

The typical motivation for additional RC data and the joint SC-RC paradigm is the need to ground the SC models in reality (Louviere *et al.* 2000). We will not discuss here the "classical" method of rescaling the SC scale by the RC scale, which became popular after the seminal work of Taka Morikawa (Morikawa 1989; Ben-Akiva and Morikawa 1990) and which is relevant for any kind of SC data (both binary and multinomial). In our context, RC data may provide some of the missing information inherent in binary SC data. In what follows, the focus will be on the correlation structure among current travel modes, which is not "observable" using binary SC data alone.

As RC data involves more than two alternatives in the respondents' choice sets, it is possible and meaningful to estimate hierarchical logit models (NL or CNL) on the RC data. The correlation obtained in a RC model can provide information needed to define plausible correlation structures in the full forecasting model. That is, if estimations on our RC model indicate a nesting structure with one degenerate car nest and another nest including all public transport alternatives (air, train and bus), then the proposed structure from our SC-data in Figure 2 would get interesting support.

4.2. Empirical illustration with RC and SC/RC models

Our RC data includes all relevant travellers in the on-side study (see section 3.4) independent of whether they left or not an e-mail address or whether they were included in the SC study (see sample size in Table 2). Based on the reported geographical information of the trips' start and ending locations, we imported zonal level-of-service data for the related O-D pairs from the Norwegian National Travel Model (Hamre *et al.* 2002[9]).

We tested all possible nesting structures for the four alternatives in the RC dataset (car, bus, train and air), including the same explanatory variables as in

[9] The LoS data contains representative values for the relevant zone pairs and are, in some instances, not updated, such that the RC data must be considered as rather imprecise.

the SC data in Table 3[10]. Somewhat surprisingly, the structure shown in Figure 4 had clearly the best fit to our RC data.

An alternative NL model, where air is nested together with bus and train did perform considerably worst (see Table A1 in the appendix for estimation results). Hence, our RC data did not provide immediate support for the correlation structure suggested from the SC-data alone (Figure 3), as air seems to be more highly correlated with car than with bus or train.

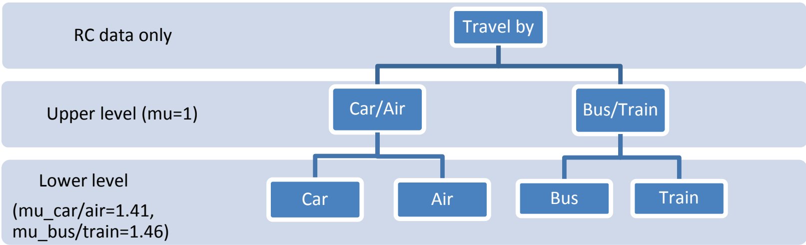


Figure 4: Nesting structure indicated from RC data alone

Combining the information from the RC and SC datasets, the nesting structure shown in Figure 5 might be proposed.

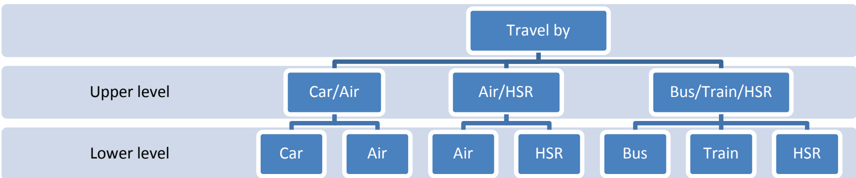


Figure 5: A possible nesting structure for a forecasting model suggested from the information in SC and RC

This structure takes into account the correlation patterns between the current modes as indicated in the RC model and the information from the SC model that

[10] The tunnel attribute was not available from zonal data and was therefore omitted in the RC dataset. Full choice sets were assumed for all decision makers except for the fact that car was only available to respondents that reported owning a car.

HSR is a closer substitute to public transport modes than to car. Both HSR and air enter two nests, thus a CNL model seems to be required.

Two CNL models with the implicit structure in Figure 5 were estimated on the pooled RC/SC data to infer the underlying parameters values (Table 5). The models assume generic coefficients in RC and SC[11]. CNL_1 uses only fixed coefficient while CNL_2 replicates the specification of model SC_HL_2, assuming normal distributed coefficients for the most important LoS-variables.

Table 5: Cross-nested logit models on pooled RC/SC data

Model Index	CNL_1		CNL_2**	
Coefficient	**Value**	**t-stat (0)**	**Value**	**t-stat (0)**
Travel cost (NOK)	-0.00264	-31.95	-0.00684	-37.66
sigma cost			0.00308	0***
Interaction: Dummy "missing income" - travel cost	-0.00061	-6.25	-0.00061	-2.04
Interaction: Dummy "did not pay" - travel cost	0.001	7.49	0.0029	13.02
In-vehicle* (min)	-0.00094	-5.25	-0.00822	-13.87
sigma in-vehicle time			0.0128	15.84
Access + egress time** (min)	-0.00533	-18.93	-0.0276	-19.13
sigma acc+eg time			0.0193	21.48
Dummy (travel time <6h)	0.389	11.25	0.188	2.16
1/frequency	-0.298	-5.74	-4.9	-7.51
sigma 1/ frequency			6.63	8.81
Tunnel share (%)	-0.00404	-2.84	-0.0181	-6.79
ASC-HSR (SC)	0.945	12.23	2.76	13.86
ASC-Air (SC)	-0.372	-4.44	0.0144	0.09
ASC-Train (SC)	0.204	3.14	1.94	8.49
ASC-Bus (SC)	0.363	4.09	2.41	4.89
ASC-Air (RC)	0.947	10.35	0.41	1.53
ASC-Train (RC)	-0.0641	-1.61	1.66	9
ASC-Bus (RC)	-0.193	-3.54	0.307	1.52
Structural parameters	**Value**	**T-stat (1)**	**Value**	**T-stat (1)**

[11] This is a restrictive assumption and, indeed, seems not to hold for our data as indicated by a likelihood ratio tests (Ortúzar and Willumsen 2011, p. 325). We suspect that the main reason for this is the different measure of attributes in RC and SC; however, it could also be that preferences change when the HSR gets available in the choice sets (see also footnote 13 in section 5).

Car/Air	1.8	5.3	7.84	12
Air/HSR	1.57	1.4	6.99	2.35
Bus/train/HSR	4.57	10.61	5.88	0.94
Allocation parameters*	**Value**	**T-stat (1)**	**Value**	**T-stat (1)**
Air to nest Car/Air	0.845	-2.18	0.232	-16.53
Air to nest Air/HSR	0.155	-11.83	0.768	-4.99
HSR to nest Air/HSR	0.527	-12.91	0.595	-6.19
HSR to nest Bus/train/HSR	0.473	-14.36	0.405	-9.09
No. of parameters		22		26
No. of observations		16852		16852
No. of respondents		9057		9057
Null-LL		-16868.4		-16868.4
Final-LL		-11099.5		-9310.01
Adjusted rho-square		0.341		0.447

* The remaining allocation parameters were fixed to 0 or 1 according to Figure 5.
** We used 500 Halton draws.
*** Seemingly some numerical issues were present in the estimate of this standard deviation.

We can use (7) to approximate the inter-alternative variance-covariance matrix for the two model versions[12] as shown in Box 2.

$$y' = \begin{pmatrix} U_{car} \\ U_{bus} \\ U_{train} \\ U_{air} \\ U_{HSR} \end{pmatrix}$$

$$\mathrm{Cov}^{CNL\_1}(y'y) \approx \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0.64 & 0 \\ 0 & 1 & 0.95 & 0 & 0.65 \\ 0 & 0.95 & 1 & 0 & 0.65 \\ 0.64 & 0 & 0 & 1 & 0.17 \\ 0 & 0.65 & 0.65 & 0.17 & 1 \end{pmatrix}$$

$$\mathrm{Cov}^{CNL\_2}(y'y) \approx \frac{\pi^2}{6}\begin{pmatrix} 1 & 0 & 0 & 0.47 & 0 \\ 0 & 1 & 0.97 & 0 & 0.62 \\ 0 & 0.97 & 1 & 0 & 0.62 \\ 0.47 & 0 & 0 & 1 & 0.66 \\ 0 & 0.62 & 0.62 & 0.66 & 1 \end{pmatrix}$$

Box 2: Correlation pattern suggested by CNL models

[12] The actual correlation structure in estimation model CNL_2 may also be affected by the random terms underlying the normal distributed error terms.

The main difference between CNL_1 and CNL_2 is that the latter model (i.e. that controls for unobserved taste heterogeneity) suggests a higher correlation between HSR and air (this corresponds to the comparison between models SC_HL_1 and SC_HL_2 in Table 4). The indicated correlation has approximately the same magnitude as the correlation between HSR, train and bus; something that seems plausible.

It has to be underlined that covariance structure cannot be transformed to other scenarios. They are particular to our data (both RC and SC) and to the specification of our model (i.e. the predefined structure of the CNL model and the chosen deterministic utility function).

With the combined RC/SC modelling, the relative scale between the upper and lower levels is estimated from the data and does not need to be assumed as in the method of translating group scale parameters (from binary SC data) to structural parameters (section 3).

5. Discussion

Modelling the choice of a new alternative (in this case HSR) is a non-trivial task. One important reason for this goes back to the limited data access to revealed choice (RC) data for new travel modes making the collection of stated choice (SC) data a necessity. The papers by Cherchi and Ortúzar (2006; 2011) and Yánez *et al* (2010) have addressed important challenges in the combined analysis of RC and (multinomial) SC data. They discuss how to fit alternative specific constants, to account for taste heterogeneity and to define inter-alternative error structures respectively, and have attempted to provide guidelines on how to cope with these challenges in practice. Our paper acknowledges that analyst's judgment is needed to determine the best way to fit models to SC data for real world forecasting of new alternatives in specific application scenarios. This applies in particular to situations where the new travel alternative is likely to change the competitive structure of the travel market, as is arguable the case for HSR in Norway.

To the extent that introducing a HSR has the potential to change the correlation structure among current modes, it is not guaranteed that information about the current correlation structure among existing travel modes - as indicated by a RC model - has guaranteed validity for future travel decision making[13].

Despite these caveats it would have been interesting to compare the correlation structure in RC and SC models more rigorously. However, a direct comparison as done by Yáñez *et al* (2010) based on multinonial SC data, is not possible with *binary* SC data, with only one alternative (the new travel mode) being common to all subgroups. Therefore, identifying the most appropriate inter-alternative correlation structure for a forecasting model (i.e. preferring the nesting structure of Figure 5 from those in figures 3 and 4) is somewhat arbitrary and subject to the assumption that the translation from "variance" into "correlation" related to the SC data on HSR is reasonable.

[13] Moreover, the introduction of HSR in the choice set may change preferences for LoS-attributes (i.e. beta-coefficients). For example, it is possible that travellers get more sensitive to travel time, in particular that of the air alternative, when HSR is available in the choice set. Therefore, it is recommended to test if the assumption of common coefficients in the RC and SC models holds (Ortúzar and Willumsen 2011, p. 325). A potential challenge arises if the measures of attributes (to whom the beta-coefficients apply) are very different in the RC and SC datasets. In our empirical case study, the nature of the RC and SC was different for two reasons: (i) the former consists of real choices and the latter of hypothetical choices; (ii) attributes in the SC set were pivoted around reported respondents' values (and are, therefore, personal-specific) while attributes in the RC set were inferred from zonal data (and, therefore, representative for the O-D zones of the trips). In this case, it is difficult to determine if the higher estimated value of time from the SC model is due to different measurement of the attributes or due to preference changes when HSR is available.

Given the shortcomings of SC binary choice data as discussed in this paper, it seems indispensable to consider having at least three alternatives in choice experiments, even though this is likely to increase the complexity of the choice tasks. Good practise is found in Yáñez *et al*. (2010) where each SP-respondent had to consider four transport modes: the current mode, the new HSR and two other transport modes that were added to the choice experiment on a random basis.

6. Conclusions

This paper discussed some methodological problems that arise when binary stated choice data between the traveller's current mode and a new option are used to build a generic forecasting model.

The most prominent methodological challenges of *binary* choices (compared to multinomial choices) is the missing data issue related with the choice of current modes; in particular, the "unobserved" correlation between current travel alternatives and the missing scale information between the upper and lower level of a typically hierarchical forecasting models.

A pragmatic approach was used in the official HSR assessment study in Norway, consisting of translating estimated group scale parameters (specific to subgroups of travellers) into alternative-specific structural parameters (Atkins 2011). We have shown that this method is mathematically incorrect and implies, in practise, some strong assumptions regarding the estimated utility scales. We also argued that the method is conceptually dubious (even if the necessary assumptions are made explicit) if one does not control for taste heterogeneity across travellers during model estimation.

By providing some empirical illustrations based on our own (richer) data, we showed that revealed choices may provide the required additional information and that a cross-nested logit model specification may be required in order to account for the variation in estimated scale parameters.

In spite of the fact that there does not appear to be an "objectively correct" method to fit stated choice data about new alternatives in real-world forecasting models, some of the required implicit assumptions might be tested if multinomial (instead of binary) stated choice data is collected.

Acknowledgements

We are grateful to Julián Arellana, Jon-Martin Denstadli, Lasse Fridstrøm, Karen Hauge, Anne Madslien, Ståle Navrud and Christian Steinsland for their help at different stages of this research. We also wish to thank the support of the Millennium Institute in Complex Engineering Systems (ICM: P05-004F; FONDECYT: FB016), the Alexander von Humboldt Foundation, the ALC-BRT Centre of Excellence, funded by the Volvo Research and Educational Foundations, and the TEMPO Project financed by the Research Council of Norway.

References

Abbe, E., M. Bierlaire and T. Toledo (2007) Normalization and correlation of cross-nested logit models. *Transportation Research* **41B**, 795-808.

Atkins (2012) Norway High Speed Rail Assessment Study: Phase III - Model Development Report. http://www.jernbaneverket.no/PageFiles/17406/Economic%20and%20Financial%20Analysis%20-Final%20Report%20Atkins.pdf. (retrieved 2.12.2014)

Bierlaire, M. (2003) BIOGEME: a free package for the estimation of discrete choice models*, Proceedings of the 3rd Swiss Transportation Research Conference,* Ascona, Switzerland.

Bierlaire, M. (2006) A theoretical analysis of the cross-nested logit model. *Annals of Operations Research* **144**, 287-300.

Bierlaire, M. (2008) An introduction to BIOGEME Version 1.8. http://biogeme.epfl.ch/v18/tutorialv18.pdf (retrieved 27.3.2013)

Ben-Akiva, M. and Bierlaire, M. (1999) Discrete choice methods and their applications to short-term travel decisions. In R. Hall (ed.), *Handbook of Transportation Science*. Kluwer Academic Publishers, Dordrecht.

Ben-Akiva, M. and T. Morikawa (1990) Estimation of switching models from revealed preferences and stated intentions. *Transportation Research* **24A,** 485-495.

Bhat, C.R. (1995) A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research* **29B,** 471-483.

Bhat, C.R. (1997) Covariance heterogeneity in nested logit models: econometric structure and application to intercity travel. *Transportation Research* **31B,** 11-21.

Caussade, S., Ortúzar, J. de D., Rizzi, L.I. and Hensher, D.A. (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research* **39B**, 621-640.

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

Cherchi, E. and Ortúzar, J. de D. (2011) On the use of mixed RP/SP models in prediction: accounting for random taste heterogeneity. *Transportation Science* **45**, 98-108.

Daly, A. and Zachary, S. (1978) Improved multiple choice models. In D.A. Hensher and M.Q., Dalvi (eds.), *Determinants of Travel Choice*. Saxon House, Sussex.

Denstadli, J.M. and Gjerdåker, A. (2011) Transportmiddelbruk og konkurranseflater i tre hovedkorridorer, Institute of Transport Economics, Oslo (in Norwegian).

Flügel, S. and Halse, A.H. (2012) Non-linear utility specifications in mode choice models for high-speed rail. *Kuhmo Nectar Conference 2012*, Berlin.

Halse, A.H., 2012. Demand for high speed rail in Norway: Survey design and data collection. Working Paper, Institute of Transport Economics, Oslo.

Jernbaneverket (2012) Høyhastighetsutredningen 2010-2012, Konklusjoner og oppsummering av arbeidet i Fase 3, Del1 (in Norwegian).
http://www.jernbaneverket.no/PageFiles/17299/Rapport_Del_1.pdf.
(retrieved 2.12.2014)

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) *Stated Choice Methods: Analysis and Application.* Cambridge University Press, Cambridge.

Manski C.F. (1973) *The Analysis of Qualitative Choice.* Ph.D. Dissertation., Department of Economics, MIT.

Morikawa, T. (1989) *Incorporating Stated Preference Data in Travel Demand Analysis*. Ph.D. Dissertation, Department of Civil Engineering, MIT.

Mayberry, J.P. (1973) Structural requirements for abstract-mode models of passenger transportation. In R.E. Quandt (ed.), *The Demand for Travel: Theory and Measurement*. D.C. Health and Co., Lexington, Mass.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (ed.), *Frontiers in Econometrics.* Academic Press, New York.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlquist et al., editor, Spatial interaction theory and residential location, pages 75–96, Amsterdam. North-Holland.

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

Ortúzar, J. de D. and Willumsen, L.G. (2011) *Modelling Transport*. Fourth Edition, John Wiley and Sons, Chichester.

Papola, A. (2004) Some developments on the cross-nested logit model. *Transportation Research* **38B**, 833–851.

Train, K.E. (2009) *Discrete Choice Methods with Simulation*. Second Edition, Cambridge University Press, Cambridge.

Vovsha, P. (1997) Cross-nested logit model: an application to mode choice in the Tel-Aviv Metropolitan Area. *Transportation Research Record* **1607**, 6–15.

Walker, J.L. (2001) *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables.* Ph.D. Thesis, Department of Civil and Environmental Engineering, MIT.

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

Wen, C.H. and Koppelman, F.S. (2001) The generalized nested logit model. *Transportation Research* **35B**, 627-641.

Yáñez, M.F., Cherchi, E. and Ortúzar J. de D. (2010) Defining inter-alternative error structures for joint RP-SP modelling: some new evidence. *Transportation Research Record* **2175**, 65-73.

