Report about results

User requirements and indications about demand volumes

June 09

kite

a knowledge base for intermodal passenger travel in europe

Deliverable D8

Report about results: user requirements and indications about demand volumes

Contract-No. TREN/06/FP6TR/S07.66711/038682-KITE

SIXTH FRAMEWORK PROGRAMME

Work Package Team

IVT

ETH
CH – 8093 Zürich
Switzerland

Tel. +41 (1) 633 4102
Fax +41 (1) 633 1057
Email frei@ivt.baug.ethz.ch
URL: www.ivt.ethz.ch

Andreas Frei

Kay W Axhausen

Document History:

Date	Author	Description
15/06/2009	Andreas Frei	Initial revision
17/06/2009	Tobias Kuhnimhof	1st revision
18/06/2009	Andreas Frei	2nd revision

Table of contents

Abbreviations

MEST – Methods for European Surveys of Travel Behaviour

INVERMO - Intermodale Vernetzung

DATELINE – Design and Application of a Survey for European Long-Distance Trips Based on a International Network of Expertise

KITE - A Knowledge Base for Intermodal Passenger Travel in Europe

CATI - Computer Assisted Telephone Interview

SP - Stated Preference

IVT - Institut für Verkehrsplanung und Transportsysteme der ETH Zürich

1 Executive Summary

The deliverable D8 has the goal to present first results from the KITE long-distance travel survey. The results are presented according to topic: Demand volume of long-distance travel, who are the long-distance travellers, modes used in long-distance travel and the user requirements.

Demand volume

The demand volume indicated by the number of long-distance journeys per person and year varies from survey to survey and within surveys, when compared by e.g. country. The results from the KITE survey are compared to the bench mark survey INVERMO and found comparable. KITE respondents report 11.67 journeys per person and year. The range of the values within the three surveyed countries Switzerland, Czech Republic and Portugal are acceptable (11.12 – 12.39), especially if considered that these three countries are very different from each other.

Who are the long distance travellers?

As already known the long-distance travel segment is dominated by a small part of the population, i.e. that 50% of the population produces over 90% of the long-distance travel demand. The typical long-distance traveller is male, in the middle working age and well educated.

Modes

The modes used for long-distance journeys are dependent on the available infrastructure. The mode shares by country vary accordingly. The public transport share in Switzerland is with 46% much higher than in the Czech Republic with 25%. The modal split is also highly dependent on distance, especially if we consider air travel, which is not an alternative for lower distances and it is the only mode for really long distances.

User requirements

The user requirement analysis shows that the modes used for long distance travel are chosen based on travel costs and time. The car shows a higher resistance against price changes and travel time changes than train or coach (value of travel time savings (VTTS) for coach: 36.29; car: 54.21; plane: 105.93; train: 46.34). Elasticities estimated show car travel to be elastic in the long distance range with a value of -1.15.

2 Introduction and background

Long-distance travel is a growing travel market segment, but reliable data and statistics about long-distance travel are rather rare. Information about long-distance travel has to provide answers to many different questions. For transport policy, cost considerations are central in the process of coordinating between different existing transport facilities and for the planning of new ones. Thus information about travel costs, prices of competitive modes and reasons for mode choice are absolutely required. In the context of transport planning, information about trip costs are equally useful as they provide insight into the cost structures, such as vehicle operating costs, user charges, taxes and tolls. But long-distance travel data is also especially needed for tourism, energy and also environmental policies.

Long-distance travel is only in rare cases a part of daily mobility (e.g. salesmen or ambassadors). Therefore, such movements are reported with low frequencies in surveys of daily mobility. However national travel surveys (NTS) dedicated to daily mobility are the only source for long-distance travel in many European countries (e.g. in Denmark, Greece or Netherlands). The problem in these surveys is the difficulty of obtaining representative statistics of long-distance travel even with relatively big samples. Therefore in most national travel surveys modules are dedicated to long-distance travel with the exceptions mentioned above. The most common format in these surveys is that respondents are asked to report their long-distance journeys in addition to their mobility on a given day, but for a longer reporting period (e.g. Great Britain NTS (2002-2004), French NTS (1993-1994), Swiss Microcensus on Travel Behavior (2005), and Swedish RES (2005/2006)). A further, but less frequent source of data for long-distance travel are surveys exclusively dedicated to long-distance travel. But past surveys show very different amounts of long-distance travel demand for the same countries, which make that those numbers hard to trust (see Dateline dataset with its consistently lower value for the number of trips per person and year).

Work package 3 of the KITE project tested in a pilot survey a suitable survey methodology that intends to close remaining information gaps about long-distance travel behaviour. These pilot surveys were carried out in Switzerland, the Czech Republic and Portugal.

A main objective of this pilot survey was to test a survey methodology in different countries in Europe and to asses the quality of information that can be obtained from it. In addition, a stated preference survey was designed to gather information about market potentials and user requirements.

The main focus of this deliverable is to give an overview of the demand volumes for long-distance travel based on these pilot surveys and compare these values to existing information about long-distance travel to get an idea about the promise of the KITE survey methodology and to sketch an impression of the user requirements based on the stated preference surveys.

First, descriptive statistics of the data are presented. In a following section the travel demand volume is analysed and compared with other data sources. The last part consists of the analysis of the stated preference dataset to calculate the value of travel time-savings for the different parts of the generalised costs of traveling.

2.1 Response rates

The response rates of the KITE survey vary from country to country. One influent factor here are the different survey protocols used. In Switzerland and Portugal the protocol used to obtain the data was a computer-aided telephone interview (CATI) while in the Czech Republic a face-to-face interview was used.