Appendix

Table 5: Multinomial and Nested logit model on revealed choice data

Model Index	**RC_MNL**		**RC_NL_1**		**RC_NL_2**	
Coefficient	Value	Rob. T-stat (0)	Value	T-stat (0)	Value	Rob. T-stat (0)
Travel cost (NOK)	-0.0064	-19.17	-0.0064	-22.04	-0.00526	-14.24
Interaction: Dummy "missing income" - travel cost	-3.03E-06	-0.01	-3.03E-06	-0.02	0.000271	1.38
Interaction: Dummy "did not pay" - travel cost	0.000943	2.3	0.000943	2.23	0.000817	2.21
In-vehicle* (min)	-0.00087	-1.42	-0.00087	-1.21	-0.00066	-1.27
Access + egress time** (min)	-0.00678	-6.99	-0.00678	-7.45	-0.00694	-7.63
Dummy (travel time <6h)	1.18	7.37	1.18	6.78	1.04	7.19
1/frequency	-0.591	-2.48	-0.591	-2.51	-0.655	-3.02
ASC-Car	0	fixed	0	fixed	0	fixed
ASC-Air	2.54	8.88	2.54	8.58	1.87	6.86
ASC-Train	-0.486	-3.26	-0.486	-3.34	-0.503	-3.73
ASC-Bus	-1.35	-7.46	-1.35	-6.52	-1.21	-7.53
	Value	Rob. T-stat (1)	Value	T-stat (1)	Value	Rob. T-stat (1)
Nest car/air	1	fixed			1.41	17.44
Nest train/bus	1	fixed			1.46	7.05
Nest car			1	fixed		
Nest air/train/bus			1.00	0.89		
No. of parameters		10		11		12
No. of observations		5406		5406		5406
No. of respondents		5406		5406		5406
Null-LL		-5767.51		-5767.51		-5767.51
Final-LL		-2360.81		-2360.81		-2337.68
Adjusted rho-square		0.589		0.589		0.593

Essay 3

International Journal of Sustainable Transportation, 9: 364–376

ISSN: 1556-8318 print/1556-8334 online
DOI: 10.1080/15568318.2013.798749

Valuation of Cycling Facilities with and without Controlling for Casualty Risk

STEFAN FLÜGEL[1,2], FARIDEH RAMJERDI[1], KNUT VEISTEN[1], MARIT KILLI[1], and RUNE ELVIK[1]

[1]*Institute of Transport Economics (TØI), Oslo, Norway*
[2]*School of Economics and Business, Norwegian University of Life Sciences (NMBU), Norway*

Received 15 November 2012, Revised 17 April 2013, Accepted 20 April 2013

Barrier effects can impact cyclists' travel time, level of comfort, and risk of accidents. When eliciting the valuation of these elements, simultaneous estimation is called for because the perceived level of comfort may depend on the accident risk. In this paper we present the results of a choice experiment in which cyclists traded off cycling time, separated tracks, intersections, and, in one additional choice experiment, casualty risk. We find that the utility of the two barrier-reducing attributes is almost halved when controlling for accident risk. We also translate the utility to a monetary scale, making the results applicable for cost-benefit analysis.

Keywords: accident, barrier, choice experiment, intersections, separated tracks, travel time

1. Introduction

Cycling as a transport mode must contend with barriers related to the infrastructure developed for motor vehicles (Jacobsen, Racioppi, and Rutter 2009; Stanley and Rattray 1978), for example, intersections and sharing the road with motorized traffic. From a transport economics perspective, these elements enter the generalized travel cost for cycling, and contribute to deselection of cycling as a mode (Elvik 2000). In addition to affecting cycling comfort/convenience and travel time, intersections and motorized traffic may also influence the risk of cycling, or cyclists' perceived worry/insecurity (Elvik 2000; Elvik et al. 2009; Jacobsen, Racioppi, and Rutter 2009; Reynolds et al. 2009). Transport authorities can contribute to increasing cycling as a transport mode by developing infrastructure, in quantity as well as in quality. The economic valuation of such publicly provided bicycling facility development can be assessed by choice experiments, whereby existing or potential bicyclists trade off—for example, separated cycling tracks and grade-separated crossings—against the time use, costs, or accident risk (Abraham et al. 2002; Börjesson and Eliasson 2012; Hopkinson and Wardman 1996; Ortúzar, Iacobelli, and Valeze 2000; Parkin, Wardman, and Page 2007; Tilahun, Levinson, and Krizek 2007; Wardman, Tight, and Page 2007; Wardman, Hatfield, and Page 1997).

In this paper we present results of an Internet-based, stated-preference survey in which cyclists faced choices between cycling route alternatives pivoted onto a recent cycle trip involving separated tracks and grade-separated crossings. In addition to these two barrier-reducing attributes, time use, and then also safety (fatalities/injuries on the section), was included in the choice sets. Due to the difficulty of establishing a credible payment vehicle for cycling facilities, the monetizing of the relative values (part worths) was obtained from between-mode choices, where the alternative mode included a cost attribute (Börjesson and Eliasson 2012; Wardman, Tight, and Page 2007; Wardman, Hatfield, and Page 1997). Our study was generalized to fit any reference cycling trip for transport reported, excluding trips under 10 minutes and access trips to public transport (also omitting recreational and exercise cycling). This generalized approach was adapted to an Internet-based survey for a sample of the Norwegian cycling population. A novel contribution from this study is the simultaneous choice-based valuation of barrier-reducing facilities and accident risk, as well as the comparison against choices not including the accident risk attribute.

The remainder of the paper is arranged as follows: The next section provides theoretical foundations for the economics of cycling for transportation and the valuation of cycling facilities. In the third section, the methodology for choice experiments is described, and the Internet-based survey material is described in the fourth section. The fifth section provides the resulting estimates of the part-worths of cycling facilities, casualty risk, and time savings, as well as the formal test results. The findings are discussed and concluded in the last section.

Address correspondence to Stefan Flügel, Institute of Transport Economics, Oslo, Norway. E-mail: sfl@toi.no

2. Theoretical and Methodological Approaches

2.1 *Barrier Effects, Bicycle Compatibility, and Bicycling Demand Effects from Facility Levels*

Roads and motor vehicle traffic create barriers to cyclist and pedestrian travel (Hine and Russel 1993; Stanley and

Rattray 1978), both in terms of impeding access and causing delay and discomfort.[1] From the transport planning and engineering approaches, several instruments for measuring "level of service" (LOS) or "compatibility" for cycling have been developed since the 1990s (Dixon 1996; Harkey, Reinfurt, and Knuiman 1998; Landis, Vattikuti, and Brannick 1997; Pikora et al. 2002, 2003). Rietveld and Daniel (2004) assessed the effect of policy-related variables controlling for geographical aspects and population characteristics, using travel survey data aggregated at the city level in the Netherlands. They found a clear negative effect on the propensity to cycle from the number of stops (or turns off) per unit distance. Furthermore, relative speed of cycling compared to car speed contributed positively to the cycling demand. They also found a significant negative effect from injury risk. Parkin, Wardman, and Page (2008) also controlled for geographical aspects and population characteristics in their analysis of the propensity to cycle to work in England and Wales, applying UK 2001 census data. They estimated a positive effect (although relatively weak) on cycling to work from the quantity of off-road cycle routes; the positive effect was more strongly related to the pavement quality on these cycle routes.

2.2 *Valuation of Bicycle Facilities Based on Choice Experiments*

Choice experiments enable hypothetical valuation of changes in the attributes of goods or services, such as bicycling trips' time use, facility/pleasance, uninterruption, injury/fatality risk, etc. (Bovy and Bradley 1985; Rizzi and Ortúzar 2003; Wardman, Hatfield, and Page 1997). The respondents carry out a series of choices (trade-offs) between two or more alternatives (options) described by the attributes (or attribute levels), and do not need to state values directly for each attribute; instead the attribute values ("part-worths") are estimated indirectly from the respondents' choices. If the hypothetical choices are related to (pivoted onto) some actual behavior ("a recent bicycle trip"), this is expected to create realism as well as mimicking how choices are carried out in real life (Ben-Akiva and Lerman 1985; Hensher and Greene 2003; Louviere, Hensher, and Swait 2000).

Wardman, Hatfield, and Page (1997) conducted choice experiments to value several cycle facilities. They included travel time and cost of the alternative mode (either car or bus), travel time for cycling, three different levels of on-road facilities for cycling and destination facilities, as well as weather. The valuation of the bicycling facilities and the bicycling time was thus achieved by dividing their coefficients by the cost coefficient for the alternative mode. The monetized value of a cycling facility is calculated as the ratio of the marginal utility of the bicycling facility over the marginal utility of bicycling time $\times$ the ratio of the marginal utility of bicycling time over the marginal (dis)utility of the cost of the alternative mode. In the case of Abraham et al. (2002), conducting a route choice experiment for cyclists (in Calgary, Alberta), the payment mode was somewhat related to bicycling. They used charges for a change room with lockers to arrive at estimates of value of bicycling time. However, this payment vehicle is not a direct cost of cycling (similar to fuel costs or tickets, for car driving and public transport). In their choice experiment they entangled stops at crossings with the time attribute, treating them as a joint variable. If there are differences in comfort/enjoyability of the time spent, that is, differences in the direct utility of travel time, this utility difference should appear in the value of time estimate from time spent in free flow versus waiting at a crossing (DeSerpa 1971).[2] Parkin, Wardman, and Page (2007) specified a logit model for the probability that cycling is acceptable.[3] From a test of the model, based on a small sample of 144 commuters in Bolton, MA, they found significantly positive effects from the proportion of off-road cycle routes as well as cycle lanes adjacent to the road. They found negative effects on route acceptability from signal-controlled junctions, as well as from turns crossing the direction of oncoming traffic. The latter feature involves a considerable accident risk for bicyclists (Elvik et al. 2009; Stone and Broughton 2003).

Regarding the elements of safety and insecurity (worry/discomfort) in bicycling choice, Hopkinson and Wardman (1996) represent an early contribution. They conducted a route choice experiment for cyclists, where particular cycling

[1]The barrier effect may be regarded as a type of congestion cost (Litman and Doherty 2009), a negative external effect from motorized transport on cycling. The choice of transport mode yields several types of negative external effects (Hanley, Shogren, and White 1997). For example, driving a car yields emissions, accident risks, and congestion that are not fully internalized by the driver; some of the costs are borne by others—by society. Cycling produces relatively minor negative external effects compared to motorized transport, and a change from car driving to cycling/walking would thus reduce external costs from transport. Increased cycling/walking may also yield additional positive external effects for society related to land use, the urban environment (liveability), and public health (Elvik 2000; Litman 2003; Pucher and Dijkstra 2000, 2003; Rietveld and Daniel 2004).

[2]Tilahun, Levinson, and Krisek (2007) describe a choice experiment in which respondents faced pair-wise choices between bicycle routes with different facilities, but in which each facility was compared with all other facilities. "For example, an off-road facility (A) is compared with a bike-lane no on-street parking facility (B), a bike-lane with parking facility (C), a no bike-lane no parking facility (D), and a no bike-lane with parking facility (E)" (290). They applied an adaptive choice design, such that the travel time for the route alternatives with better facilities was changed according to previous choices, with initial times of 40 min. for the best facility route and 20 min. for the worst facility route. The route alternatives were described by video clips, plus an indication of travel time. They estimated a bike lane facility valuing 16.41 min., a no in-street parking valuing 9.27 min., and an off-road facility valuing 5.13 min., in terms of being willing to add this travel time relative to a 20-min. trip lacking these facilities.

[3]The logit model had the following form: $\Pr(\mathrm{A}) = 1 \big/ 1 + e^{Z_{ij}^U - Z_{ij}^A}$, where i refers to routes and j to junctions, Z_{ij} represents the overall risk of a journey; and the "utility of cycling being unacceptable (U), Z_{ij}^U, is arbitrarily set to zero and the utility of cycling being acceptable (A), Z_{ij}^A, is a linear function of the variables" (Parkin et al. 2007, 370). The specification of Z_{ij}, with types of routes and junctions represented by dichotomous (presence of particular condition) or continuous variables (intensity of particular condition), would involve contributory effects to cycling demand (or acceptability). In addition to journey attributes, the function can also include individual characteristics, for example, age and gender.

facilities would bring about safety improvements. The facility development involved a "high quality, totally segregated cycle path along the railway line which would be tarmaced and under continual camera surveillance," remarking that "given that the value of cycling facilities, for example, in terms of risk reduction, can be expected to depend on the time spent using the facility, cycling facility and time are treated as a joint variable" (243). They also introduced a payment vehicle—a charge to use the segregated cycle path—making it a kind of toll road for bikes. Ortúzar, Iacobelli, and Valeze (2000) analyzed mode choice in relation to a proposed segregated network of cycle ways in Santiago de Chile. The mode alternatives were bicycle or bicycle combined with metro (access trip by bike to station), against the current mode (car or public transport). Both bicycle and other transport mode alternatives were described by travel time and cost. For the hypothetical new cycle network, the payment vehicle was a bicycle shelter charge. They used the choice experiment for cycle demand estimation, given the construction of segregated cycle ways, finding that it could increase the bicycle share from approximately 1.5% to nearly 6%, while stressing the importance of trip length (in time) as a limiting factor.

Wardman, Tight, and Page (2007) combined revealed preference and choice experiment data to elicit valuations and predict demand effect of measures encouraging cycling. Somewhat similar to Hopkinson and Wardman (1996) and Abraham et al. (2002), Wardman, Tight, and Page (2007) treated cycling facility and time as an interaction variable. A main purpose of the choice experiment was to assess the effect on the time value for cycling with different cycling facilities and to estimate bicycling demand. One subsample valued time on three of the following cycling road facility specifications: major roads with no cycling facilities, minor roads with no cycling facilities, nonsegregated on-road cycle lanes, segregated on-road cycle lanes, and completely segregated cycle lanes. The other subsample valued time with the following destination facilities: parking facilities at destination (outdoors and indoors), and shower/changing facilities at destination. The estimated value of cycling time was the same for major and minor roads with no cycling facilities, while it was reduced to approximately one-third in the case of nonsegregated cycle lanes, close to the time value for motorized commuting transport, estimated at 6.5 pence per min., or GBP 3.90 per hour, in 1999 values.[4] The difference of approximately 12 pence per min., more than 7 GBP/h, yields the estimated value of nonsegregated on-road cycle lanes. The value of time was further halved, approximately, for segregated cycle lanes; and relative to no cycling facilities, the estimated implicit value is approximately 9.50 GBP/h for segregated cycle lanes. Regarding destination facilities, outdoor parking was found equal to 2.5 min. travel time-saving (ca 48 pence per min.), indoor parking 4.3 min., and shower/changing facilities (in addition to indoor parking) 6 min. (about 115 pence per min.).

In a recent paper, Börjesson and Eliasson (2012) present a two-step choice experiment for bicyclists in Stockholm. In the first step the cyclists chose between cycling and an alternative (second-best) transport mode involving travel time, travel cost for the alternative mode, and the share of the bicycle time on a separated path. In the second step, the cyclists faced a pair-wise choice between bicycle routes, "differing in terms of travel times, number of signalised intersections, total waiting time at those intersections and whether there was a bicycle parking facility at the destination" (677). Similar to Wardman, Tight, and Page (2007), Börjesson and Eliasson (2012) found considerably higher value of cycling time, in mixed street traffic, than value of time on the alternative motorized mode (mostly public transport); that is, 15.9 EUR/h compared to 8.7 EUR/h (or 17.6 EUR/h compared to 9.3 EUR/h evaluated at average sample income and baseline travel time below 40 min.). The value of cycling time is reduced to 10.5 EUR/h (or 12.2 EUR/h) for cycling on a separate cycling path, implying a value of cycling on a separate path, relative to cycling in mixed traffic, equal to EUR 5.4 per hour. They find a value of one signalized intersection equal to 1.02 bicycling minute (or 1.1 bicycling minutes for one signalized intersection in addition to the delay). Using the bicycling value of time in mixed traffic, 15.9 EUR/h, the monetized value of a signalized intersection is estimated at 0.27 EUR (0.29 EUR).

2.3 Modeling Choice Experiments Involving Policy-Related Attributes of Bicycling Compatibility—Separate Paths, Elimination of Crossings, and Reduced Accident Risk

Transport authorities can respond to bicycling facility demand by developing infrastructure, in quantity as well as in quality. The policy measure related to stops/crossings involves the construction of grade-separated crossings that will eliminate the need to stop (Elvik 2000, Elvik et al. 2009). Building on the referred choice experiments of bicycling for transport (Börjesson and Eliasson 2012; Ortúzar, Iacobelli, and Valeze 2000; Wardman, Tight, and Page 2007; Wardman, Hatfield, and Page 1997), a two-step modeling of the valuation of policy-related attributes for bicycling is proposed. Practically, this is due to the challenge of finding an appropriate payment vehicle for cycling; there is no direct "out-of-pocket" cost of cycling (e.g., road tolls for cycling lanes) in Norway, and hypothetical payment vehicles could create a credibility problem for the choice scenario.[5] Thus, a first step comprises choices between cycling and an alternative mode involving out-of-pocket costs, similar to the approaches by Börjesson and Eliasson (2012), Ortúzar, Iacobelli, and Valeze (2000), and Wardman, Hatfield, and Page (1997), and, which will yield an implicit valuation of time in cycling. In the second step, the time valuation is applied for valuation of other bicycling attributes. The following random utility functions apply to the first step choices between cycling and

[4]The ratio of the cycling travel time coefficient with no facilities (Time-Y) and the motorized commuting travel time coefficient (Time) is 2.97, from the multinomial logit (MNL) in Table 1, in Wardman, Tight, and Page (2007), thus implying a cycling travel time value of slightly more than 19 pence.

[5]We did develop a prototype scenario involving a bicycling toll road, but it was not implemented.

an alternative paid transport mode (suppressing notation for individuals):

$$\begin{aligned} U_B &= \mathrm{ASC}_B + \beta_{t,B,SEG} \cdot t_{B,SEG} + \beta_{t,B,NO} \cdot t_{B,NO} + \varepsilon_B \\ U_A &= \beta_{t,A} \cdot t_A + \beta_{c,A} \cdot c_A + \varepsilon_A \end{aligned} \tag{1}$$

where U_B and U_A refer to, respectively, bicycling utility and alternative mode utility; $t_{B,SEG}$ and $t_{B,NO}$ refer to, respectively, bicycle travel time on segregated cycle path and bicycle travel time in mixed street traffic; t_A is the travel time on the alternative motorized mode; c_A is the cost of the alternative mode; and $\beta_{t,B,SEG}$, $\beta_{t,B,NO}$, $\beta_{t,A,}$ and $\beta_{c,A}$ are the corresponding coefficients; ASC_B refers to an alternative-specific constant; and ε_B and ε_A refer to error terms assumed to be iid extreme value (Gumbel) distributed, yielding a logit structure of the choice model (McFadden 1974).

The following random utility functions apply to the second-step choices between different cycling trips:

$$U_j = \beta_{t,j} \cdot t_j + \beta_{\mathrm{SEP},j} \cdot \mathrm{SEP}_j + \beta_{\mathrm{CRO},j} \cdot \mathrm{CRO}_j + \beta_{\mathrm{CAS},j} \cdot \mathrm{CAS}_j + \varepsilon_j \tag{2}$$

where U_j ($j \in C$, the choice set) refers to cycling trip alternatives; t_j is the travel times; SEP_j is the share of separated cycle path (off-road cycle/walking path or on-road cycle lane); CRO_j is the number of crossings with motorized transport (per trip length); CAS refers to the number of fatalities and serious injuries per trip length; $\beta_{t,j}$, $\beta_{\mathrm{SEP},j}$, $\beta_{\mathrm{CRO},j,}$ and $\beta_{\mathrm{CAS},A}$ are the corresponding coefficients; and ε_j refers to the error term assumed to be iid extreme value distributed. The value of bicycle travel time in mixed street traffic, from the between-mode choice experiment, can be applied for monetized valuation of separated cycle path, crossings, and casualties in the within-mode choice experiments.

Consistent with the fundamental axiom of consumer theory, a choice of alternative j implies that $U_j > U_k$, or $\Pi_j = \Pi\{U_j > U_k\}$, for all $k \neq j$ (Beggs et al. 1981). The implicit valuation of the attributes in the alternatives can be estimated by logit modeling. Heterogeneous preferences for these attributes can be handled by randomized parameters, such that these follow a distribution in the population. This yields a mixed logit model specification, where choice probabilities have to be simulated:

$$E(\Pi_j) = \int_{\boldsymbol{\beta}} \frac{\exp(V_j)}{\sum_{k \in C} \exp(V_k)} f(\boldsymbol{\beta}) d\boldsymbol{\beta} \tag{3}$$

where V_j refers to the index (deterministic) portion of the random utility function; $\boldsymbol{\beta}$ is the parameter vector, and $f(\cdot)$ represents a density function, the mixing distribution of the parameters. The mixed logit model also allows correlation among choices made by an individual. Not allowing random parameters would reduce (3) to an MNL (Train 2009).

We will assume that all random parameters follow a normal distribution. Fixing the denominator (the cost parameter in the first-step between-mode choice and the time parameter in the second-step choice between generic bicycling alternatives), the additive utility function implies that parameters can be interpreted as marginal valuations; and the value ratios will follow the same distribution as the numerator (Revelt and Train 2000; Ruud 1996).

2.4 Choice Experiment Design

De Jong et al. (2007) describe attribute design for pair-wise choices: a near-orthogonal design avoiding choices with dominant alternatives. It applies to hypothetical choices pivoted on reported trip characteristics, yielding reference (base) levels for the attributes. In each choice pair one of the alternatives includes the base level for the attributes of the choice alternatives. For all attributes, there are two levels with a higher value than the base value and two levels with a lower value than the base value. Travel time refers to door-to-door journey time, while travel cost is the "out-of-pocket" cost for the one-way journey. Because the respondent may not been able to calculate the exact cost of the journey, the researcher could adjust respondents' stated cost, based on reported trip distance, fuel type, and perceived fuel efficiency. Attributes other than time and cost can also be built around reported levels of a reference trip, or estimated with respect to, for example, base distance (travel time) level.

3. The Survey Design

3.1 An Overview of the Bicycling Attributes and Choice Experiment

The following choice experiments will be applied for valuing bicycling time savings, as well as bicycling facilities, related to barrier and insecurity effects, testing for the effect when also including a specific accident risk attribute:

1. Mode choice experiment between cycle and car or public transport, involving the attributes total cycle time (t_B), total in-vehicle time (t_A), "out-of-pocket" cost of the trip with car or public transport (c_A), and binary attribute variable segregated cycle path (SEG)
2. Within-mode choice experiment involving the attributes total cycle time (t), share of separate cycle path (SEP), and number of stops/crossings (CRO)
3. Within-mode choice experiment involving the attributes total cycle time (t), share of separate cycle path (SEP), number of stops/crossings (CRO), and number of casualties (CAS)

There are eight pair-wise choices per respondent in the between-mode choice experiment, and there are six pair-wise choices per respondent in the within-mode choice experiments.

3.2 Design for the Between-Mode Choice Experiment

The first choice experiment involving cycling is a between-mode choice experiment, an alternative-specific choice between bicycling (B) and a second-best alternative paid transport mode (A), either car or public transport, against the reported cycle trip. The purpose of this

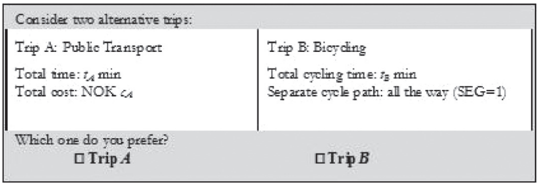


Fig. 1. Presentation of choice pairs for between-mode choice experiment.

experiment is to establish a value of travel time for bicycling, which will then be used for implicit valuation of other bicycling attributes in within-mode cycling choices omitting a payment vehicle.

In the between-mode choice experiment, the "base time" was not based on the reference bicycle trip time, but was randomly assigned to respondents (either 15, 21, 26, 32, 38, 43, 52, 58, 61, or 68 min.). The respondent is asked to choose an alternative mode (car or public transport) to replace the bicycle trip (Fig. 1). The "base time" of the alternative mode (car or public transport) is 0.4 times the "base time" of bicycling, while the "base cost" of the alternative mode, in NOK, is 1.4 times the "base time" of bicycling, below 30 min.; 0.6 times the "base time" of bicycling, between 30 and 60 min.; and 0.35 times the "base time" of bicycling, above 60 min. There were three levels of each of these attributes: ±30% for "base time" below 30 min., ±25% for "base time" of 30 min. or less; and ±30% relative to "base cost." The choice experiment will be a randomized factorial design, with no blocks of choices with redundant pairs of alternatives, following de Jong et al. (2007).

Table 1. The time attribute (t) in choice experiments (CE), with base levels from cyclists' reported trip length (min) on reference trip

Base time of trip (min)	Change in time of the trip relative to base time (min)				
	level −2	level −1	level 0	level 1	level 2
10–19	−4	−1	0	3	6
20–44	−4	−1	0	3	6
45–74	−5	−2	0	3	8
75–119	−10	−5	0	5	18
120–179	−12	−5	0	8	20

Note: No cyclist reported trips above 120 minutes.

Table 3. The crossing attribute (CRO) in choice experiments (CE), with base levels from cyclists' reported no. of crossings on reference trip

Base no. of stops	Change in no. of stops relative to base no.				
	level −2	level −1	level 0	level 1	level 2
0	0	0	0	1	2
1	−1	−1	0	1	2
2	−2	−1	0	1	2
3	−3	−1	0	1	2
4	−4	−2	0	1	2
5	−5	−3	0	2	3
6–8	−6	−3	0	2	4
9–12	−9	−5	0	3	5
13+	−13	−7	0	4	7

3.3 Design for the Within-Mode Choice Experiments

3.3.1 *Time, Share of Separate Cycling Paths, and Number of Crossings*

The following table displays the bicycling travel time attribute base levels and variation with respect to the base level, for application in the second-step within-mode choice experiment (Table 1).

For the share of separate cycling path, it was deemed necessary to simplify the attribute structure, including a disconnection from stated reference levels (Table 2).

The design of attribute levels for number of stops (due to crossings) followed more closely the reported levels of the respondents' reference trip, although this implied a problem for the variation over attribute levels when the reported base was 0 or 1. When the base number of stops is zero, "level −1" and "level −2" also have to be zero, and similarly when the base number of stops is 1, the relative change has to be − 1 in both "level −1" and "level −2" (Table 3).

In a first within-mode choice experiment, the respondent is asked to choose between cycling trips involving the attributes total cycle time (t), share of separate cycle path (SEP), and number of crossings (CRO), altogether six pair-wise choices (Fig. 2).

3.3.2 *Including a Safety Attribute—the Number of Casualties Estimated for the Reference Trip*

A pertinent issue for our research was the extent to which barrier/insecurity elements could be valued in choice

Table 2. The separate cycling path attribute (SEP) in the second and third choice experiments (CE), with simplified base levels from cyclists' reported share of separate cycling path on reference trip

Base level of separated bicycling path, in percent of distance	Default base levels and upward and downward attribute levels, in percent of distance				
	level −2	level −1	level 0	level 1	level 2
>50%	80%	50%	30%	15%	0%
<=50%	60%	40%	20%	10%	0%

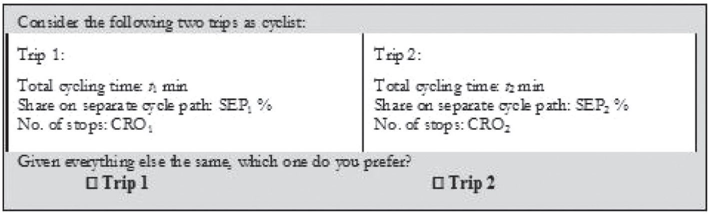


Fig. 2. Presentation of choice pairs for the first within-mode choice experiment.

experiments, conditioning versus not conditioning on a particular accident risk attribute. The within-mode choice experiment without an accident risk attribute (Fig. 2) could be tested against a choice experiment including an accident risk attribute. A way of presenting accident risk in choice experiments is to present numbers of casualties in cycling accidents on the reference route section of the cyclist (Rizzi and Ortúzar 2003). However, in contrast to travel time, separated cycle sections, and grade-separated crossings, the number of casualties—or casualty risk—is not a type of information that the cyclist can be expected to report. In a generalized choice experiment for different reference trips, the casualty risk (the annual casualty number) can be estimated from the other trip information. More precisely, our approach attaches an annual expected number of fatalities and serious injuries on a cycle road of a certain length with a certain motor vehicle density (annual average daily traffic, AADT) at shared facilities and/or intersections. Initial AADT levels were assigned to each respondent based on the urbanization level at the respondents' place of residence, simplifying to three levels only: 12,000 for city, 6,000 for other densely populated areas, and 2,000 for rural areas (Elvik 2008). The initial AADT level could be adjusted one level upward or downward by the respondents' own assessment of traffic density.

Reported trip time is applied for calculating trip length, assuming some average bicycling speed, 12 km/h. The conversion from reference trip time (midpoints) to trip length means that in 15 min. a trip by cycle will cover 3 km [(15/60) • 12], and so on (Elvik 2008). Table 4 shows the procedure of estimating base levels of casualties (serious injuries and fatalities) on cycle road sections of different length and with different AADT levels at shared facilities and/or intersections.

The basis for the calculations shown in Table 4 was the actual casualty numbers in Norway from 1998 to 2005, but adjusted for underreporting, that is, the number of serious injuries was multiplied by 1/0.7 (Elvik and Borger Mysen 1999). Further adjustment upward was considered necessary for the shortest trips to ensure large enough integer values to allow variation up and down from the base level (Elvik 2008). Furthermore, it is known that bicycle injuries are incompletely reported in official Norwegian accident statistics, although less incomplete for serious injuries than for slight injuries (Veisten et al. 2007).

Regarding the casualty attribute range and levels, it followed the design for pair-wise choices from de Jong et al. (2007), similarly to the time attribute, with two lower levels than the base and two higher levels than the base. The two levels with higher values (worse levels) were set

Table 4. Base levels of casualty attribute (CAS), i.e., fatalities and serious injuries, in choice experiments, derived from cyclists' actual trip length (in time).

			Mean annual expected number of casualties					
			Initial estimation			Upward adjustment due to underreporting in official statistics, and adaptation to choice experiment		
Base time (min)	Mean time (min)	Distance (km)	AADT 12,000	AADT 6,000	AADT 2,000	AADT 12,000	AADT 6,000	AADT 2,000
10–19	15	3	0.66	0.49	0.26	3	2	2
20–44	32	6,4	1.40	1.05	0.56	3	2	2
45–74	60	12	2.63	1.97	1.05	4	3	2
75–119	90	18	3.94	2.96	1.58	6	4	2
120–179	150	30	6.57	4.93	2.63	9	7	5

Note: Casualties refer to fatalities and serious injuries. The reported base time of the actual trip was first converted to trip length, by assuming reasonable mean travel speeds, 12 km/hour, based on the national travel behavior survey from 2005 (Denstadli et al. 2006). For the estimation of fatality/injury risk per trip length, it was assumed that the route used would have an injury/fatality risk close to the mean value for all cycling routes (Elvik 2008). No cyclist reported trips above 120 minutes.

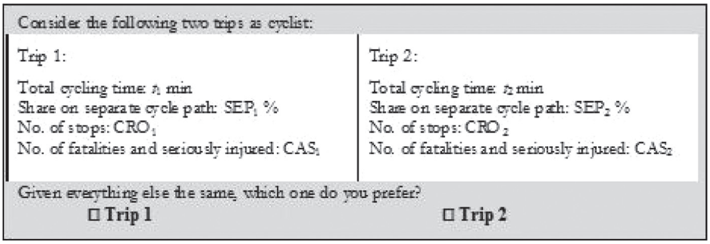


Fig. 3. Presentation of choice pairs for the second within-mode choice experiment.

to, respectively, 15% and 30% above the base level (rounded to integer), while the two lower levels (better levels) were set to, respectively, 15% and 30% below the base levels in Table 1. The exception is for base levels below four casualties, where the increases were set to 1 and 2, and reductions set to −1 and −2, from the base levels (because the low base level would not yield differentiation applying 15% and 30% changes). The first and second within-mode choice experiments, including the specific accident attribute, were presented over six pair-wise choices (Fig. 3).[6]

3.4 Survey Development

In our project we followed the same respondents in two waves of surveying. In the first wave they described a recent cycling trip that yielded reference values for the choice experiments. They also entered the between-mode choice experiment and the first between-mode choice experiment. In the second wave the accident attribute was introduced, with reference level defined from the time use on the reference trip reported in the first wave.

The development of the survey was initiated in 2008. At the end of April 2008 draft scenarios of road safety measures and examples of risk change descriptions related to the second wave of the bicycling survey were presented in a focus group of eight participants. Although the participants indicated understanding of probability communication devices, such as grids with black squares representing fatalities (Alberini and Chiabai 2007), we opted for the approach by Rizzi and Ortúzar (2003), presenting and altering fatality/injury numbers instead of fatality/injury risk figures. In May 2009 a small pretest of the first wave of the bicycling survey, including the registration of the bicycling reference trip as well as the between-mode and the first within-mode choice experiments, was carried out among colleagues. Although no specific pilot testing was carried out among bicyclists, similar survey and choice experiment structures were tested for other transport modes, in two waves, during the first part of 2009 (Ramjerdi et al. 2010; Veisten et al. 2012). For the second wave, the pilot testing of the accident attribute among car drivers resulted in a reduction from 25% and 50% up and down from the base level of casualties, to 15% and 30% up and down from the base level, due to the indicated strong preference for the alternatives with lowest number of casualties.

Our main two-wave survey was applied to a fairly large sample of the Norwegian population, and was carried out first in June and July 2009. Due to a mismatch of the routing of the respondents from wave 1 to wave 2 (Samstad et al. 2010), a new two-wave survey was carried out in April and May 2010. Only the results of the latter 2010 data, in which respondents were correctly routed between wave 1 and wave 2, are reported here. The two-wave Internet-based survey was carried out via e-mail recruiting from the national Internet panel of Synovate Norway. There were 2,408 bicycling respondents in wave 1, and 1,573 (65.32%) of them also participated in wave 2. In the analysis we will only consider those respondents responding to both wave 1 and wave 2 ($n = 1573$). Figure 4 illustrates the sampling procedure of the study.[7]

In addition to questions about the reference bicycle trip and choice experiments, the surveys also included other elements. The questionnaire structures were the following, respectively, for wave 1 and wave 2:

Wave 1:

- Introductory questions about individual characteristics
- The reference bicycle trip
- Between-sample choice experiment
- Within-sample choice experiment (omitting accident risk attribute)
- Questions about bicycling for transport
- Debriefing questions and more questions about individual characteristics

[6]The wave 2 choice experiment also included an opt-out ("do not know") option; and while this might be included in the analysis (Veisten et al. 2012), it will be omitted in this study, primarily for the purpose of comparing the wave 1 choice experiment (omitting the safety attribute) with the wave 2 choice experiment (including the safety attribute).

[7]According to Synovate Norway, our response rates were common for their Internet panel, and they applied techniques to adjust the sample to population figures, that is, distributions of gender, age, and regional appurtenance. Synovate Norway, formerly MMI (Markeds- og Mediainstituttet) AS, was part of the international opinion research company Synovate when carrying out our survey. Synovate Norway joined the Ipsos Group on 1 January 2012 and is now called Ipsos MMI (http://ipsos-mmi.no).

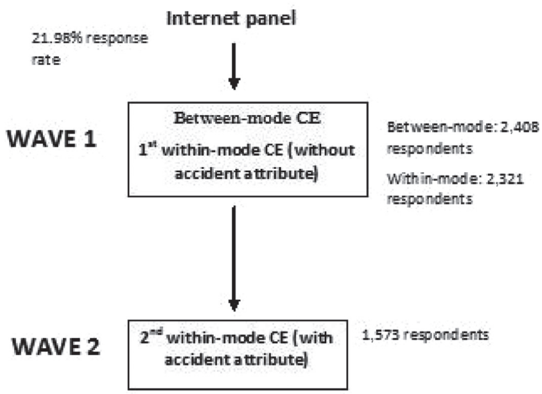


Fig. 4. Two-wave Internet-based survey.

Wave 2:

- Introduction to the issue of fatality/injury risk and casualty numbers
- Scenario for change in casualty numbers
- Within-sample choice experiment including accident risk attribute
- Respondents' income/ability to pay
- Debriefing questions (fatality/injury risk beliefs, accident experience)

4. Results

4.1 Descriptive Analysis

Table 5 lists the means and ranges of individual characteristics.

Before the statistical analysis of the choice experiments, some respondents were excluded if not meeting certain requirements related to their reference trip. Respondents were excluded from wave 1 if they reported a reference trip of over 100 minutes, or over 50 km. Respondents who reported what was considered an unrealistically high number of intersections, that is, more than 10 per km, were also excluded. Finally, respondents for whom an average speed of over 30 km/h was calculated were also excluded, based on the assumption that their reported trip was recreational (for the purpose of exercising) rather than cycling for transport (Ramjerdi et al. 2010).