The response rates are given in Table 1.

Table 1 Response rates of the KITE long-distance travel survey

	Switzerland		Portugal		Czech Republic	
	[%]	Sample	[%]	Sample	[%]	Sample
Base sample	100.0	4’160	100	5'333	100.0	1’933
Interviews conducted	24.3	1’010	21.6	1'152	64.0	1’237
Non-response reasons						
Problem with the phone number	20.6	855	2.1	112		
Refusal	20.4	847	15.3	816		
Non-contact	1.3	56	50.5	2'693		
Age-problem	21.3	888	4.3	229		
Language-problem	10.9	454	0	0		
Other	1.2	50	6.2	331		
Mean interview duration (min)		16.5		16.5		37.0

The response rates have to be analyzed against the background of a maximum of 15 calls over several weeks. The samples are random samples and representative for the population controlled for age, gender and income. Because of the relatively small sample size, the samples are not representative for different regions within the countries.

For the Czech sample a quota sample was used. An exact non-response statistic is therefore not available for the data. The high number of non-contacts in the Portuguese sample occurs because they stopped to try to reach a part of the sample after they reached the requested numbers of interviews. But the persons in the base sample had at least one call attempt. The high share of problems with the phone number in the Swiss sample is caused by the change of an available official address and telephone number sample. This caused adjustments in the address-databases of the commercial survey companies, which has an effect on the sample quality and causes these phone-number problems.

The response rates are acceptable given the relative long duration of the interviews. Especially because the refusers have a share of only 20.4 % in Switzerland and 15.3% in Portugal, which is low if we compare this with other long-distance surveys, where refusal rates were e.g. 27 % in Norway (Denstadli, 1999).

The average interview duration for the telephone interviews is remarkably shorter than for the face-to-face interviews. This is not directly comparable, because the software used in a CATI makes the interview process much more efficient.

Based on the answers in the first part of the survey, in the second wave a customized stated preference survey was sent to self-identified respondents, which means that those respondents which had undertaken a long-distance journey during the last 8 weeks, which was not a regular journey (regular was defined: at least once per week; or journeys with the same destination during the last eight weeks), were asked if they are willing to participate in a written survey based on this telephone interview. For the response rates, see Table 2. An ex-ante estimate for the SP-questionnaire’s response rate in Switzerland, predicted a response rate of 54% (see Axhausen, 2007 for details about the ex-ante estimate of the response burden and the response rate derived from it). The lower actual response rate (45%) can be explained by the fact, that a commercial company conducted the interviews and the relative high response burden the respondents already had with the first parts of the survey.

Table 2 Response rates of the follow up KITE SP-survey

	Switzerland		Portugal		Czech Republic	
	[%]	Sample	[%]	Sample	[%]	Sample
Interviews conducted		1’010		1'152		1’237
Agreed to participate	100.0	507	100.0	230	100.0	524
Response	45.2	229	16.1	37	97.5	511

In total 777 SP-questionnaires were returned with each two times four choice situations, which gives us overall 6’216 observed choices.

2.2 Descriptive statistics

Table 3 shows the socio demographic characteristics of the respondents in the KITE pilot survey by country. For the Swiss KITE data the numbers in the brackets provide a comparison with the Swiss Microcensus Travel 2005, the Swiss representative national travel diary survey (Swiss Federal Statistical Office and Swiss Federal Office for Spatial Development, 2007). The comparison shows that the Swiss population is a little bit older (46.4 years old), slightly less educated, has a higher car availability and slightly fewer public transport season tickets. Obviously, the data does not show a systematic bias. A reweighting of the data seems not necessary given the relatively small deviations. Moreover, reweighing often entails the danger of producing other biases in the dataset.

The differences between the different countries are obvious and not surprising. While the age structure is nearly the same, the main differences are in the household income, where Switzerland has the highest, followed by the Czech Republic and Portugal. As the income was collected in categories, which are common to each country, the recoding of the income is rather misleading in the case of Portugal, where the highest income category is 3’000 Euros. As the education system in the surveyed countries is different, the categories cannot be compared directly with each other, especially as vocational training as such does not exist in Portugal.

Car availability does not differ significantly between those countries. Portugal has compared to Switzerland and the Czech Republic a slightly lower availability and the people, which have only frequently access to a car are in Switzerland remarkably more frequent than in Portugal and the Czech Republic. This is most probably coming from car sharing programs in Switzerland, which are among the most successful in the world. The public transport season ticket ownership is the category, which differs most obviously in this comparison. While the share in Switzerland is high with 54.6%, the shares in the Czech Republic (12.6%) and Portugal (3.9%)

indicate that public transport is less developed and not as affordable as it is in Switzerland.

Table 3 Socio-demographic characteristics by country

	Switzerland			Czech Republic			Portugal		
Variable	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Age	44.3	45	15.6	43.5	43	15.7	44.0	43	16.2
Household income (€)	4580	4640	1910	2490	2170	1440	2240	1100	3670
Males	45.42%(47.87%)			50.14%			48.14%		
Education									
N.A.	3.24%(12.46%)			0.22%			4.82%		
Obligatory schooling	12.58%(19.22%)			19.76%			44.43%		
Vocational training	46.72%(31.34%)			34.83%			-		
Highschool diploma	6.01% (9.18%)			34.22%			30.21%		
Further technical training	8.42%(10.68%)			1.05%			1.94%		
University degree	22.46%(17.12%)			9.91%			18.48%		
Car available									
Always	76.47%(69.83%)			78.42%			63.78%		
Frequently and rarely	18.13%(18.42%)			0.00%			6.76%		
Never	5.33%(12.84%)			21.63%			29.12%		
Public transport season tickets									
Ownership Yes	54.61%(52.22%)			12.64%			3.87%		
Sample size N	1011			1274			1257		

3 Number of long-distance journeys

As already stated in the introduction long-distance journeys are rare events (in the case of the 1995 *American Travel Survey* 4.0 journeys/year over 100 miles and about one percent of all journeys[1] fall into this category (BTS, 1997)). As rare events are not captured in sufficient numbers in daily mobility surveys of national travel surveys (NTS) reliable data about long-distance journeys, even only basic demand volumes, are not or only rarely available. The main differences which produce biased figures are the survey methodology used. Therefore a summary about the survey challenges of long-distance travel and how the KITE survey tried to overcome these problems is given here.