4.2. Statistical Analysis

Table 6 shows the results of the logit modeling of the between-mode choice. Two different models were estimated, where the rightmost model represents equation (1), including two bicycle time parameters, one for no segregation and one for segregation (indirect valuation of segregated cycle path, via time valuation). The leftmost model includes a single bicycle time parameter and a specific segregation variable (direct valuation of segregated cycle path). The mixed logit version includes a normally distributed error component in the utility function of bicycle, with standard deviation (SIGMA_bic) estimated from the data. This specification is chosen in order to control for differences in the preference for cycling as such, then presumably measuring the marginal utility of the attributes more precisely.

The leftmost models fit the data slightly better, and might also be considered more in line with the choice experiment presentation. The goodness-of-fit measures of the mixed logit models are substantially better than those of the MNL versions. The value of SIGMA_bic is relative high compared to the ASC_bic, indicating that there is a relatively high (unobserved) heterogeneity in preferences for cycling compared to the alternative mode. The value of time in the alternative mode, based on mixed logit modeling, is (NOK 87 per hour in the leftmost model and) NOK 86 in

Table 5. Descriptive statistics for individual characteristics ($n = 1{,}573$)

	Mean	Minimum	Maximum
Age	45.3	17	81
University degree	0.709	0	1
Net personal monthly income (NOK)	23,088	2500	55,000
Income missing	0.0493	0	1
Gender (1 for males)	0.587	0	1
Daily travel distance by bicycle, km*	8.59	0	232
Bicycle trip length (>10 min), km	6.38	0.2	38
Number of crossings per bicycle trip (>10 min)	7.29	0	99
Number of crossings per km traveled	1.53	0	10
Zero share of the reported trip on separated cycling facility	0.19	0	1
Less than half of the reported trip on separated cycling facility	0.26	0	1
About half of the reported trip on separated cycling facility	0.18	0	1
More than half of the reported trip on separated cycling facility	0.30	0	1
All the reported trip on separated cycling facility	0.07	0	1

Note: This only includes the respondents responding to both wave 1 and wave 2. The daily distance cycled would drop to 8.32 km if reported daily distances of over 60 km were excluded.

Table 6. Logit modeling of first-step between-mode choice experiment

	Single bicycle time parameter and specific segregated path parameter				Two bicycle time parameters, with and without segregated path			
	multinomial logit		mixed logit (error component)		multinomial logit		mixed logit (error component)	
	Estimate	Rob st. error	Estimate	Rob st. error	Estimate	Rob st. error	Estimate	Rob st. error
ASC_bic	1.16	0.0596	2.42	0.182	1.67	0.0580	3.28	0.180
SIGMA_bic			2.53	0.0998			2.48	0.100
B_cost_alt	−0.0302	0.00239	−0.0534	0.00429	−0.0303	0.00237	−0.0527	0.00421
B_t_alt	−0.0437	0.00444	−0.0748	0.00815	−0.0441	0.00442	−0.0760	0.00801
B_t_bic	−0.0783	0.00184	−0.146	0.00522				
B_t_bic_seg1					−0.0673	0.00177	−0.124	0.00484
B_t_bic_seg0					−0.0904	0.00213	−0.167	0.00566
B_dummy_sep	1.05	0.0431	1.90	0.0890				
Obs.	12,000		12,000		12,000		12,000	
Respondents	12,000 (pseudo)		1,500		12,000 (pseudo)		1,500	
Null Log L (LL_0)	−8317.766		−8317.766		−8317.766		−8317.766	
Constant Log L (LL_C)	−8312.485		−8312.485		−8312.485		−8312.485	
Final Log L (LL_F)	−6648.060		−5273.747		−6714.433		−5374.465	
Adj. rho-square	0.200		0.365		0.192		0.353	

Note: All models were estimated using BIOGEME (Bierlaire 2003). Robust *t*-tests were computed taking into account the repeated observations nature of the data.

the rightmost model, with 95% CI of (64,109). The bicycling value of time estimates from the mixed logit models are higher than those from the MNL, but the differences are not statistically significant. The "average" value of bicycle time, from the leftmost model, is 164 (140,188). The value of bicycle time in mixed traffic is NOK 190 (162,218) per hour, while it drops to NOK 141 (120,170) on segregated paths.[8] The value of time for alternative modes is estimated lower than for cycling (in the interval between 84 and 87 NOK/hour, depending on model specification). This is most likely due to the lower comfort level in cycling (more physical effort, exposure to bad weather conditions, etc.) compared to car and public transport. The indirect value of a segregated path, per trip, is then NOK 49 per hour traveled; while the direct valuation, from the leftmost model, yields NOK 87.3 per hour traveled.[9]

Table 7 displays results of the second-step within-mode experiments, representing equation (2) without (wave 1) and with (wave 2), respectively, the accident risk attribute. We present pooled models including the choices in both waves. We differentiate between generic modeling, with common parameters in both waves for time use, crossings, and separated paths/lanes, and wave-specific modeling for all parameters except the bicycling time parameter. We include a scale parameter for data of wave 2 in order to measure the relative differences in error variance.[10] Note that the scale parameter does not affect the WTP values because the scale simply cancels out when calculating parameter ratios. The specification of the mixed logit model is somewhat different from the approach in the between-mode modeling. Here we assume random coefficients that are normally distributed. We allow thereby for unobserved taste variation with respect to separate paths, crossings, and casualty risk. The time attribute is assumed to be fixed in order to make the calculation of the WTP values easier.[11] The constant term (ASC) is now related to the left-hand-side alternative and it is expected to be statistically insignificantly different from zero.

The scale parameter for wave 2 is lower than one, as expected, implying higher error variance in wave 2 choices than in wave 1 choices. The mixed logit specification is preferable, compared to MNL, in light of the substantial improvement in goodness-of-fit and the fact that most standard deviations of the random parameters are relatively high, except for the one related to crossings in wave 2 (SIG_CRO_w2). The calculation of parameter ratios is therefore based on the mixed

[8]These value estimates are approximately 20%–30% higher than similar estimates based on the 2009 data (Ramjerdi et al. 2010).

[9]The calculation is based on average reported bicycle time of 27.8 minutes. The value of 35.6 NOK is multiplied by 60/27.8 to obtain the value of 87.3 NOK. Note that the average travel distance value is rather high because only trips over 10 minutes are included in the study.

[10]It is likely that the error variance in the within-mode experiments differs between wave 1 and wave 2. The hypothesis is that the error variance is higher in wave 2 because the inclusion of the casualty attribute adds complexity to the choice decision. The relative impact of unobservable factors is therefore likely to be more prominent in wave 2. The scale parameter for wave 1 is normalized to one, while the scale parameter for wave 2 is estimated from the data. A lower scale parameter will indicate higher error variance (Train 2009).

[11]For estimation of willingness-to-pay (WTP) based on choice experiments where all parameters are random, see, for example, Hensher and Greene (2003) and Daly, Hess, and Train (2012).

Table 7. Logit modeling of second-step within-mode choice experiment, including accident risk attribute in wave 2 (w2) but not in wave 1 (w1).

	Joint parameters				Wave-specific parameters except for the bicycling time parameter			
	multinomial logit		mixed logit (random coefficient)		multinomial logit		mixed logit (random coefficient)	
	Estimate	Rob st. error	Estimate	Rob st. error	Estimate	Rob st. error	Estimate	Rob st. error
ASC_left	(−0.0146)	−0.67	(−0.0156)	0.0300	(−0.0155)	0.0199	(−0.00735)	0.0285
B_t_w1w2	−0.0902	0.00355	−0.166	0.0108	−0.0840	0.00380	−0.158	0.0112
B_CRO_w1w2	−0.100	0.00596	−0.199	0.0146				
SIG_CRO_w1w2			0.209	0.0177				
B_CRO_w1					−0.104	0.00637	−0.223	0.0164
SIG_CRO_w1							0.225	0.0183
B_CRO_w2					−0.0575	0.00893	−0.124	0.0175
SIG_CRO_w2							(0.00821)	0.0165
B_SEP_w1w2	0.0405	0.00117	0.0754	0.00422				
SIG_SEP_w1w2			0.0799	0.00424				
B_SEP_w1					0.0412	0.00114	0.0830	0.00444
SIG_SEP_w1							0.0796	0.00422
B_SEP_w2					0.0251	0.00202	0.0452	0.00496
SIG_SEP_w2							0.0634	0.00723
B_CAS_w2	−2.04	0.111	−3.88	0.300	−1.42	0.103	−2.90	0.291
SIG_CAS_w2			2.77	0.221			2.06	0.214
Scale_w2	0.516	0.0304	0.510	0.0445	0.750	0.0574	0.689	0.0737
Obs.	17,594		17,594		17,594		17,594	
Respondents	17,594 (pseudo)		1,500 (2968)		17,594 (pseudo)		1,500 (2968)	
Null Log L (LL_0)	−12195.231		−12195.231		−12195.231		−12195.231	
Constant Log L (LL_C)	−12194.504		−12194.504		−12194.504		−12194.504	
Final Log L (LL_F)	−9304.911		−8242.755		−9280.718		−8201.474	
Adj. rho-square	0.237		0.323		0.238		0.326	

Note: All models were estimated using BIOGEME (Bierlaire 2003). Robust *t*-tests were computed taking into account the repeated observations nature of the data. The actual number of respondents in the analyses was 1,500 in the wave 1 part and 1,484 in the wave 2 part (since some respondents always choose the opt-out in wave 2). Regarding the mixed logit (random parameter logit) modeling, BIOGEME does not allow that a random variable has different scales for different choices of the same respondent, such that the respondent ID has to be split up, that is, doubled. There is one respondent for whom no information about this within-mode game in wave 1 was available, supposedly because the respondent dropped out after the between-mode choice experiment in wave 1, but afterward the respondent participated in wave 2, completing the wave 2 within-mode choice experiment.

logit results, whereby all numbers are average values of the estimated distribution.

First, we test if the inclusion of the casualty attribute would have a significant impact on the marginal utility of barrier effects, that is, (lack of) separated paths/lanes and crossings. Applying a likelihood ratio test, we tested the null hypothesis of equal parameters for the bicycle barrier parameters. The underlying null hypothesis in the (restricted) generic model is:

$$H_0 : B\_CRO\_w1 = B\_CRO\_w2 \text{ and } B\_SEP\_w1 = B\_SEP\_w2$$

For the mixed logit model versions (where the null hypothesis also comprises equal standard deviations of the random parameters), the likelihood ratio test statistic is:

$$-2(-8242.755 + 8201.474) = 82.562 > 9.49$$

where 9.49 is the critical value for $p = 0.05$ with four degrees of freedom.[12] Thus, we can reject the null hypothesis and conclude that the marginal utility of barrier effects is not independent of the inclusion of the casualty attribute, because, in Table 7, the valuation of barrier effects is higher in wave 1 than in wave 2.

One stop (elimination of one stop/crossing) obtains a value of 1.20 cycling minutes in the joint model, but the wave 1 estimate is 1.41 min. compared to 0.78 min. in wave 2. Applying the average value of bicycling time of 164 NOK/h (2.73 NOK/min.) yields 3.27 NOK per elimination of stop per cycling trip in the joint model, 3.85 NOK in wave 1 vs. 2.14 NOK in wave 2.[13] A marginal increase of the share of separate cycling

[12] The likelihood ratio test statistic for the MNL versions is: $-2(-9304.911 + 9280.718) = 48.39 > 5.99$, where 5.99 is the critical value for $p = 0.05$ with two degrees of freedom.

[13] Börjesson and Eliasson (2012, 680) make an argument for applying the "values of time for cycling on a separate bicycle path when converting the values" of facilities per minute "[s]ince 90% of the respondents had access to a separated bicycle path on more than half of the trip"; and this was also the procedure applied by Ramjerdi et al. (2010). However, Börjesson and Eliasson (2012) actually apply the value of time for cycling in mixed traffic when converting the values per minute to Euro values. We applied the average bicycling value of time, irrespective of cycling facility, since our sample of cyclists reported a variation of the share of the trip on separated facilities going from none to all and approximately equal numbers with less than half as with more than half of the trip on separated cycling facilities (Table 5).

paths, either a 1% increase of on-road cycle lanes or off-road cycle paths, obtains a value of 0.50 cycling minutes in the joint model, or 1.24 NOK, 0.49 min. and 1.34 NOK in wave 1 vs. 0.29 min. and 0.78 NOK in wave 2. For an average trip length of 6.4 km (Table 5), the implicit value (for the joint model) of a km separate cycling facility is then $1.24 \cdot (100/6.4) = 19.4$ NOK per trip.[14]

The specific wave 2 estimate of the casualty per cycling minute is 18.35 versus 23.37 in the joint mixed logit model. Multiplying by 2.73 (the value of cycling time in minutes), this yields 50.10 NOK and 63.81 NOK, respectively, per trip. For the calculation of values of statistical casualties (VSC), VSL, and VSSI, we will apply an AADT of 5,500 for the roads that the bicyclists followed and/or crossed, because this represents the weighted average for the three AADT classes (2,000, 6,000, 12,000) that the respondents were assigned to, and could partly correct. A WTP of 50.10 NOK per casualty reduction per trip yields a value of a statistical casualty (VSC) of 100.6 million NOK, while a WTP of 63.81 NOK yields a VSC of 128.1 million NOK. Assuming that the relative risk of fatality and serious injury for every casualty is 0.2 vs. 0.8, and that the death rate equivalent (DRE) of a serious injury is 0.2, the VSC estimates yield VSL of NOK 279.5 mill and NOK 355.8 mill, respectively, while the VSSI estimates are NOK 55.9 mill and NOK 71.1 mill, respectively. These value estimates are quite sensitive to the AADT estimate (Hensher et al. 2009; Veisten et al. 2012).[15]

5. Discussion and Conclusions

We have carried out simultaneous choice-based valuation of barrier-reducing facilities and accident risk, as well as the comparison against choices not including the accident risk attribute. To our knowledge, such choice experiments have not been described in the literature. Similar to the approach in former choice experiments involving bicycling, the monetary valuation of cycling qualities was established from choices between cycling and an alternative mode that included a cost attribute (Börjesson and Eliasson 2012; Wardman, Tight, and Page 2007; Wardman, Hatfield, and Page 1997). The stated choice experiment was, to a considerable degree, pivoted to respondents' actual trips, applying the reported reference levels for time use, and creating a casualty (safety) attribute from the time use combined with accident statistics (Elvik 2008). The other attributes (barrier-reducing facilities) were also to some extent varying around the level reported for the reference bicycling trip in order to enhance the realism of the choice settings. Because the main purpose of this research was to estimate unit values at the Norwegian national level for time savings, km of separate cycling paths, eliminated stops/crossing, and casualty reductions (Samstad et al. 2010), our Internet-based approach was generalized to fit any reference cycling trip above 10 min. (but omitting recreational and exercise cycling).

The within-mode choice experiment for cyclists was carried out with and without the casualty attribute. This specific feature of the study allowed us to investigate to what degree the valuation of cycling facilities (separate cycling facilities and elimination of intersections with motorized traffic) encompassed a perceived safety gain provided by the facilities. Based on formal testing, we reject the null hypothesis of equal valuation independently of the inclusion of the casualty attribute. Indeed, the valuation of the two facilities almost halved. The estimated valuation for an eliminated intersection reduces from 3.85 NOK to 2.14 NOK and the valuation of a 1% increase in separate path reduces from 1.34 NOK to 0.78 NOK. The reduced values can then be seen as valuation of the facilities itself (i.e., an improvement in comfort and convenience) ‘‘controlling for’’ the utility associated with casualty risk-reduction. Although casualty risk might not include all perceived aspects of safety, these valuations might encompass utility to avoid less-severe accidents (e.g., slight injury from a collision with a pedestrian).

The disutility of longer travel time or delays due to barriers such as crossings can be supposed to have been controlled for through the travel time attribute, which was included in all presented choice experiments. For cost-benefit analysis it is seemingly desirable to obtain valuations of isolated effects, for example, the inconvenience of crossings/intersections when controlling for casualty risk and time use effects, because there already exist time and casualty valuations in the cost-benefit analysis tools of the transport sectors (e.g., in Norway). In this way the potential for double counting will be reduced.

In general, the valuation of separate cycling paths was substantial, and thus well in line with results from former research (Börjesson and Eliasson 2012; Parkin, Wardman, and Page 2007; Wardman, Tight, and Page 2007).[16] Although an implicit safety valuation for separate paths might be considered objectively wrong (Elvik et al. 2009; Stone and Broughton 2003; Veisten, Sælensminde, and Hagen 2005), there might still be some convenience and comfort quality in separate paths that comes in addition to both safety and time use when including the safety attribute. This bicycling mobility quality, or barrier-reducing facility, was also shown

[14]A slightly different model specification was also tried, in terms of rearranging the attribute-specific constant (ASC) to either the ‘‘safer’’ or ‘‘riskier’’ alternative (based on the casualty attribute), where the difference between the ASC can be interpreted as preference for safety per se when traveling (Veisten et al. 2012). The resulting estimated WTP for a casualty reduction, based on this alternative modeling, was NOK 21.53 (15.45, 27.60) per trip, which yields a VSC of NOK 43 million, a VSL of NOK 120 million (86,153), and a VSSI of NOK 24 million (17, 31). The value of a crossing (stop/intersection) was 0.99 cycling minutes, yielding 2.52 NOK (1.92, 3.11) per elimination of stop per cycling trip. And finally, the marginal value of a separated cycling facility was 0.34 cycling minutes, or 0.92 NOK (0.78, 1.06), and for an average trip length of 7 km, the implicit value of a km separate cycling facility then became $0.92 \cdot (100/7) = 13.1$ NOK (11.1, 15.0).

[15]For example, an AADT of 7000 would yield VSC estimates of 43 million NOK for a WTP of 16.88 NOK per casualty reduction per trip and NOK 58 million for a WTP of 22.62 NOK. This would yield VSL of NOK 120 mill and NOK 160 mill, respectively; while the VSSI estimates would be NOK 24 mill and NOK 32 mill, respectively.