The core of the design problems for long-distance travel surveys and the reason for the differences in the numbers about the demand is their very limitation to journeys with a minimum distance or duration. The movements to be reported are rare events requiring long reporting periods to increase the chance that the respondent can report at least one journey and that the contact is not wasted in terms of capturing information about travel. Counterbalancing this is the problem of recalling events, which might have happened weeks ago, in some detail, which limits the reporting period to a range of four to eight (twelve) weeks, given the relatively low salience of routine long-distance travel for many above average frequency travellers (Armoogum and Madre, 1997).

3.1.1 Long-Distance Travel Definitions

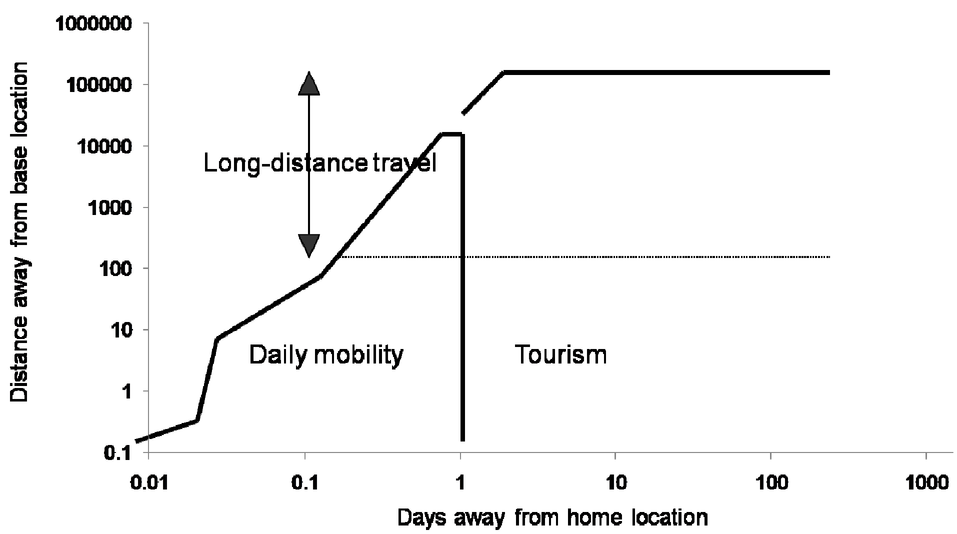
In surveys of daily mobility the study objective is clear: the capture of all movements of the respondents for a day, excluding only movements within large facilities, such as shopping centres or factories. Even this basic question is open to discussions in the case of long-distance travel. Because the division between movements relevant to long-distance travel and the related decision making and the irrelevant local movements needs to be defined, as it is impossible to ask the respondents to report all movements undertaken during a multi-day long-distance journey (See Axhausen, 1996 for a more thorough discussion).

---

1 Stage: a continuous movement with one mode/means of transport; trip: a sequence of stages between two activities; tour: a sequence of trips starting and ending at the same location; journey: a tour starting and ending at the current base (e.g. home)

The difficulty for the survey design is to avoid complex definitions and to find a natural description, which invites the respondents to report all relevant movements, by minimising the difference between the reported movements and those which should have been reported. As mentioned in section 2.2 there are mainly two different focuses in long-distance travel surveys, which leads to two different styles of long-distance travel definitions. Figure 1 shows these two possible views on long-distance travel in a schematic. On the one hand it is possible to differentiate long-distance travel from daily mobility by the duration of being away from home; on the other hand it is possible to define it by a minimum distance travelled from a certain base location. While the duration of a stay is mainly in focus when looking at tourism where the data is needed for supply and marketing decisions, the duration of stay is not crucial for transport planning, where the data need is more focused on route/mode choices which are determined in part by distance.

Figure 1 Long-distance travel definition



Source: Adapted from Axhausen, 1996

While the decision to use distance as a criteria is widely accepted in transport planning long-distance surveys, the exact cut-off and type (crow-fly or network distance) of distance were never harmonized and vary from country to country.

3.1.2 Recall problems

The definitions in long-distance surveys have next to the problems above further implications on data of long distance travel. As long-distance journeys are rare events it is indispensable to have a relatively long duration reporting period and therefore recall problems occurs.

The duration of the reporting period interacts as a design variable with the basic unit of the reporting chosen (i.e. stage, trips or journey) and therefore the level of detail, as this implies a certain recall burden for the respondent. The wish of the analyst for detail has to be traded off against the response burden and recall difficulty of stages or even trips undertaken some time ago. A four-week reporting period might be compatible with stages while a twelve-week reporting period only with journeys.