[16]Akar, Fischer, and Namgung (2013, 347) found that ‘‘women are more sensitive to being close to bicycle trails and paths.’’ See Larsen, Patterson, and El-Geneidy (2013) for a study using geographic information systems in identifying locations for bicycling facilities.

to be highly valued in a recent Swedish study (Börjesson and Eliasson 2012). Elimination of forced stops due to intersections with motorized traffic yields both improved safety and more convenience and comfort in bicycling (Abraham et al. 2002, Parkin, Wardman, and Page 2007). The valuation of the casualty reduction, when specified as an attribute, was substantial as well, and if brought forward to some demand model or cycling propensity model (Parkin, Wardman, and Page 2007), would clearly predict increased cycling as a response to a safety improvement, with all other things being equal (Ortúzar, Iacobelli, and Valeze 2000).

Finally, investigations into preferences for cycling in transport remain relatively sparse compared to investigations into car driving. We believe more bicycling route choice experiments are warranted, with more diversification in experimental design. We also call for testing of payment mechanisms for cycling facilities, for example, toll roads compared to shelter charges (Ortúzar, Iacobelli, and Valeze 2000), or work-related or school-related subsidies, for example, reduction of subsidies related to improved cycling route facilities.

Acknowledgments

We thank Svenn Jensen, Harald Minken, Juan de Dios Ortúzar, Luis Rizzi, and Hanne Samstad for contributions at various stages of this research. Finally, we acknowledge the comments of three anonymous referees. All remaining errors and omissions are entirely our own responsibility.

Funding

The data collection for this research was funded by the Norwegian Public Roads Administration, the Norwegian National Rail Administration, Avinor, the Norwegian Coastal Administration, and the Norwegian Ministry of Transport and Communications, through the project "Valuation Study."

References

Abraham JE, McMillan S, Brownlee AT, Hunt JD. 2002. Investigation of cycling sensitivities. Paper presented at the 81st Annual Meeting of the Transportation Research Board, 13–17 January, Washington, DC.

Akar G, Fischer N, Namgung M. 2013. Bicycling choice and gender case study: The Ohio State University. International Journal of Sustainable Transportation 7(5):347–365.

Alberini A, Chiabai A. 2007. Discount rates in risk versus money and money versus money tradeoffs. Risk Analysis 27(2):483–498.

Beggs S, Cardell S, Hausman JA. 1981. Assessing the potential demand for electric cars. Journal of Econometrics 16(1):1–19.

Ben-Akiva M, Lerman SR. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA: MIT Press.

Bierlaire M. 2003. BIOGEME: A free package for the estimation of discrete choice models. Paper presented at the 3rd Swiss Transportation Research Conference, 19–21 March, Monte Verita, Ascona, Switzerland.

Börjesson M, Eliasson J. 2012. The value of time and external benefits in bicycle appraisal. Transportation Research Part A: Policy and Practice 46(4):673–683.

Bovy PHL, Bradley M. 1985. Route choice analysed with stated preference approaches. Transportation Research Record 1037:11–20.

Daly A, Hess S, Train K. 2012. Assuring finite moments for willingness to pay in random coefficient models. Transportation 39(1):19–31.

de Jong G, Tseng Y, Kouwenhoven M, Verhoef E, Bates J. 2007. The Value of Travel Time and Travel Time Reliability, Survey Design (Final Report prepared for The Netherlands Ministry of Transport, Public Works and Water Management). Significance, Leiden, Netherlands.

Denstadli JM, Engebretsen Ø, Hjorthol R, Vågane L. 2006. 2005 National travel survey – Key results. TØI Report 844/2006. Oslo: Institute of Transport Economics (TØI).

DeSerpa AJ. 1971. A theory of the economics of time. The Economic Journal 81(324):828–846.

Dixon L. 1996. Bicycle and pedestrian level-of-service performance measures and standards for congestion management systems. Transportation Research Record 1538:1–9.

Elvik R. 2000. Which are the relevant costs and benefits of road safety measures designed for pedestrians and cyclists?. Accident Analysis & Prevention 32(1):37–45.

Elvik R. 2008. Risk Information in Value of Safety Study (Working Paper SM/1962/2008). Oslo, Norway: Institute of Transport Economics (TØI).

Elvik R, Borger Mysen A. 1999. Incomplete accident reporting: Meta-analysis of studies made in 13 countries. Transportation Research Record 1665:133–140.

Elvik R, Høye A, Vaa T, Sørensen M. 2009. The Handbook of Road Safety Measures. Bingley, UK: Emerald.

Hanley N, Shogren JF, White B. 1997. Environmental Economics in Theory and Practice. New York: MacMillan.

Harkey DL, Reinfurt DW, Knuiman M. 1998. Development of the bicycle compatibility index: A level of service concept. Transportation Research Record 1636:13–20.

Hensher D, Greene W. 2003. The mixed logit model: The state of practice. Transportation 30(2):133–176.

Hensher DA, Rose JM, Ortúzar JD, Rizzi LI. 2009. Estimating the willingness to pay and value of risk reduction for car occupants in the road environment. Transportation Research Part A: Policy and Practice 43(7):692–707.

Hine J, Russel J. 1993. Traffic barriers and pedestrian crossing behaviour. Journal of Transport Geography 1(4):230–239.

Hopkinson P, Wardman M. 1996. Evaluating the demand for new cycle facilities. Transport Policy 3(4):241–249.

Jacobsen PL, Racioppi F, Rutter H. 2009. Who owns the roads? How motorised traffic discourages walking and bicycling. Injury Prevention 15(6):369–373.

Landis BW, Vattikuti VR, Brannick MT. 1997. Real-time human perceptions: Towards a bicycle level of service. Transportation Research Record 1578:119–126.

Larsen J, Patterson Z, El-Geneidy A. 2013. Build it. But where? The use of geographic information systems in identifying locations for new cycling infrastructure. International Journal of Sustainable Transportation 7(4):299–317.

Litman TA. 2003. Economic value of walkability. Transportation Research Record 1828:3–11.

Litman TA & Doherty E. 2009. Transportation Cost and Benefit Analysis: Techniques, Estimates and Implications. 2nd ed. Victoria, BC: Victoria Transport Policy Institute.

Louviere JJ, Hensher DA, Swait JD. 2000. Stated Choice Methods: Analysis and Application. Cambridge, UK: Cambridge University Press.

McFadden D, ed. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. New York: Academic Press.

Ortúzar JD, Iacobelli A, Valeze C. 2000. Estimating demand for a cycle-way network. Transportation Research Part A: Policy and Practice 34(5):353–373.

Parkin J, Wardman M, Page M. 2007. Models of perceived cycling risk and route acceptability. Accident Analysis & Prevention 39(2):364–371.

Parkin J, Wardman M, Page M. 2008. Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation 35(1):93–109.

Pikora TJ, Bull FCL, Jamrozik K, Knuiman M, Giles-Corti B, Donovan RJ. 2002. Developing a reliable audit instrument to measure the physical environment for physical activity. American Journal of Preventive Medicine 23(3):187–194.

Pikora TJ, Giles-Corti B, Bull FCL, Jamrozik K, Donovan RJ. 2003. Developing a framework for assessment of the environmental determinants of walking and cycling. Social Science & Medicine 56(8):1693–1703.

Pucher J, Dijkstra L. 2000. Making walking and cycling safer: Lessons from Europe. Transportation Quarterly 54(3):25–50.

Pucher J, Dijkstra L. 2003. Promoting safe walking and cycling to improve public health: Lessons from The Netherlands and Germany. American Journal of Public Health 93(9):1509–1516.

Ramjerdi F, Flügel S, Samstad H, Killi M. 2010. Value of Time, Safety and Environment in Passenger Transport: Time, Reliability, and Comfort (TØI Report 1053-B/2010). Oslo, Norway: Institute of Transport Economics (TØI).

Revelt D, Train KE. 2000. Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier (Working Paper Series). Berkeley, CA: University of California Institute of Business and Economic Research.

Reynolds CCO, Harris M, Teschke K, Cripton PA, Winters M. 2009. The impact of transportation infrastructure on bicycling injuries and crashes: A review of the literature. Environmental Health 8:47. doi: 10.1186/1476–069X-8-47.

Rietveld P, Daniel V. 2004. Determinants of bicycle use: Do municipal policies matter? Transportation Research Part A: Policy and Practice 38(7):531–550.

Rizzi LI, Ortúzar JD. 2003. Stated preference in the valuation of interurban road safety. Accident Analysis & Prevention 35(1):9–22.

Ruud P. 1996. Approximation and simulation of the multinomial probit model: An analysis of covariance matrix estimation. Berkeley, CA: University of California Institute of Business and Economic Research.

Samstad H, Ramjerdi F, Veisten K, Navrud S, Magnussen K, Flügel S, Killi M, Halse AH, Elvik R, San Martin O. 2010. Den norske verdsettingsstudien, Sammendragsrapport [Value of Time, Safety, and Environment in Norwegian Passenger Transport, Summary Report] (TØI Report 1053/2010). Oslo, Norway: Institute of Transport Economics (TØI).

Stanley J, Rattray A. 1978. Social severance. In: Pearce DW (ed.), The Valuation of Social Cost. London: George Allen & Unwin, pp. 140–163.

Stone M, Broughton J. 2003. Getting off your bike: Cycling accidents in Great Britain in 1990–1999. Accident Analysis & Prevention 35(4):549–556.

Tilahun NY, Levinson DM, Krizek KJ. 2007. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. Transportation Research Part A: Policy and Practice 41(4):287–301.

Train K. 2009. Discrete Choice Methods with Simulation. Cambridge, UK: Cambridge University Press.

Veisten K, Flügel S, Rizzi L, Ortúzar JD, Elvik R. 2012. Valuing casualty risk reductions from estimated baseline risk. Research in Transportation Economics 43(1):50–61.

Veisten K, Sælensminde K, Alvær K, Bjørnskau T, Elvik R, Schistad T, Ytterstad B. 2007. Total costs of bicycle injuries in Norway: Correcting injury figures and indicating data needs. Accident Analysis & Prevention 39(6):1162–1169.

Veisten K, Sælensminde K, Hagen K-E. 2005. Bicycle Injuries, Risk of Cycling, and the Tool for Cost-Benefit Analysis of Measures Towards Cycling. Oslo, Norway: Institute of Transport Economics (TØI).

Wardman M, Hatfield R, Page M. 1997. The UK national cycling strategy: Can improved facilities meet the targets? Transport Policy 4(2):123–133.

Wardman M, Tight M, Page M. 2007. Factors influencing the propensity to cycle to work. Transportation Research Part A: Policy and Practice 41(4):339–350.

Essay 4

Research in Transportation Economics 47 (2014) 50–60

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Accounting for user type and mode effects on the value of travel time savings in project appraisal: Opportunities and challenges

Stefan Flügel [a, b, *]

[a] School of Economics and Business, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway
[b] Institute of Transport Economics (TØI), Gaustadalléen 21, NO-0349 Oslo, Norway

ARTICLE INFO

Article history:
Available online 18 October 2014

JEL classification:
D61
D63
R40

Keywords:
Value of travel time savings
User type effect
Mode effect
Project appraisal
Distributional weights

ABSTRACT

Differences in Value of Travel Time Savings (VTTS) between travel modes can play a decisive role in the ranking of projects that affect the travel time of different travel modes. Conceptually, between-mode differences in VTTS can be decomposed into the user type effect (UE) that accounts for differences in characteristics of user groups (e.g. income differences) and the mode effect (ME) that accounts for differences in travel modes (e.g. the comfort level). Several studies have disentangled and quantified these two effects. However, their potential use for project appraisal has not been thoroughly discussed in the literature.

Two opportunities of using information about ME and UE in appraisal are discussed: (i) Removing the UE from national mode-specific VTTS in order to obtain a set of VTTS that only differs by the comfort level of the modes (ii) Provide the VTTS in travel modes taking into account user type effects of travellers that switch modes after project implementation.

The former arguably improves on the equity approach in project appraisal under the normative argument of valuing individual's time saving equally. The latter can improve the overall precision of user benefit representation in project appraisal compared to the standard mode-specific approach, where mode switchers are assumed to have the same VTTS in the new mode independent of which original user group they belong to.

1. Introduction

Differences in Value of Travel Time Savings (VTTS) between travel modes can play a decisive role in the ranking of projects that effect the travel time of different travel modes. In a cost benefit analytical (CBA) framework, a social planner will, *ceteris paribus*, favour projects that reduce the in-vehicle travel time of travel modes with the highest VTTS, i.e. the travel mode for which its users have the highest willingness-to-pay (WTP) for travel time savings. As VTTS differences between modes may be substantial (e.g. in Norway the VTTS in air is 204 NOK/h while it is just 74 NOK/h in long-distance bus trips (Samstad et al., 2010)), there is a possibility that certain projects are prioritized even though alternative projects safe more time in total and/or imply lower investment costs. As the general public and policy makers may regard such project ranking as opaque or unjust, it seems important to understand and possibly address the origin of between-mode differences in VTTS in order to define optimal standards for project appraisal.

Mode-specific VTTS differ not only because modes themselves differ (e.g. in the level of comfort), but also because the travel modes' user groups differ in characteristics (e.g. average income level). This difference in characteristics is due to self-selection of travellers to travel modes. Wealthy travellers are, for example, more likely to drive their own car (as they can afford to own one), while travellers with high preference for time savings (again, typically wealthy persons) are likely to choose air instead of slower modes like bus. Conceptually, between-mode differences in VTTS can be decomposed into mode and user type effects (Wardman, 2004).[1] User type effects (UE) are empirically obtained by the differences in VTTS of two user groups in a given travel mode, while mode effects (ME) are differences in VTTS in two modes for the same user

* School of Economics and Business, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway. Tel.: +47 41520576.
E-mail address: sfl@toi.no.

[1] Wardman actually called the ME "mode valued effect" which might be more precise. Other terms used to describe the same effect are "pleasantness-effect" (Mackie et al., 2001) and "comfort effect" (Fosgerau et al., 2010).

http://dx.doi.org/10.1016/j.retrec.2014.09.018

group. Several studies have disentangled and quantified these two effects (e.g. Fosgerau, Hjorth, & Lyk-Jensen, 2010; Ramjerdi, Flügel, Samstad, & Killi, 2010; Ramjerdi, Sælensminde, Rand, & Sætermo, 1997), however, their potential use for project appraisal has not been thoroughly discussed in the literature.

Three CBA approaches regarding the degree of segmentation of VTTS by mode and user group can be distinguished.[2] This is not an exhaustive list, but rather an illustrative one, with two contrary approaches and one compromise approach.

1) The so-called *equity approach* that applies the same VTTS for all travel modes and user groups in a population. An equity VTTS is hard to justify through economic reasoning and is likely to yield to a misallocations of resources as it will seldom match the actual WTP of travellers affected by a specific project. However, the approach has its normative justification under the political paradigm of not taking into account individual differences in applied CBA. This approach is for example currently used in Denmark.
2) The *project-specific approach* that tries to account as precisely as possible for the WTP of the actual beneficiaries of the project. It is in clear contrast to the equity approach as it includes mode specific VTTS, geographical segmentation in VTTS[3] and opens up for specific valuation studies among the actual users of the infrastructure provided by the project. The Official Norwegian Report on cost benefit analysis has recently suggested that national values should only be used when the information about the VTTS of actual users is insufficient (NOU, 2012, page 12). While this approach may appear theoretical appealing, there are practical limits to investigating project-specific VTTS.
3) The *(standard) mode specific approach* uses national VTTS values but it segments the market by travel modes. This approach might be seen as a practical compromise between the equity and the project specific approach, but arguably it lacks good conceptual reasoning. As it ignores direct user group segmentation (e.g. by geography) it might be defended under an equity argument. However, travel mode segmentation accounts indirectly for traveller's characteristics and preconditions because of the above mentioned self-selection. Note that this approach accounts for UE and ME by their sum but does take into account information about their relative sizes.

In this paper two opportunities for using information about ME and UE in order to define alternative VTTS approaches in project appraisal are discussed.

(i) *ME-dependent equity value*, that removes the UE from (standard) mode specific VTTS values, to obtain a set of VTTS values that only differs by the comfort level of the modes (i.e. differences due to the ME). This arguably improves on the equity approach while remaining consistent with the normative argument of valuing every individual's time saving (of same size) equally.
(ii) *VTTS of switching modes* provides VTTS values for situations where travellers switch travel modes. Knowledge about UE and ME enables a differentiation of VTTS by user groups defined by the (expected) choice of mode before and after project implementation. In principle this improves the overall precision of user benefit representation in project appraisal compared to the standard mode-specific approach, where mode switchers are assumed to have the same VTTS in the new mode independent of which original user group they belong to.

None of the two outlined implementations is without methodological and practical challenges, as will be discussed in the paper.

The paper provides definitions and theoretical background in Section 2. Section 3 describes ways to improve current VTTS approaches by incorporating knowledge about UE and ME in project appraisal. Section 4 gives an empirical example and illustrates the possible impact of different VTTS approaches on project ranking by means of a stylized case study. Section 5 concludes.

2. Definitions and theory

2.1. Mode- and user type effects

2.1.1. Definitions

For a formal definition of the two effects, the differences in the representative VTTS[4] between mode k and l with associated current user groups g_k and g_l are as set out below:

$$\Delta VTTS_{k,l} \equiv VTTS_{k,g_k} - VTTS_{l,g_l} = VTTS_{k,g_k} - VTTS_{l,g_k} + VTTS_{l,g_k} - VTTS_{l,g_l} \tag{1}$$

The mode effect (ME) is defined as the VTTS difference between two modes for a given user group:

$$ME_{k,l,g_k} \equiv VTTS_{k,g_k} - VTTS_{l,g_k}, \tag{2}$$

The user-type effect (UE) is defined as the VTTS differences of two user groups in a given travel mode:

$$UE_{l,g_k,g_l} \equiv VTTS_{l,g_k} - VTTS_{l,g_l} \tag{3}$$

Then, one can write between-mode differences in VTTS as the sum of the two effects:

$$\Delta VTTS_{k,l} = ME_{k,l,g_k} + UE_{l,g_k,g_l}. \tag{4}$$

Furthermore, assume that every individual q in the population can be associated with one user group g_q. Typically groups are identified by the current mode choice of q but other segmentation rules may be applied. We define the average mode effect over all groups g_q in a population as:

$$\overline{ME_{k,l}} \equiv \frac{1}{N} \sum_{g_q} N_{g_q} ME_{k,l,g_q}, \tag{5}$$

with N_{g_q} being the amount of members in group g_q and N being the total amount of individuals in the population.

For later discussion we define an average VTTS in a transport mode, say k, as the weighted average of the representative VTTS in mode k over all groups in the population.

[2] This paper does not discuss other segmentation of VTTS, e.g. by trip purpose, trip length or travel time components. When not stated differently, the VTTS of a travel mode relates to the in-vehicle time.

[3] The VTTS in cities is often found higher than in rural areas (see Abrantes & Wardman, 2011; Börjesson & Eliasson, 2014; and Østli, Halse, & Ramjerdi, 2012 for empirical evidence from the UK, Sweden and Norway).

[4] The representative VTTS will in application typically be inferred as some mean value derived from a sample of travellers in the groups. In this section we are only making heterogeneity across groups explicit. Treating heterogeneity within groups is subject of discussion in later sections.

$$\overline{VTTS}_k = \frac{1}{N}\sum_{g_q} N_{g_q} VTTS_{k,g_q}. \tag{6}$$

The VTTS in (6) can be seen as the representative VTTS for the whole population *in a given transport mode*. It is a modification of the "equity VTTS" ($\overline{VTTS}$) and is therefore referred to as the "ME-dependent equity value".

Note that the difference between the VTTS of two transport modes calculated by (6) will equal the average mode effect (in 5).

$$\overline{VTTS}_k - \overline{VTTS}_l = \overline{ME}_{k,l}. \tag{7}$$

User groups may not only be defined by their current travel mode as in (1)–(4), but also by subgroups identified by their previous transport mode choice. We use the notation $g_{k|l}$ to identify an user group that consists of travellers having switched from mode k to mode l. Travellers that already previously have chosen l ("non-switchers") are labelled $g_{l|l}$.

We define the resulting VTTS in a transport mode l subject to mode switches of travellers (referred to as *VTTS of switching modes*, $VTTS_l^{S.M.}$) as the weighted average of different user groups now using l.

$$VTTS_l^{S.M.} = \frac{1}{N_l}\left(N_{g_{l|l}} VTTS_{l,g_{l|l}} + \sum_k N_{g_{k|l}} VTTS_{l,g_{k|l}}\right) \tag{8}$$

with $N_l = N_{g_{l|l}} + \sum N_{g_{k\_l}}$, where $N_{g_{l|l}}$ is the number of non-switchers and $N_{g_{k|l}}{}^k$is the number of travellers that are transferred from different travel mode k.

$VTTS_l^{S.M.}$ will in general differ from $VTTS_{l,g_l}$, i.e. the representative VTTS for the generic users of transport mode l. Intuitively the difference is connected to user type effects.

Indeed, defining a weighted average of user type effects between different subgroups identified by their switching behaviour ($g_{k|l}$) and a generic user group (g_l) as

$$\overline{UE}_l^{S.M.} \equiv \frac{1}{N_l}\left(N_{g_{l|l}} UE_{l,g_{l|l},g_l} + \sum_k N_{g_{k|l}} UE_{l,g_{k|l},g_l}\right) \tag{9}$$

it can be shown that:

$$VTTS_l^{S.M.} - VTTS_{l,g_l} = \overline{UE}_l^{S.M.}. \tag{10}$$

2.1.2. Theoretical components and interpretation

The VTTS of individual q can be derived from time allocation models (Becker, 1965; DeSerpa, 1971; Jara-Díaz & Guevara, 2003). From one of the early models (Oort, 1969) the following decomposition of VTTS is obtainable.[5]

$$VTTS_q = w_q + \frac{MU_q^W}{\lambda_q} - \frac{MU_q^T}{\lambda_q}. \tag{11}$$

Hence, the VTTS is the sum of the nominal wage rate (w_q) and the marginal utility of work time (MU_q^W), minus the marginal utility of travel time (MU_q^T). The latter two terms are normalized by λ_q, the marginal utility of income, in order to transfer utility onto a monetary scale.

MU_q^T expresses the direct utility (disutility in most cases) from spending time in transport; it is likely to depend on the specific travel mode (in general $MU_q^{T_k} \neq MU_q^{T_l}$).

The expression $\left(w_q + \frac{MU_q^W}{\lambda_q}\right)$ represents the total value of work and can be seen as the opportunity cost of travelling (OCT), i.e. the value which could have been obtained from activities other than travelling.[6] The OCT is independent of the travel mode ($OCT_q^k = OCT_q^l$) but differs across individuals and thereby also, on average, between user groups ($OCT_{g_k} \neq OCT_{g_l}$).

The VTTS in travel mode k of group g_k (current users of travel mode k), can be expressed as:

$$VTTS_{k,g_k} = OCT_{g_k} - \frac{MU^{T_{k,g_k}}}{\lambda_{k,g_k}}. \tag{12}$$

where $OCT_{g_k}, MU^{T_k,g_k}, \lambda_{k,g_k}$ are representative values for the user group g_k.

Combining (12) and (2), the mode effect can be written as:

$$ME_{k,l,g_k} = -\frac{1}{\lambda_{g_k}}\left(MU^{T_k,g_k} - MU^{T_l,g_k}\right). \tag{13}$$

The ME will therefore depend on the difference in direct disutility obtained from time spent in the two travel modes (scaled by the marginal utility of income of the given user group). The direction (sign) of the ME will indicate which travel mode offers a more useful/pleasant travel time and can be regarded as more "comfortable" in a wide sense. This might relate to the actual comfort of seats but also the possibilities to read, sleep, eat, work, use entertainment devices etc., as well as the perceived level of safety. λ_{g_k} will have an effect on the absolute size of ME but not on its sign. A user group with high average income (implying lower λ_{g_k}) will have higher ME even when the perceived "comfort difference" of modes is identical to the perception of user groups with lower income.

Combining (12) with (3) we obtain:

$$UE_{l,g_k,g_l} = \left(OCT_{g_k} - OCT_{g_l}\right) - \left(\frac{MU^{T_l,g_k}}{\lambda_{g_k}} - \frac{MU^{T_l,g_l}}{\lambda_{g_l}}\right). \tag{14}$$

The UE will depend on the differences in representative opportunity costs in the two user groups and the (scaled) differences of the marginal utility of time spent in the given travel mode.

UE is due to self-selection as traveller characteristics will influence both the VTTS (by Equation (11)) and the choice of travel modes. Wealthy persons are more likely to self-select to expensive travel modes, while busy persons (high OCT) are likely to use fast travel modes. Self-selection also works through different perceptions of comfort level in the chosen mode compared to the discarded travel modes across user groups (e.g. train users might not drive as they have higher preferences for taking a nap than the average car user).

2.2. Project appraisal

2.2.1. Welfare effect of travel time savings

In general terms, projects should be compared (ranked) by their effect on social welfare, which is assumed to be expressible in terms of utility values of all individuals q.

[5] While VTTS equals just the wage rate in Becker's original model, most of the later proposed time allocation models, where the time spend in activities enter utility directly, obtain decompositions of VTT similar to the one by Oort (1969); see Jara-Díaz (2007, page 68,69) for an overview.

[6] This would also include leisure time which equals the total value of work under the assumption that decision makers can freely assign time to activities that are remunerated.

$$W_s = W_s(U_1, \ldots, U_q, \ldots, U_N) \tag{15}$$

Following Gálvez and Jara-Dıaz (1998) in a compact and generally applicable specification, individual utility is a function of goods X_{iq} which again is a function of generalized income I_q and prices P.

Then, the welfare change due to time savings (or losses) is given by:

$$dW_s = \sum_q \frac{\partial W_s}{\partial U_q} \frac{\partial U_q}{\partial I_q} \frac{\partial I_q}{\partial t_q} dt_q = \sum_q \mathbf{\Omega_q} \lambda_q dB_q \tag{16}$$

where

- $\Omega_q = \frac{\partial W_s}{\partial U_g}$ is the (normative) social weight, expressing how much an utility unit of individual q contributes to social welfare
- $\lambda_q = \frac{\partial U_q}{\partial I_q}$ is the marginal utility of income of individual q
- $dB_q = \frac{\partial I_q}{\partial t_q} * dt_q$ is the user benefit from time savings, i.e. the monetary value of q's consumer surplus variation (in a sense of Hicksian compensating variation, see e.g. Jara-Díaz, 2007, page 99).

dB_q is approximately given as $VTTS_q \Delta t_q$, so that (16) can be written as (Gálvez and Jara-Dıaz, 1998; Mackie, Jara-Dıaz, and Fowkes (2001)):

$$\Delta W_s = \sum_q \Omega_q \lambda_q VTTS_q \Delta t_q. \tag{17}$$

Hence, the welfare change of an individual's travel time savings is due to the size of the time saving multiplied by the (subjective) VTTS, scaled by the marginal utility of income and multiplied by a normative weight set by the social planner.

2.2.2. Cost Benefit Analysis (CBA)

The standard approach in CBA is the "willingness-to-pay (WTP) calculus" (Sugden, 1999), where the benefits of travel time saving projects (here without cost and taxes) are represented by the sum of individual WTP:

$$\Delta W_s^{CBA} = \sum_q dB_q = \sum_q VTTS_q \Delta t_q \tag{18}$$

This equation can be rewritten as:

$$\Delta W_s^{CBA} = \sum_{g_q} \sum_{q \in g_q} \frac{VTTS_q}{\overline{VTTS_{g_q}}} \overline{VTTS_{g_q}} \Delta t_q, \tag{19}$$

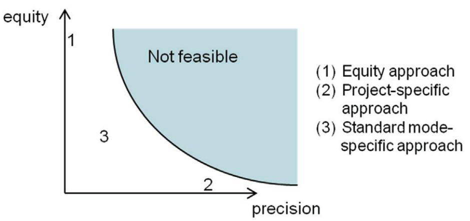


Fig. 1. Illustration of VTTS-approaches on the equity and precision dimension.

where $\overline{VTTS_{g_q}}$ is the average VTTS of the group g individual q belongs to. In applied CBA the factor $\frac{VTTS_q}{\overline{VTTS_{g_q}}}$ is omitted (i.e. the heterogeneity in VTTS within groups is not regarded):

$$\Delta W_s^{appl.CBA} = \sum_{g_q} \sum_{q \in g_q} \overline{VTTS_{g_q}} \Delta t_q = \sum_{g_q} \overline{VTTS_{g_q}} \Delta t_g \tag{20}$$

with $\Delta t_g = \sum_{q \in g_q} \Delta t_q$ being the aggregated (net) time savings of group g.

When talking about *distributional* weights in CBA one is usually referring to a factor α_q or α_{g_q} put as a weighting factor on individual or group specific user benefits;

$$\Delta W_s^{w.CBA} = \sum_q \alpha_q VTTS_q \Delta t_q, \tag{21}$$

or in applied studies

$$\Delta W_s^{w.appl.CBA} = \sum_{g_q} \alpha_{g_q} \overline{VTTS_{g_q}} \Delta t_g. \tag{22}$$

Comparing (21) with (17) we see that α_q corresponds to $\Omega_q \lambda_q$ and is therefore a combination of a (normative) social weight and the marginal utility of income.

As pointed, out by Gálvez and Jara-Dıaz (1998), CBA standards that set $\alpha_q = 1$ correspond to a situation where Ω_q are set to $1/\lambda_q$, i.e. where social weights are inversely proportional to the marginal utility of income. As the marginal utility itself is decreasing with income, CBA standards imply higher weights for wealthy persons when transferring utility to wealth. Or, framing it differently, unweighted CBA ignores differences in the marginal utility of income among individuals when transferring WTP to wealth.

With a positive economic perspective (implying $\Omega_q = 1$), the classical argument is that the un-weighted sum of actual WTP should be used in project appraisal and that redistribution of welfare should be done by means of a (second-best) income tax (given that lump sum transfer from "winners" to "losers" are not feasible). It is argued that distributional weights in CBA will lead to efficiency losses compared to a redistribution of wealth via the income taxes (Harberger, 1978). Johansson-Stenman (2005) questions this proposition and shows cases (i.e. model assumptions) for which factors α_q (being equivalent to λ_q given $\Omega_q = 1$) are called for to account for the fact that poor persons profit more from marginal decreases in income tax. Considering that the appropriateness of model assumptions, the means of redistributing wealth and the sources of funding[7] differ from project to project it seems difficult to conclude generally about the efficiency of weights in CBA.

Comparing (22) with (19), it is evident that taking averages of groups (the VTTS segmentation in user groups) has a likely distributional element. The most extreme case, i.e. the equity approach of VTTS that takes a grand average of VTTS over all persons.

$$\Delta W_s^{appl.CBA,equity} = \overline{VTTS} \sum_{q \in N} \Delta t_q, \tag{23}$$

implies in standard CBA (α_{g_q}=1) that $\Omega_q = \frac{VTTS_q}{\lambda_q \overline{VTTS}}$. This expression resembles a strong normative weight and is likely to yield to a misallocation of resources in an economic sense. For instance, user benefits from travel time savings of a transport project that accrue

[7] See Börjesson and Eliasson (2012, 2014) for a discussion on the importance of funding source for transport project appraisal.

mainly to travellers with above-average VTTS $\left(\frac{VTTS_q}{VTTS_{gq}} > 1\right)$ will be underestimated with the equity approach.

3. Concepts of accounting for ME and CE in project appraisal

3.1. Directions of improving VTTS-approaches in project appraisal

Before outlining ways to account for UE and ME in project appraisal, this section is intended as a general discussion on directions to improve current VTTS-standards in project appraisal.

The three current VTTS-approaches introduced in Section 1, i.e. (1) equity approach, (2) project-specific approach and the (3) standard mode specific approach can be compared by the two dimensions precision and equity. A schematic illustration is given in Fig. 1.

Precision is referred to as the degree to which a VTTS-approach is able to account for the actual VTTS of travellers. A VTTS-approach that multiplies individual time savings with the actual VTTS for every traveller (without taking any averages) would "score" highest. A VTTS-approach would score lower when it takes averages over heterogeneous travellers. Precision would also be lower when factors influencing VTTS are ignored in the derivation of VTTS (e.g. if the approach would ignore the mode effect or the user type effect). From the three bespoken VTTS-approaches, the project specific approach has in general the highest precision as it does not apply national averages of VTTS but strives to account for the VTTS of the particular users of a transport project. The standard mode specific approach is more precise than the equity approach as it accounts for the mode- and the user type effect (while the equity approach ignores both effects).

Equity is referred to the degree to which VTTS-approaches imply equal VTTS for travellers in the population. The equity VTTS approach scores highest here as it applies the same VTTS for all possible travellers. The standard mode specific approach scores lower on equity because different travellers will be assigned different VTTS according to their transport mode choice which in many cases will be due to self-selection (such that rich travellers get assigned a higher VTTS on average). The project specific approach is the least equitable of the three as it applies different VTTS values for different locations/projects in the nation.

Obviously a VTTS-approach cannot be very precise and equitable at the same time. However, it might be possible to find improvements on either of the two dimensions. For instance, an equity approach that calculates a faulty average over the population can be improved on the precision dimension by replacing it by the more precise one. In that case the level of equity is fully retained while precision is improved. Improving equity while retaining the existing level of precision seems hard to achieve, but it might be possible to improve a lot of the equity dimension, without losing much precision (see discussion later).

Which VTTS-approach is most appropriate/preferable for project appraisal does obviously depend on how the social welfare function is specified. For instance, assuming a welfare function in line with normative weights of $\Omega_q = \frac{VTTS_q}{\lambda_g VTTS}$ (compare above), the equity approach is preferable. In this regards, it has to be underlined that equity and precision in Fig. 1 refer to VTTS (and user benefits given an objective measure of Δt_q) and not to welfare units. One should always be interested in a most precise representation of the specified welfare function for project appraisal.

The objective to score high on the precision dimension is straightforwardly motivated in standard CBA ($\alpha_q = 1$) and from a positive economic perspective ($\Omega_q = 1$). Each imprecision in the user benefit representation can bias the project ranking and can lead to a misallocation of resources in an economic sense.

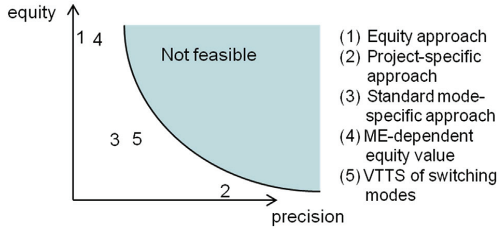


Fig. 2. Illustration of VTTS-approaches (including the two proposed approaches) on the equity and precision dimension.

The objective to score high on the equity dimension can be motivated by the general notion of fairness.[8] Indeed, the underlying motive for the VTTS equity approach seems to be the alleged tendency of travel modes with high estimated VTTS (usually car and air) to be more frequently used by wealthier persons. In this sense the equity approach is a radical way to eliminate any between mode differences in VTTS due to traveller's characteristics and preconditions. In the Danish Value of Time Study the following explanation for the choice of the equity approach is found: "In line with the discussion concerning income differences, the steering group for the project has expressed the view that the cost-benefit analysis will be considered most relevant by policy makers if the analysis treats everybody equally. It has therefore been decided to use the grand average of 67 DKK per hour as the central value to be applied to all transport modes (Fosgerau, Hjorth, & Lyk-Jensen, 2007, page 19/20). Hence, besides the general fairness arguments (e.g. that wealthier people should not get assigned higher VTTS just because they are wealthier), also a presumed gain in political acceptance of CBA seems to motivate this approach. For instance a distinct geographical segmentation of VTTS is likely to be difficult to implement and maintain politically in case of resistance by stakeholders from disadvantaged regions. Even ethnical reasons might be put forward to motivate the equity approach, for instance by referring to the value of statistical life (VSL) in CBA-calculus for which it is endorsed to use a common value for all travellers (as if not CBA-calculus might imply the saving a life of an affluent persons is worth more for society than saving a life from a poor person). For VTTS the ethical argument seems less convincing but the principle that every individual unit of time saving should be worth the same for society (independent of the characteristics and preconditions of the persons by whom the time savings are experienced) is hard to discard on the basis of lacking valid normative reasoning.

3.2. ME-dependent equity values

Assume that a social planner's normative credo is that every individual unit of time saving should contribute equally to the total benefit of transport projects, but that he is critical of a single equity VTTS because he acknowledges that utility of spending time in travel modes will differ by the travel modes' comfort level. In this case, empirical information about UE and ME will enable him to correct for the user type effect and to obtain a set of VTTS that varies

[8] There exists a rich literature of the concept of fairness (or justice) in political philosophy. The equity approach seems well compatible with the "difference principle" (Rawls, 1971, 2001) arguing that economic inequality should be arranged such that it profits the least-advantaged members of society.

between modes only on account of the mode effect (i.e. the travel mode's comfort level in a wide sense).