In postal questionnaires this issue is compounded by the issue of how to provide for the "standard" trip or journey: a simple out-and-return journey or a complex trip involving multiple stages. A paper form cannot accommodate certain levels of complexity in a self-completion context, which limits the possibility to choose the base unit and to define it. Related to this is the question of how frequent travellers or repeated trips can be supported. In the first case, one would like to reduce the response burden by either simplifying the task or by reducing the reporting period.

3.1.3 Fatigue effects

A relatively long reporting period for surveying rare events creates the further problem of fatigue effects. The response burden distribution in long-distance travel surveys is highly skewed in contrast to daily mobility surveys because of the highly skewed distributions of travel frequency. To avoid the tedium of repeating the description of very similar journeys for frequent travellers one could offer shortcuts. While both things can be achieved in CATI/CAPI contexts, they are not as easily possible on paper forms without inviting other respondents to use these short-cuts. In addition, one is interested in the details of those journeys of frequent travellers, if one has doubts about the identity of those repeated journeys.

The same questions reoccur when looking at the design of the question sets for each reporting unit: number of items, complexity of the items, and complexity of the available pre-coded answers. The designer has to trade-off desired detail against respondent boredom and response burden. This issue interacts with the design of the questions on the page, where multiple units on each page save postage and reduce the footprint of the forms, while equally generating the impression of complexity through the busyness of the page.

This brief discussion has highlighted the special difficulties inherent in conducting surveys of long-distance travel, where the complexity of the subject, the resulting response burdens and the data needs have to be balanced, so that valid and useful data are obtained at reasonable cost.

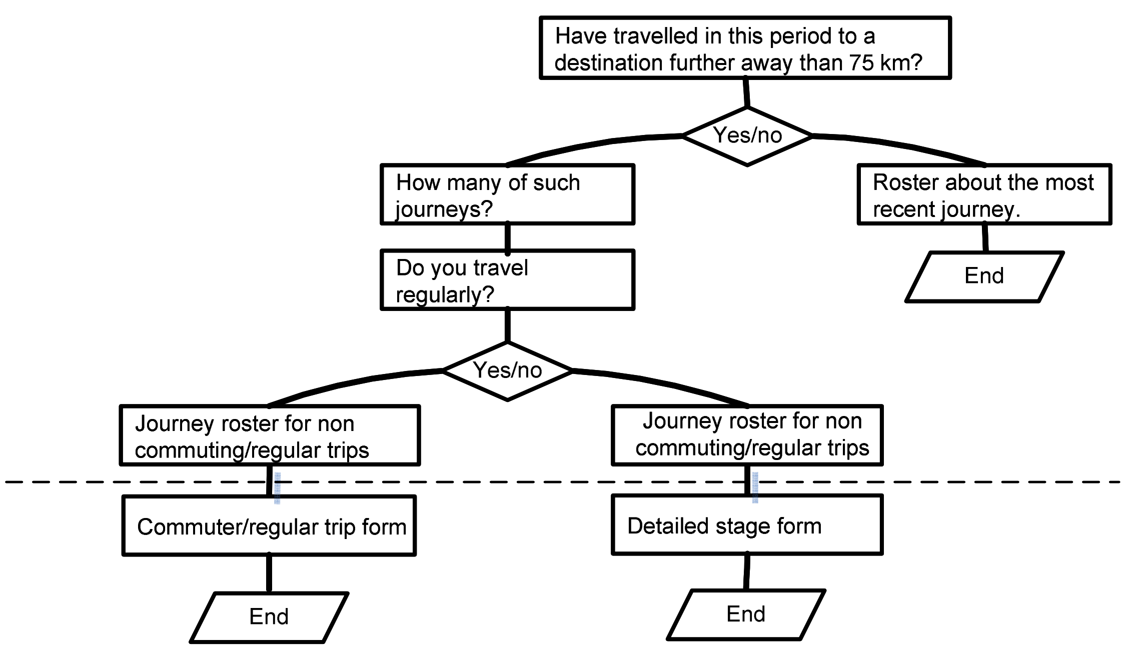
3.1.4 The KITE survey methodology

In addition to the normal household data, person data and travel diary in the survey, a protocol of three steps is used.

As discussed above there is a conflict in balancing the analysts' wish for a high level detail and the response and recall burden. The chosen relatively long reporting period is inconsistent with the need of capturing stages and with the low cut-off of 75 km as the boundary of long-distance travel, because of fatigue effects. To avoid this imbalance, two arrangements from the MEST and the INVERMO survey were implemented into the survey protocol. First, a screening by CATI/Face-to-Face is implemented to screen the relevant respondents and filter frequent long-distance travellers to help the respondent to abbreviate the description of repeated journeys. Second, a journey roster is employed which obtains the basics of the journeys undertaken during the reporting period, such as start/end, date of departure/return, destination, main mode, main purpose and size of party. The detailed stage description then follows for the three most recent journeys. In addition to avoid repeated questions for frequent long-distance travellers and to lower their response burden a special commuter/regular trip questionnaire asks only for one of these repetitive journeys. This should reduce the refusal rate of this very important segment of travellers. The logic of the protocol is graphically presented in Figure 2.

In Switzerland and Portugal the surveys were carried out as a CATI and in the Czech Republic as a face-to-face interview. For more details about the field work see Deliverable D7.

Figure 2 Logic of the protocol



3.2 Calculating the number of long-distance journeys

The simplest way to calculate the number of long-distance journeys as a indicator for long distance travel demand is extrapolating the mean number of regular and non-regular long-distance journey reported per person during the reporting period (8 weeks in the KITE survey) to a whole year.