A direct way to achieve this is to calculate the "ME-dependent equity value" defined in (6) as the weighted average of the VTTS in a given transport mode over all user groups in the population. It can be seen as an improvement of the equity approach on the precision dimension as it includes information about the mode effect (going from 1 to point around 4 in Fig. 2 below). From another perspective, it can also be seen as an improvement over the standard mode specific approach on the equity dimension as this VTTS can be thought of as being controlled for self-selection (going from 3 to 4 in Fig. 2 below).This is because all possible travellers are included in the calculation independent of their mode choice. The ME-dependent equity value has a lower precision compared to the standard mode specific approach (as the UE is obviously important for a precise VTTS representation), but the ME-dependent equity value retains the normative credo of "value every individual unit of time saving equally". It comes in the modified version of: "value every individual unit of time saving *in a given transport mode* equally".

In practical applications one will often lack information about a travel mode's VTTS for at least some user groups. In the Norwegian Value of Time Study the design of the stated preference (SP) study includes only choice experiments for which the VTTS in car is possible to estimate for current car user and for every respondent that has car as his (first-choice) alternative mode (i.e. a subsample of all user groups). Hence, some self-selection is likely to remain. While this is an empirical problem which can be overcome by stated preference (SP) studies with a sufficient number of observations in each user group, there is also a conceptual point of discussion, namely if all citizens or just the potential users should be included in the calculation. This can be seen as a normative question, which may be illustrated by the example of certain affluent persons never taking the bus (they would rather take a cab). If this group of affluent persons would be left out of Equation (6), one would underestimate $\overline{VTTS}_{bus}$ and one would not have gotten entirely rid of the self-selection problem. Therefore, departing from the normative credo of the equity approach, it is recommended to try to include as many user groups as possible, so as to establish VTTS values that only differ by the mode effect. Then also λ_q, impacting the absolute size of ME as shown in Equation (13), can be thought of being representative for all citizens.

3.3. VTTS of switching modes

Some travel time saving projects will have a notable effect on traveller's mode choice. In usual CBA practice, travellers that shift from mode k to l are assumed to shift VTTS from $VTTS_{k,g_k}$ to $VTTS_{l,g_l}$. As this change in VTTS is due to the sum of ME and UE, individuals shift not only travel mode but are also treated as they shift their characteristics influencing UE, e.g. the income level. Obviously, this assignment is restrictive and Mackie et al. (2001) claims even that "[u]ntil the income effect [UE] can be properly disentangled from the 'pleasantness' effect [ME], there is more to be lost than gained from subdividing in-vehicle time savings by modes" (page 102).

Given knowledge about the ME of user group g_k one can express the change in VTTS after mode shift as going from $VTTS_{k,g_k}$ to $VTTS_{l,g_k}$. If all travellers of a mode k were forced to switch mode to l, $VTTS_{l,g_k}$ would indeed be the value to consider. However, mode choice is usually voluntary so that there is a self-selection between travellers that switch modes and those who choose not to. The correct VTTS to use is therefore$VTTS_{l,g_{k/l}}$, where subscript k/l indicates the user group defined by the former mode k and the new mode l. The change in VTTS is due to user type effects within current users ($VTTS_{k,g_k}$ to $VTTS_{k,g_{k/l}}$) and the mode effect of user group$g_{k/l}$($VTTS_{k,g_{k/l}}$ to $VTTS_{l,g_{k/l}}$).

Given information about the different $VTTS_{l,g_{k/l}}$ and the expected number of mode shifts after project implementation it is then possible to calculate a (project specific) VTTS in mode l as the weighted average of VTTS over all user groups now using l; see (8) in Section 2.1.1. for the mathematical expression.

In principle, the VTTS of switching modes ($VTTS_l^{S.M.}$) improves over the (standard) mode specific VTTS ($VTTS_{l,g_l}$) on the precision dimension (moving from 3 to point around 5 in Fig. 2 above) as it takes into account the new composition of travellers with their characteristics underlying the user type effect However this implies that one is able to get precise estimates of the values of $VTTS_{l,g_{k/l}}$. For this, current users of all transport mode k that will switch to transport mode l must be identified and their VTTS in the new transport mode l must be estimated. In principle, SP techniques would be susceptible to provide such information on a project-to-project basis, given that respondents are truthfully telling and revealing both their (intended) switching behaviour as a result of a concrete project and their VTTS in the new travel mode. However there is a potential danger of strategic behaviour on the part of respondents directly affected by the project. Thus such estimates might lack validity.[9] Another issue is that SP-studies are costly and many project-specific SP-studies will have a small sample if they are warranted at all. This can easily lead to inconsistencies between SP-studies conducted for different projects.

In absence of project specific VTTS estimates, one faces a conceptual challenge in how to obtain good proxies of the different user groups based on national studies. Using data from the Norwegian Value of Time study, one assumes that travellers switching from k to l coincide with respondents having l as a (first choice) alternative mode to k. This may or not may be a good approximation (obviously depending on the specific project).

Note also that when national values (usually obtained not more than every 5–10 years) are used as an approximation, then one should idealistically define user groups by their sequence of mode shifts over consecutive projects. However it seems difficult (practically impossible) to keep track of the composition of different user groups. The VTTS of switching modes is therefore rather a short term concept of the evaluation of user benefits.

Some transport model systems are capable of predicting mode choice behaviour and user benefits (by log-sums) simultaneously. Here the VTTS relevant for project appraisal (or parameters underlying VTTS) are included as explanatory variables in the mode choice model (i.e. they are known before mode choice is predicted). This is however not possible for the concept of *VTTS of switching modes* as its value depends on the predicted mode choice (weighted averages over user groups defined by their switching behaviour are used in Equation (8)). This methodological problem seems hard to resolve. However, common practice in many countries (also in Norway) is to calculate user benefits in CBA sequentially after the mode choice is predicted. In this case, *VTTS of switching modes* can be calculated given reliable information about the number of switchers from different travel modes and their values of travel time savings.

4. Application

4.1. Estimation

Although it is possible to estimate the VTTS from time allocation models (Jara-Díaz, Munizaga, Greeven, Guerra, & Axhausen, 2008),

[9] For the national VTTS studies in Denmark and Norway, where choice experiments are framed independent of specific project, no empirical evidences for strategic behaviour could be found (Flügel et al., 2011; Fosgerau et al., 2010).

the clear majority of studies that estimate VTTS is based on discrete choice data, either of the revealed preference (RP) or stated preference (SP) type.[10] To disentangle ME and UE one needs the VTTS for one user group in different travel modes and the VTTS in one travel mode for different user groups (compare definitions (2) and (3)). Cross-sectional data with just one choice observation per decision maker will in general not provide this information. An experimental design used in recent SP-studies that disentangle ME and UE, is to let respondents go through two sequences of route choice tasks; one in their current mode and one in their (first-choice) alternative mode (Fosgerau et al., 2007; Ramjerdi et al,. 2010).[11] In those studies, route alternatives were characterized just by two attributes, travel time and travel cost, facilitating the "integrated approach" to VTTS estimation (Fosgerau, Hjorth, & Lyk-Jensen, 2006) in which VTTS can directly be parameterized with covariates. Fosgerau et al. (2010) specified the logarithm of VTTS of respondent n in choice task t as:

$$\log VTTS_{nt} = \beta' \boldsymbol{X}_{nt} + \delta' \boldsymbol{D}_{nt} + u_n \tag{24}$$

where u_n is the normally distributed, person specific random term, X_{nt} are socio-economic or design variables for respondent n in choice task t, and D_{nt} are dummy variables representing respondent n's current and alternative mode choice and whether or not t refers to a choice task in n's current or alternative mode. The estimated relative sizes of elements in δ' give information about user type and mode effects (see details in Fosgerau et al., 2010). When using socio-economic variables in X_{nt} it is important to realize that ME and UE are statistically controlled for the effects of these variables, so that the full size of the ME and UE will not be obtained (Flügel, Hjorth, & Ramjerdi, 2011). For instance, Fosgerau et al. (2010) controlled for income and several background variables. In this case it is likely that the size of the UE is reduced.[12] A great advantage of the integrated approach is that one can control for design variables (SP artefacts) such as the absolute size of the travel time saving (Δt), which often has a direct effect on empirical VTTS estimates.[13] This makes the estimation of UE and ME more consistent and valid.

4.2. Empirical example

This section discusses some empirical evidence based on data from the Norwegian Value of Time study (Ramjerdi et al., 2010). The subsample of private long distance trips (>100 km) within Norway with travel modes car, bus (coach), rail and air is considered. For interpretation of the mode effect, two elements are important: (1) only car drivers (not passengers) are included (drivers are likely to perceive relatively more discomfort due to lack of sleeping and reading possibilities) and (2) only in-vehicle time is considered, although for the air mode, the time spent at the airports is included.

The model specification does not control for socio-economic variables, so that the full size ME and UE are obtained. Table 1 gives relative values as estimated from the integrated estimation model described above. Groups g_{car}, g_{bus} g_{rail} and g_{air} consist of respondents that were not routed into the alternative mode choice experience. As the routing was random, these groups are regarded as the representative user group of the corresponding user modes. For the other groups, the first travel mode in the subscript indicates the current mode and the second transport mode the alternative mode.

Comparing line by line one obtains the ME; comparing column by column one obtains the UE. One can distinguish between user type effects across and within current users (UEa and UEw). Comparing e.g. $VTTS_{car,g_{car/bus}}$ with $VTTS_{car,g_{bus/car}}$ gives the UEa (current car versus current bus users), while comparing e.g. $VTTS_{car,g_{car/bus}}$ with $VTTS_{car,g_{car/rail}}$ gives UEw (alternative mode bus versus alternative mode rail). UEa is due to self-selection to current modes, while the UEw effect is due to self-selection to alternative modes (Flügel et al., 2011).

The 95% confidence intervals in Table 1 are rather broad and overlapping in most cases. However, there is a clear pattern of relative VTTS across user groups and across modes, which is easier to spot when scaling and rearranging VTTS in "pairs of user groups" (Table 2).

As a reading example: For the $g_{car/bus} - g_{bus/car}$ pair we see that current car users (with alternative mode bus) have a VTTS in car of 19.7 €/h, while current bus users (with alternative mode car) have a VTTS in bus of 12.8 €/h. Hence $\Delta VTTS_{car,bus}$ is 6.9 €/h. As shown in Equation (4) this change can be associated with UE (here UEa) and ME. The UEa is −4.3 and indicates that current car users have higher opportunity cost than bus users (income differences and self-selection); the ME of −2.6 tells us that sitting in bus is more comfortable (in a wide sense) than driving a car.[14]

Studying the direction of the ME, the following consistency over all user groups is observed: Time in mode air (including waiting times at airports) is perceived least "comfortable" and car (as a driver) is less comfortable than bus and rail (the differences between rail and bus being quite small). Studying the direction of the UEa it is evident that car drivers (bus passengers) have higher (lower) VTTS than other current users in any given travel mode.

Another way of looking at the estimation results from Table 1 is to regard the VTTS of the alternative mode as the value which would be most applicable for user group after switching travel modes. Table 3 gives the values which may be established under this assumption.[15]

As a reading example: Current rail users - that have a representative VTTS of 14.3 €/h[16] – have a VTTS of 22.2 €/h in their alternative mode air. The difference of 7.9 €/h can then be thought of being the best estimate (with the data available) for the change in VTTS after a mode switch from rail to air. The differences is decomposable into a UEw of 1.3 € (current rail users that (can

[10] SP studies are often used for the specific purpose of VTTS estimation. Among VTTS-SP-studies, route choice experiments are in general preferred over mode choice experiments as the alternative specific constants in mode choice models are likely to capture some of the "comfort effect", which should conceptually be associated with the disutility of travel time.

[11] This approach was earlier used in studies described in Algers, Dillen, and Widlert (1996) and Ramjerdi et al. (1997).

[12] See Flügel et al. (2011) for some empirical tests regarding the influence of controll variables on the estimation of UE and ME.

[13] In Flügel et al. (2011), the elasticity of Δt on VTTS was estimated at around +20%. Given that this is a real phenomenon and not just an SP specific finding, it would also have an effect on the appraisal of projects that provide different absolute sizes of travel time savings. However, this issue is not further discussed in this paper.

[14] The direction of the mode effect between bus and car might be a particularity of long distance travel. Fosgerau et al. (2010) finds evidence that the reversed direction of ME between car and bus. The Danish study does not segment in long and short distance and presumably most bus trip are short-distance, where busses are often crowded.

[15] Similar tables were previously reported in the Norwegian Value of Time Study (Flügel & Minken, 2011; Ramjerdi et al., 2010; Samstad et al., 2010). There the 'official' VTTS for each travel mode was used to calculate the VTTS for switching modes. The results in this table are based on results in Table 1 scaled the 'official' VTTS of car. The average values for other mode than car do not equal the official values in Norway. However, the ordinal ranks of VTTS are the same ($VTTS_{air} > VTTS_{car} > VTTS_{rail} > VTTS_{bus}$).

[16] This is the VTTS in rail for group g_{rail} in Table 1 multiplied by 18.75 to convert it into absolute numbers.

Table 1
Relative Value of Travel Time Saving (VTTS) with 95% confidence intervals based on Flügel et al. (2011, page 16[a]).

User group	Relative VTTS in …			
	Car	Bus	Rail	Air
g_{car}	1.00 (normalized)			
$g_{car/bus}$	1.050 (0.890–1.239)	0.905 (0.767–1.067)		
$g_{car/rail}$	1.168 (0.998–1.366)		1.004 (0.859–1.174)	
$g_{car/air}$	1.105 (0.757–1.613)			1.927 (1.318–2.819)
g_{bus}		0.627 (0.548–0.717)		
$g_{bus/car}$	0.821 (0.660–1.022)	0.682 (0.511–0.843)		
$g_{bus/rail}$		0.607 (0.471–0.782)	0.601 (0.462–0.782)	
$g_{bus/air}$		0.530 (0.376–0.748)		0.934 (0.653–1.337)
g_{rail}			0.763 (0.668–0.873)	
$g_{rail/car}$	1.061 (0.879–1.280)		0.944 (0.784–1.137)	
$g_{rail/bus}$		0.724 (0.590–0.888)	0.712 (0.586–0.887)	
$g_{rail/air}$			0.830 (0.678–1.017)	1.183 (0.962–1.455)
g_{air}				1.384 (1.231–1.556)
$g_{air/car}$	1.016 (0.836–1.235)			1.412 (1.169–1.706)
$g_{air/bus}$		0.597 (0.432–0.843)		1.145 (0.810–1.616)
$g_{air/rail}$			0.647 (0.544–0.769)	1.048 (0.885–1.242)

[a] Additional user groups that rejected the offered alternative mode are not displayed here.

Table 2
Value of Travel Time Savings, mode effects and user type effects, Comparison of "pair of user groups".

In € per hour[a]	VTTS in car	VTTS in bus	ME		VTTS in car	VTTS in rail	ME		VTTS in car	VTTS in air	ME
$g_{car/bus}$	19.7	17.0	−2.7	$g_{car/rail}$	21.9	18.8	−3.1	$g_{car/air}$	20.7	36.1	15.4
$g_{bus/car}$	15.4	12.8	−2.6	$g_{rail/car}$	19.9	17.7	−2.2	$g_{air/car}$	19.1	26.5	7.4
UEa	−4.3	−4.2		UEa	−2.0	−1.1		UEa	−1.7	−9.7	
	VTTS in rail	VTTS in air	ME		VTTS in bus	VTTS in air	ME		VTTS in bus	VTTS in rail	ME
$g_{rail/air}$	15.6	22.2	6.6	$g_{bus/air}$	9.9	17.5	7.6	$g_{bus/rail}$	11.4	11.3	−0.1
$g_{air/rail}$	12.1	19.7	7.5	$g_{air/bus}$	11.2	21.5	10.3	$g_{rail/bus}$	13.6	13.4	−0.2
UEa	−3.4	−2.5		UEa	1.3	4.0		UEa	2.2	2.1	

[a] Using 18.75 €/h for the normalized groups which roughly corresponds the 150 NOK/h, the recommended unit value for private long distance car-trips in Norway (Ramjerdi et al., 2010).

afford to) switch to air are likely to be wealthier than the average current rail user), and an ME of 6.6 €/h (rail more "comfortable" than air[17]). As discussed earlier the UEw of 1.3 € should be included in the VTTS-change after mode shift whenever it is reasonable that mode choice is voluntary (in most cases presumably).

The ME for air is particularly high and explains most of the VTTS differences to the other modes. A rather low overall impact of the UE might be explained by the relatively low income disparity in Norway.

4.3. Stylized case study

For an illustration of the concepts described, imagine the following hypothetical scenario.

- A road section has recently been upgraded from 2 to 3 lanes and the extra lane will soon open
- The road administration (thereafter: social planner) wonders if they should dedicate the extra lane (or part of it) to bus services only.
- Denote by $0 \le \pi \le 1$ the share of length of the extra lane dedicated to bus services. For $\pi = 1$ the whole lane is dedicated for bus services and no private car is allowed to drive there. $\pi = 0$ means free use of the extra lane, which will mostly benefit car drivers.
- The resulting time savings (in hours) are assumed to be $\Delta T_{bus} = \pi$ and $\Delta T_{car} = 1 - \pi$.
- Before opening the extra lane, there are 10,000 daily users of the road with market shares $P^0(\text{car}) = 0.5$ and $P^0(\text{bus}) = 0.5$.
- After opening the extra lane, market shares will be $P^1(car) = \frac{1}{1+e^{\theta(2\pi-1)}}$ and $P^1(bus) = \frac{1}{1+e^{\theta(2\pi-1)}}$ with $\theta \ge 0$ being a sensitivity parameter for the effect of time saving on demand (for $\theta = 0$ there will be no transferred demand between car and bus despite changes in the relative travel times).[18]
- No induced demand and no transferred demand from other travel modes.
- Travel time savings are the only user benefits and no third-party benefits and cost exist.
- The social planner wants to maximize the total benefits of the extra lane by finding the optimal project specification as given by the value of π.

With the empirical information of Table 3, the social planner considers four sets of VTTS estimates:

(i) *Equity VTTS.* $\overline{VTTS}$ is calculated as the average of the VTTS of the representative current users in all four modes: $VTTS_{car,g_{car}} = 18.8€/h$, $VTTS_{bus,g_{bus}} = 11.8€/h$, $VTTS_{train,g_{train}} = 14.3€/h$ and $VTTS_{air,g_{air}} = 26.0€/h$. As this is a hypothetical setting, it is conveniently assumed that all modes have the same market shares nationwide; hence an (un-weighted) averages of $\overline{VTTS} = 17.7€/h$ is applied.