The mean numbers of long-distance journeys per person and year calculated as described above are 10.06 in total, 8.76 in Switzerland, 12.41 in the Czech Republic and 8.70 in Portugal. The relative difference among the countries is large and seems to be unreasonable, as there is no logical explanation why people in the Czech Republic should travel more than people in Switzerland or Portugal.

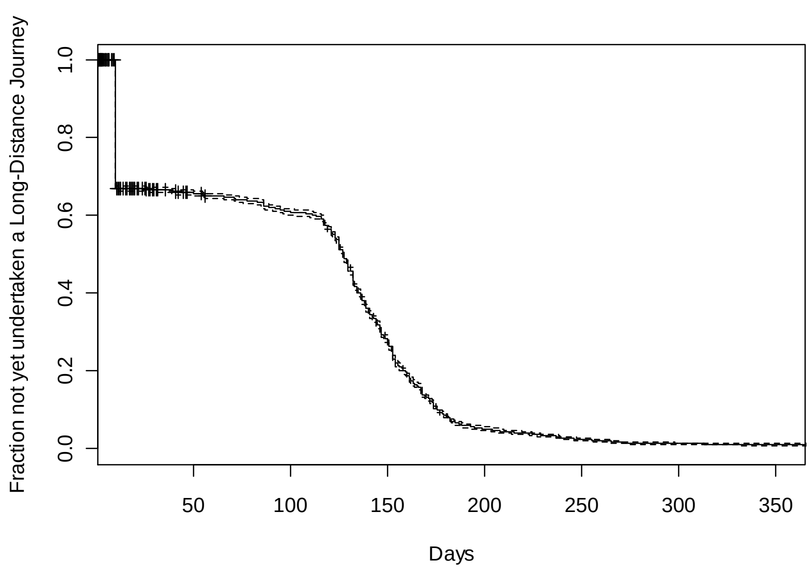
The data shows that specifically the number of persons which did not report a long-distance journey during the reporting period is much lower in the Czech Republic (18.7%) than in Switzerland (48.05%) or Portugal (49.6%). As compared to the standard method, the KITE questionnaire design allows use of additional and more precise information to calculate the numbers of long-distance journeys, e.g. the last long-distance journey undertaken, if a interviewee did not undertake such a journey during the reporting period. KITE data yields information about regular and non-regular long-distance journeys during the reporting period. The frequency of regular journeys and date of non-regular journeys are known. All of this information is combined using a survival function *S(t)=Pr(T>t) to calculate the mean time between long-distance journeys.* Information about the period between two events or undertaken long-distance journeys is available as completely observed periods or as left- or right-censored periods, meaning that the start or end of the periods are not observed. Table 4 gives an overview about the data used to calculate the mean survival time between two long-distance journeys.

Table 4 Journey intervals in the KITE survey

Data source		Censoring
Last long-distance journeys before the 8 week reporting period		Right censored; the end of the period is not observed
Regular journeys		Non censored
Non-regular journeys	First one reported	Left censored; the beginning of the period is not observed
	between	Non censored
	Last one reported	Right censored

The survival function used is a non-parametric survival function determining the probability that the time between two long-distance journeys is longer than some specified time. It is the inverse of the cumulative distribution function and imputes the unknown, censored information. Figure 4 shows the survival rate calculated for the period between two long-distance journeys.

Figure 3 Survival function of the time period between two long-distance journeys for all journeys



The survival rate is clearly not parametric, which has its origin in the three different data sources, where the regular journeys produce a very small survival time, the non-regular journeys the middle survival times and the last long-distance journey undertaken the tail of the function which decreases very slowly.

The calculated mean survival time is 31.26 days, which corresponds to 11.67 journeys per person and year. The effect on the total number of long-distance journeys compared to the mean calculated above is rather small (+16%). The differences between the countries are still there, but they are smaller as expected, because the interviewees in Switzerland and Portugal had a significantly larger number of people, which did not undertake a long distance journey during the reporting period.

3.3 Comparison of Long-Distance Travel Surveys

There is no data direct available from a long-distance survey for a direct comparison of the data for Switzerland, Portugal and the Czech Republic. Deliverable 4 from the KITE project contains comprehensive overview which uses DATA from different sources to compute the best guess for long-distance figures among various countries. The borderline used in the KITE project to define long-distance journeys is 100 km, which is used to calculate the figures in D4. To compare the survey outcome in D8 with available Data, all containing information in the Dataset is used because of the relatively small sample. DATELINE contains information for Switzerland and Portugal, but is not reliable for demand volumes, national travel surveys also do not have comparable long-distance data, e.g. the Swiss microcensus contains information about journeys with overnight stay. But to get an idea about the quality of the data obtained in the KITE project and the consistency of the survey methodology used, a short comparison between the long-distance travel demand figures of KITE and of the INVERMO survey is made. The German INVERMO long-distance panel travel survey was carried out at the national level independent from a regular national travel survey. Table 5 presents the basic figures of these two different data sources.

The long-distance data from the German Survey INVERMO is one of the most reliable data sources for long-distance travel, because it produced very similar figures compared to the German NTS MiD (Mobilität in Deutschland) and because the survey methodology used was state of the art and as a panel survey the surveyed persons had feed-back during the whole reporting period.

The definition of long-distance travel is different between INVERMO and KITE, but as the reported distance is comparable to network distance and network distance is around 1.3 times crow-fly distance as shown by Chalasani et al., 2004, the definitions are comparable for the travel demand volume. Overall the figures of the KITE survey produce are rather similar to INVERMO. The KITE survey produces in total 11.67 reported long-distance journeys per person and year compared to 8 journeys in INVERMO.

Table 5 Comparison of long-distance journeys per person and year and modal split

Definition of long-distance travel	INVERMO Germany	KITE Switzerland	Czech Republic	Portugal
Journeys > 100 km reported distance per person and year (without commuting)	8			
Journeys > 75 km crow-fly distance (without daily long distance commuting)		11.41	11.12	12.39

The comparison between the different countries within KITE shows that the standard deviation and mean numbers in Switzerland and Portugal are very similar. It is rather unlikely that people in the Czech Republic undertake more long-distance journeys than people in Switzerland respectively in Portugal. The differences in the number of long-distance travel between the countries within KITE are not huge, but still seem to be systematic and have their origin in the survey methodology, where in the Czech Republic the data was collected as a face-to-face interview in contrast to Switzerland and Portugal, which used a CATI interview. This leads to the conclusion, that there is a share of persons, which do not provide as comprehensive information in a CATI interview as in a face-to-face interview. The last long-distance journey, if they did not undertake one during the reporting period, has helped to correct for some of those omissions.

The comparison of the distance distribution of the journeys between the three countries (see Table 6) shows that the Czech data provides a much higher share of trips just above the border of 75 km crow-fly distance. Even if the interviewers did not provide further information about distances e.g. with a map, the help of an interviewer and the possibility to discuss such a distance helps the interviewees to estimate the distance between a origin and destination.

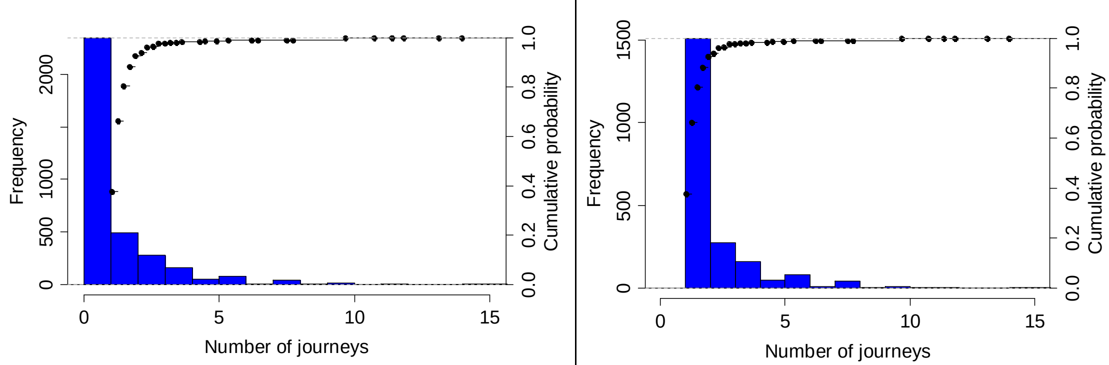
This brief discussion has highlighted the special difficulties inherent in conducting surveys of long-distance travel, where the complexity of the subject, the resulting response burdens and the data needs have to be balanced, so that valid and useful data are obtained at a reasonable cost. Even if there seems to be still a bias due to the face-to-face interviews, the numbers of journeys undertaken are quite similar and higher in comparison with DATELINE, the last survey which was carried out in different European countries, where for Germany three long-distance journeys per person and year were estimated and for Switzerland only two (For a comparison of existing conditioned data, see Deliverable 4).

3.4 Modelling the number of long-distance journeys

For modelling the numbers of long distance journeys are the non-corrected data used. Figure 4 a) shows the distribution of the number of long-distance journeys. The distribution is highly left skewed with an excess of zeros. Figure 4 b) shows the same distribution but without the persons which did not undertake a journey within the reporting period. The data is still highly left skewed.

Figure 4 Distribution of the number of long-distance journeys per person and 8 weeks

a) Histogram with zero journey respondents | **b) Histogram without zero journey**