[17] Many elements of "comfort" seem to favour train over air, e.g. more sitting space, the possibility to walk around, no annoying security checks, better (perceived) safety, possibility to talk on the mobile. phone etc.

[18] The formula for the market shares after opening the lane can be derived from an incremental logit model with generic time coefficient. This is a simplification ignoring the fact that time coefficients are likely to differ between travel modes (and user groups). Note that there is an inconsistency present when mode specific VTTS are used in project appraisal.

Table 3
Value of Travel Time Savings after switching travel modes.

In € per hour	VTTS after switching mode to …				Change in VTTS compared to representative user group (UEw + ME)			
Current mode	Car	Bus	Rail	Air	Car	Bus	Rail	Air
car	*18.8*[a]	17.0	18.8	36.1		−1.8 (0.9−2.7)	0.1 (3.2−3.1)	17.4 (2.0−15.4)
bus	15.4	*11.8*[a]	11.3	17.5	3.6 (1.0 + 2.6)		−0.5 (−0.4−0.1)	5.8 (−1.8 + 7.6)
rail	19.9	13.6	*14.3*[a]	22.2	5.6 (3.4 + 2.2)	−0.7 (−1+0.2)		7.9 (1.3 + 6.6)
air	19.1	11.2	12.1	*26.0*[a]	−6.9 (0.5−7.4)	−14.8 (−4.5−10.3)	−13.8 (−6.3−7.5)	

[a] Estimated Value of the representative user groups.

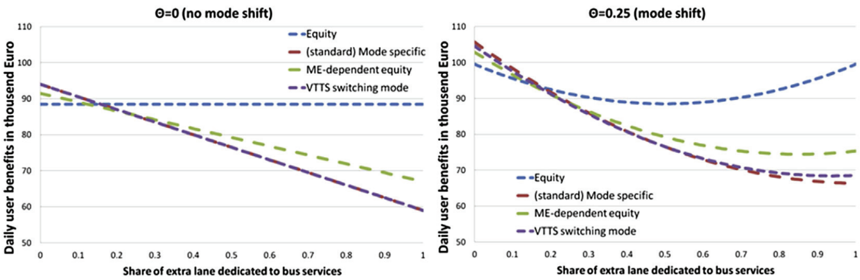


Fig. 3. Total user benefits of travel time savings for different values of π.

(ii) *Standard mode specific VTTS.* This approach uses the VTTS values given in (i) separately.

(iii) *ME-dependent equity values.* $\overline{VTTS}_{car}$ is here calculated as the un-weighted average of $VTTS_{car,g_{bus/car}} = 15.4€/h$, $VTTS_{car,g_{rail/car}} = 19.9€/h$, $VTTS_{car,g_{air/car}} = 19.1€/h$ and $VTTS_{car,g_{car}} = 18.8€/h$. Note that this calculation does not include all user groups in the population and idealistically one would also like to include the VTTS in car for user groups that do not use car as their current or first-best alternative mode. But, for instance $VTTS_{car,g_{rail/bus}}$ is not available in the Norwegian Value of Time study. Hence, $\overline{VTTS}_{car}$, calculated at $18.3€/h$, will still contain some self-selection. The same applies for $\overline{VTTS}_{bus}$ which is calculated at 13.4 €/h. Note that $\overline{VTTS}_{bus}$ is greater than $VTTS_{bus,g_{bus}}$ because the former is (at least partly) corrected for the user type effect.

(iv) *VTTS of switching modes.* It is given as the set of values: $VTTS_{car,g_{car\_car}} = 18.8€/h$, $VTTS_{bus,g_{car\_bus}} = 17.0€/h$, $VTTS_{car,g_{bus\_car}} = 15.4€/h$ and $VTTS_{bus,g_{bus\_bus}} = 11.8€/h$. As information about $VTTS_{car,g_{car\_car}}$ (i.e. the VTTS in car for users that use car before and after project implementation) is not available, $VTTS_{car,g_{car}}$ (i.e. the VTTS in car for the representative user of car) is used as an approximation. The same is applied for bus.

Figure 3 below depicts the total user benefits of time savings (in 1000€/day) obtained by Equation (18), i.e. calculated as the sum of WTP for individual time savings (depending on π) using each of the four sets of VTTS. The left graph refers to the situation without mode shift ($\theta = 0$), while the right graph represents a situation where there is some mode shift (the arbitrary choice of $\theta = 0.25$ implies that for $\pi = 1 (\pi = 0)$ the market share of bus (car) increases from 50% to 56.2%).

Considering the left graph ($\theta = 0$), the equity approach has the same user benefits independently of the chosen π. This underlines the potentially strong redistribution effect of the equity approach, as other arguments for bus (e.g. a slightly more favourable environmental impact compared to car) might tilt the decision in favour for bus users. Using mode specific values, the social planner should set $\pi = 0$, which maximizes time savings for car drivers. This is a simple consequence of the higher VTTS in car in the Norwegian data. Also under the ME-dependent equity values, it is indicated that the social planner should make the lane free for use. Here, the $\pi = 1$ project specification comes out relatively better compared to the (standard) mode specific VTTS, as UE has been removed from VTTS, increasing $VTTS_{bus}$ somewhat (from 11.8 €/h to 13.4 €/h). That the $\pi = 0$ project specification is still favourable (even after excluding UE) is due to the mode effect between car and bus which for the Norwegian data was found positive (implying that time in bus is more useful (comfortable) than time spend driving the car). The VTTS of switching modes corresponds to the mode specific set of values, under the assumption in this example that car (bus) users that do not switch have the same VTTS as representative car (bus) users.[19]

Considering the right graph in Fig. 3, representing a situation with mode shift (for all values of π but 0.5), the equity approach implies that π should be set either to zero or to one (these are equally good project specification assuming the same VTTS in bus and car). For the other three sets of VTTS, it is again shown that $\pi = 0$ gives the highest total benefits. Now there is a difference between the (standard) mode specific and the VTTS of switching modes, as "switchers" have been assigned a different VTTS value. With the chosen θ value, the difference is rather small. The gap between the two approaches increases when applying a higher value of θ (i.e. when assuming a higher number of "switchers" compared to "non-switchers").

Suppose now that the social planner wants to account for the fact that bus users have (on average) higher marginal utility of income. Assume that discrete choice models indicate that the marginal utility of the average bus user is 20% higher than for average car users. After some discussion a distributional weight of 1.2 for

[19] As mentioned above, this assumption was made because there were no specific VTTS estimates available for non-switchers. If they would be available and if the value would differ compared to the VTTS of the representative car drivers this two approaches might very well imply different user benefits.

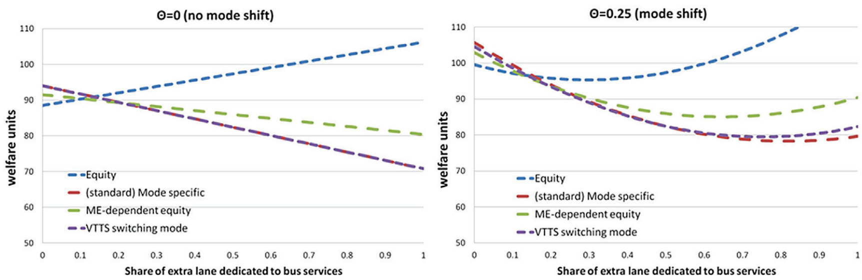


Fig. 4. Welfare of travel time savings for different values of π using distributional weight for bus users of 1.2.

bus users is therefore agreed upon. Fig. 4 depicts the resulting effects on welfare (which no longer can be expressed on a monetary scale).

Now $\pi = 1$ is proposed as the best project specification using the equity VTTS. Bus users are prioritized twice here, (i) because they are assigned the same VTTS as in car even though actual WTP (indicated by the Norwegian Data) is lower and (ii) because of the distributional weights chosen in this example. The former is defendable under the normative argument and the latter might be defended under an economic perspective (given that the marginal utility of income (and welfare function) is represented well with the weights). However, combining the two arguably lacks normative as well as economic reasoning. The other sets of VTTS come to the same project ranking as without distributional weights. This is because the ratio of VTTS is higher than 1.2. However, the gap in relative difference between project specification $\pi = 0$ and $\pi = 1$ is smaller given the choice of distributional weights.

Note that the (standard) mode specific set of VTTS is close to the ME-dependent equity VTTS without distributional weights (Fig. 3). And obviously they would coincide when weights where chosen $\alpha_{car} = \frac{\overline{VTTS}_{car}}{VTTS_{car,g_{car}}} = 0.973$ and $\alpha_{bus} = \frac{\overline{VTTS}_{bus}}{VTTS_{bus,g_{bus}}} = 1.136$. This shows that information on ME-dependent equity VTTS can also be used to set distributional weights that account for differences in user groups. This might be interesting from a normative perspective but also with an economic rational given that differences in the marginal utility of income (which were shown to be important elements of social welfare) are represented well by such distributional weights.

5. Conclusion

This paper discussed two conceptually appealing strategies to account for the mode and user type effects in assessing the value of travel time savings in project appraisal.

The *ME-dependent equity VTTS* was argued to be preferable over a single equity value, as it accounts for the mode effect, which was shown to be an important element of between-mode differences in VTTS. It does correct for the user type effect (self-selection) and is therefore consistent with the normative principles of the equity approach. With this set of VTTS, a unit of travel time saving *in a given travel mode* is valued equally for all types of persons. In a framework where one uses (standard) mode specific VTTS and wants to correct for the marginal utility of income, it was shown that the information about *ME-dependent equity VTTS* can also be used to calculate distributional weights. From this perspective, it implicitly (but only approximately) takes account of the marginal utility of income, and might therefore better fit the purpose of project appraisal, where changes in welfare matter (rather than the single monetary value of aggregated willingness-to-pay). To rigorously establish *ME-dependent equity VTTS*, one needs *VTTS* of a representative sample of the population for all travel modes. To avoid long and complicated SP questionnaires, an idea would be to take random subsamples, e.g. every respondent goes through a sequence of route choice tasks in two (random) travel modes.[20] In the Norwegian Value of Time Study the modes considered for route choice experiments depend on the current and (first-best) alternative travel mode. When calculating *ME-dependent equity VTTS* based on this value, only part of the self-selection is controlled for.

The *VTTS for switching modes* is in principle preferable to (standard) mode specific VTTS in situations where projects lead to changes in mode choice. This is because it takes into account the fact that there are user type effects between "switchers" and "non-switchers". In principle, it therefore provides a more detailed representation of actual WTP. However, when project specific valuation studies are not feasible, one has to rely on proxies for the actual (project specific) group of switchers. As this may involve a notable inaccuracy, it is not certain that one would actual gain precision in every project appraisal. Besides, for most projects, having a minor impact on mode choice, it does not seem worthwhile to account for it (especially in countries with rather low user type effects as in Norway). However, for big projects own valuation studies might be called for in order to assess the resulting VTTS for the (predicted) post-project market.

A hypothetical case study was used to illustrate that the chosen VTTS approach might have a strong impact on decision making. In this stylized example, the establishment and inclusion of different VTTS sets were rather straightforward. The applicability in more realistic and complicated settings is left for further research.

Acknowledgements

I want to thank Farideh Ramjerdi for introducing me to the topic and her contribution to the Norwegian Value of Time Study (without this project, this paper would not have been written). I also thank Harald Minken for his contribution to an unpublished working document on the VTTS of switching modes, Lasse Fridstrøm for proof-reading and Ståle Navrud for providing some useful comments. I am also grateful for comments and suggestions from Jonas Eliasson and two anonymous reviewers. All remaining errors, omissions and misconceptions are of course my own responsibility.

References

Abrantes, P. A. L., & Wardman, M. R. (2011). Meta-analysis of UK values of travel time: an update. *Transportation Research Part A: Policy and Practice, 45*, 1–17.

[20] However, even in this case it might be that relatively many respondents drop out of the questionnaire when the travel mode is not relevant for them. Hence some self-selection is likely to remain.

Algers, S., Dillen, J. L., & Widlert, S. (1996). The National Swedish value of time study. In *Paper presented at PTRC International Conference in Value of Time*.
Becker, G. (1965). A theory of the allocation of time. *The Economic Journal, 75*, 493–517.
Börjesson, M., & Eliasson, J. (2012). *Experiences from the Swedish value of time study*. Working papers in Transport Economics from CTS – Centre for Transport Studies Stockholm (KTH and VTI).
Börjesson, M., & Eliasson, J. (2014). Experiences from the Swedish value of time study. *Transportation Research A, 59*, 144–158.
DeSerpa, A. (1971). A theory of the economics of Time. *The Economic Journal, 81*, 828–846.
Flügel, S., Hjorth, K., & Ramjerdi, F. (2011). Between-mode differences in value of travel time estimates: evidence from Norway. In *Paper presented at Kuhmo Nectar Conference*. Stockholm: Sweden. http://indico.conferences.dtu.dk/getFile.py/access?contribId=130&sessionId=84&resId=0&materialId=paper&confId=75.
Flügel, S., & Minken, H. (2011). *Praktisk bruk av tidsverdier for overført persontrafikk i nytteberegning*. Oslo: Institute of Transport Economics (available on demand).
Fosgerau, M., Hjorth, K., & Lyk-Jensen, S. V. (2006). *An integrated approach to the estimation of the value of travel time*. Association for European Transport and Contributors.
Fosgerau, M., Hjorth, K., & Lyk-Jensen, S. V. (2007). *The Danish value of time study*. Copenhagen: Danish Transport Research Institute.
Fosgerau, M., Hjorth, K., & Lyk-Jensen, S. V. (2010). Between-mode-differences in the value of travel time: self-selection or strategic behaviour? *Transportation Research Part D: Transport and Environment, 15*, 370–381.
Gálvez, T., & Jara-Dıaz, S. (1998). On the social valuation of travel time savings. *International Journal of Transport Economics, 25*, 205–219.
Harberger, A. C. (1978). On the use of distributional weights in social cost-benefit analysis. *Journal of Political Economy, 86*, S87–S120.
Jara-Díaz, S. R. (2007). *Transport economic theory*. Emerald Group Publishing Limited.
Jara-Díaz, S. R., & Guevara, C. A. (2003). Behind the subjective value of travel time – the perception of work, leisure, and travel from a joint mode choice activity model. *Journal of Transport Economics and Policy, 37*(Part 1), 29–46.
Jara-Díaz, S. R., Munizaga, M. A., Greeven, P., Guerra, R., & Axhausen, K. (2008). Estimating the value of leisure from a time allocation model. *Transportation Research Part B: Methodological, 42*, 946–957.
Johansson-Stenman, O. (2005). Distributional weights in cost-benefit analysis—-should we forget about them? *Land Economics, 81*, 337–352.
Mackie, P. J., Jara-Dıaz, S., & Fowkes, A. S. (2001). The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review, 37*, 91–106.
NOU. (2012). *Samfunnsøkonomiske analyser*. Norges offentlige utredninger, 2012: 16 http://www.regjeringen.no/pages/38064911/PDFS/NOU201220120016000DDDPDFS.pdf.
Oort, C. (1969). The evaluation of travelling time. *Journal of Economics and Policy, 3*, 279–286.
Østli, V., Halse, A. H., & Ramjerdi, F. (2012). *Regional variasjon i verdien av reisetid*. Oslo: Institute of Transport Economics.
Ramjerdi, F., Flügel, S., Samstad, H., & Killi, M. (2010). *The Norwegian value of time study*. Oslo: Institute of Transport Economics.
Ramjerdi, F., Sælensminde, K., Rand, L., & Sætermo, I. F. (1997). *Summary: The Norwegian value of time study, Part 1*. Oslo: Institute of Transport Economics.
Rawls. (1971). *A theory of justice*. Cambridge, Massachusetts: Belknap Press.
Rawls. (2001). *Justice as fairness: A restatement*. Cambridge, Massachusetts: Belknap Press.
Samstad, H., Ramjerdi, F., Veisten, K., Navrud, S., Magnussen, K., Flügel, S., et al. (2010). *Den norske verdsettingsstudien - Sammendragsrapport*. Oslo: Institute of Transport Economics.
Sugden, R. (1999). *Developing a consistent cost-benefit framework for multi-modal transport appraisal*. University of East Anglia.
Wardman, M. (2004). Public transport values of time. *Transport Policy, 11*, 363–377.

Stefan Flügel

Stefan Flügel was born in Freiburg i. Brsg., Germany, in 1981. He holds a Diploma in economics from the Humboldt University of Berlin (2008). He works as a research economist at the Institute of Transport Economics in Oslo (TØI).

The thesis consists of an introduction and four independent essays. It concerns the modeling and application of transport-related choice data and contributes on the field of specifying utility function in travel mode choice and estimating and applying willingness-to-pay (WTP) measures and user benefits for economic appraisal.

Essay 1 is about the curvature of marginal utility functions of Level-of-Service attributes in travel mode choice models. It presents the concept of self-selection to attribute values in travel mode choice models which is argued to be a potential explanation for counter-theoretical empirical results estimated on cross-sectional data.

Essay 2 concerns the correlation structure among travel mode alternatives. The essay contributes to the current literature by identifying and discussing the limitations and caveats in deriving the error structure of the forecasting model from the estimation models based on binary stated choice data between travel's current mode and a new alternative (here: high-speed rail). The use of additional data and some advanced discrete choice models is proposed as a way to improve.

In Essay 3, we analyze choices made by cyclists in different types of choice experiments and elicit their WTP for cycling facilities such as separated cycling path and reduction of crossings. The novel element of this paper is that we include a casualty risk attribute. We find that WTP is close to halve when controlling for casualty risk.

Finally, Essay 4 provides a discussion about opportunities and challenges of including information about user type- and mode effects on between-mode differences in value of travel time savings in project appraisal. In this context, I argue that the proposed approach of "mode effect dependent equity value", which acknowledges mode effects due to comfort difference of travel modes but controls for user type effects due to self-selection, may help to provide optimal standards in economic appraisal

Professor Ståle Navrud was Stefan's supervisor

E-mail: sfl@toi.no

School of Economics and Business
Norwegian University of Life Sciences (NMBU)
P.O Box 5003
N-1432 Ås, Norway

Telephone: +47 6496 5700
Telefax: +47 6496 5701
e-mail*: hh@nmbu.no*
http:/www.nmbu.no/hh

ISSN: 1894-6402
ISBN: 978-82-575-1300-9