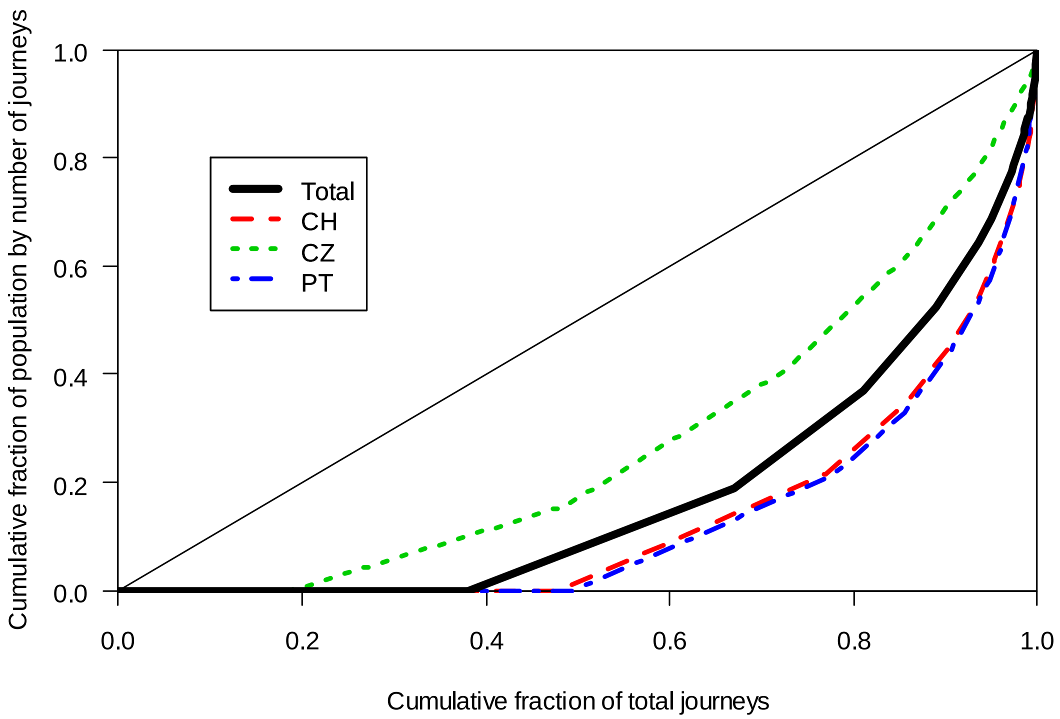


This highly skewed distribution is a further problem of long-distance travel data. For example, Last, Manz and Zumkeller (2003) found in the INVERMO data that half of the population in their sample produced over 90% of the long-distance journeys. The most mobile percent of the population undertook long-distance journeys ten times more frequently than the national average.

Figure 5 shows the Lorenz curve of the long-distance journeys undertaken to compare the results in INVERMO presented above with the results in KITE. The Lozenz curve shows the fraction of persons, which produce a given fraction of long-distance journeys. It can be seen, that for Switzerland and Portugal the same share of persons (50% of the population produces a little bit over 90% of the long-distance journeys).

The sample from the Czech population is again different, and again because they have a much lower share of people who do not undertake long-distance journeys. This can be seen in Figure 5 were the green line starts at a little bit under 20% whereas the Lorenz curve for Switzerland and Portugal begins around 50%. The further progress is similar for all countries and is a reflection of the skewed distribution in Figure 4. If the distribution were uniform, the Lorenz curve would be equal to the diagonal.

Figure 5 How many people produce how many long-distance journeys (Lorenz curve)?



The number of long-distance journeys reported by socio-demographic attributes of the respondents are reported in Table 7. The first column for each statistic contains the values for all interviewees and the second contains only the numbers for interviewees who at least did one regular long-distance journey during the reporting period (regular is defined as: regular according to the time-interval with a minimum of once per week or regular according to space with at least two long-distance journeys with the same destination in the last 8 weeks). Age seems to make a difference, with the middle aged people (31 to 40) undertaking more long-distance journeys than the older (41 to 60) and the under 30 ones. With the retirement age of about 61+ immobility increases because of growing infirmities, lower disposable incomes, absence of work requirements (from under 2% for people under 40, 5% for 41-60 and 10% for 61+), which leads to a lower number of long-distance journeys. Gender has a surprisingly high influence on the numbers of long-distance journeys. The average number of long-distance journeys which are only undertaken on a regular basis are with an average of 45.6 journeys per person and year for males higher than for females with 39.1 and the share with regular long-distance journeys is higher with 26% for males then 23% for females. The number of non-regular long-distance journeys does not differ between the genders. The education level indicates that a higher level of education increases the number of long-distance journeys. The category further technical training is an exception and exists only in Switzerland as common education category, so it cannot be compared directly. The income level shows like the education level a clearly visible trend where higher incomes increase the number of long-distance journeys.

Table 7 Average number of total and regular long-distance journeys by socio-demographic characteristics (not corrected by hazard-model analysis)

Variable	Mean		Min.		Max.		Median		SD	
Category	Tot	Reg	Tot	Reg	Tot	Reg	Tot	Reg	Tot	Reg
Age										
Up to 30	12.2	21.5	0.0	1.0	326.1	326.1	6.5	13.0	27.9	39.4
31 to 40	17.5	25.4	0.0	1.0	391.3	391.3	6.5	13.0	46.8	51.7
41 to 60	12.6	22.2	0.0	1.0	326.1	326.1	6.5	13.0	29.2	43.9
61 and older	8.5	13.1	0.0	1.0	260.9	260.9	6.5	6.5	16.8	24.0
Sex										
Female	9.3	16.0	0.0	1.0	391.3	391.3	6.5	13.0	21.2	30.9
Male	16.3	25.9	0.0	1.0	326.1	326.1	6.5	13.0	39.5	49.6
Education										
N.A.	7.0	12.0	0.0	1.0	52.2	52.2	0.0	6.5	11.5	13.1
Obligatory schooling	8.2	11.1	0.0	1.0	326.1	326.1	0.0	6.5	24.6	24.9
Vocational training	11.7	20.2	0.0	1.0	260.9	260.9	6.5	13.0	25.5	44.5
Highschool diploma	11.0	20.7	0.0	1.0	391.3	391.3	6.5	13.0	37.5	44.8
Further technical training	11.1	22.8	0.0	1.0	65.2	65.2	6.5	19.6	15.4	20.3
University degree	19.3	28.8	0.0	1.0	260.9	260.9	6.5	13.0	42.3	52.5
Income [Euro/month]										
0-1999	11.6	16.4	0.0	1.0	326.1	326.1	6.5	6.5	34.5	42.4
2000-5999	12.2	21.8	0.0	1.0	260.9	260.9	6.5	13.0	23.2	35.1
6000+	18.1	35.4	0.0	1.0	391.3	391.3	6.5	13.0	44.4	65.8

To capture the wide range in the number of long-distance journeys further analysis is necessary to determine the probability distribution, which represents the data set. Figure 4 shows that the count data follow a left skewed curve. A possibility to represent count data is the discrete Poisson distribution. The variance for the number of long-distance journeys is much higher than the mean (see Table 5). This indicates over-dispersion. To deal with the skew of the number of long-distance journeys a negative binomial distribution can be used instead of a Poisson distribution. The data has also an excess of zero and therefore in addition to predicting the number of long-distance journeys undertaken, there is interest in predicting the existence of people with no long distance journeys, i.e., the probability that a person does not undertake long distance journeys during the reporting period. The zero-inflated negative binominal regression is a two step model with two processes, where the first process is a logit model which is chosen for all interviewees with no long.distance journeys during the reporting period and process two, the negative binominal model, when the interviewee has undertaken at least one long-distance journey.

Socio-demographic, travel related and survey-specific dummy variables are employed to explain the number of long distance journeys using a zero-inflated negative binomial regression model. After removing variables, which correlate highly with each other (limit = 0.5; e.g. working status and place of work), variables with a significance level below 0.05 were removed step-wise. The parameter estimates are reported in Table 8.

Table 8 Parameter estimates for the zero-inflated negative binominal regression of the number of long-distance journeys

Negative binominal regression model coefficients (with log link)				
Variable (Base Categorie)	Beta	b/St. err	Sign.	
Constant	0.49	0.12	4.24e-5	***
Country (Switzerland)				
Czech Republic	0.29	0.09	3.31 e-5	***
Portugal	0.34	0.09	3.64 e-5	***
Gender female [y/n]	-0.38	0.06	-6.72 e-5	***
Age (Up to 30)				
31 to 40	0.21	0.08	2.57e-2	*
41 to 60	0.06	0.07	0.85	
61 and older	-0.22	0.09	-2.51e-2	*
Car available (Always)				
Frequently and rarely	0.09	0.13	0.69	
Never	-0.59	0.08	-7.25e-5	***
Frequent flyer card [y/n]	0.60	0.13	4.75e-6	***
Educ. (Obligatory schooling)				
Vocational training	0.22	0.09	2.49e-2	*
Highschool diploma	0.33	0.08	4.15e-5	***
Further technical training	0.10	0.17	0.58	
University degree	0.42	0.09	4.78e-5	***
Log(theta)	-0.19	0.04	8.75e-6	***
Zero-inflation model coefficients (binomial with logit link)				
Constant	-0.73	0.56	0.197	
Car available (Always)				
Frequently and rarely	0.75	0.39	0.056	.
Never	0.58	0.41	0.157	
Educ. (Oblig. schooling)				
Vocational training	-1.51	0.56	0.007	**
Highschool diploma	-0.62	0.32	0.056	.
Further technical training	-1.31	0.78	0.092	.
University degree	-13.93	796.97	0.986	

N = 3196 Adjusted R^2 = 0.53

Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

The goodness-of-fit is, as expected from the descriptive statistics, relatively good (R^2=0.53), and the comparison of the current model with a null model without predictors using chi-squared test on the difference of log likelihoods shows that the model is highly significant (less then 0.001). We can see in the model that the dispersion parameter Log(theta) is significantly different from zero. This suggests that the outcome is overdispersed and that a negative binomial model is more appropriate than a Poisson model. The Vuong statistic suggests that the zero-inflated model is a significant improvement over a standard negative binomial model with a p-value of 0.0008.

The count predictors show an expected change in log(count) compared to the base category, e.g. female persons (female = 1) had an expected log count -0.38 higher than male persons (male = 0).

The country of origin show a significant influence on the numbers of long-distance journeys, whereas Portugal and the Czech Republic show a positive influence on the number of long-distance journeys compared to Switzerland. The gender shows the same results as in Table 7. The age of the respondents shows a quadratic-shaped influence. Younger people undertake fewer long-distance journeys than those aged thirties and more then people in retirement age. The car availability has a strong positive influence on the number of long-distance journeys as well a frequent flyer card. A higher education leads to more long-distance journeys.

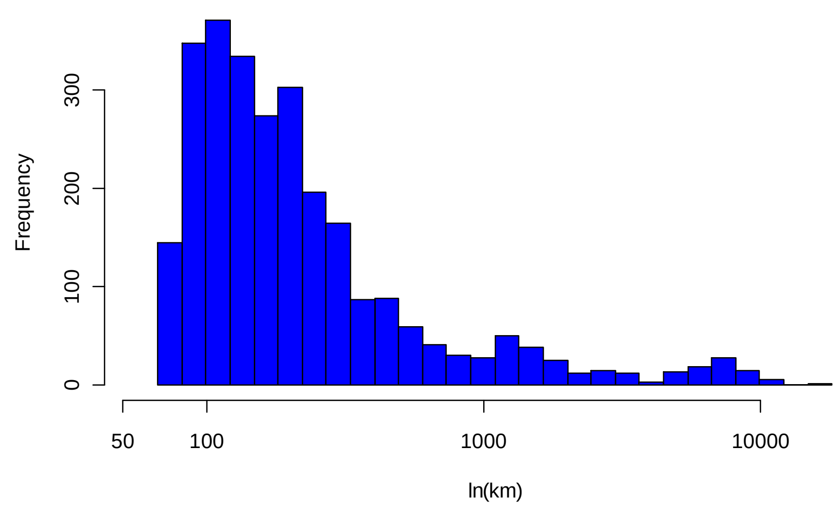
Below the model results for the negative binominal regression model, a block of output containing the logit coefficients for predicting excess zeros along with their standard errors and p-values. The logit coefficients only measures if a interviewee has undertaken a long-distance journey during the reporting period or not by giving the probability for this binary variable (section 6 explains the logit model more precise). A negative parameter indicates a lesser likelihood to have not undertaken a long-distance journey during the reporting period. The estimates for the logit model are similar to the parameter estimated for the negative binominal model, which is expected.

4 Distance and long-distance journeys

Figure 6 shows the distance distribution and it is clearly visible, that the distance is high left skewed even on a logarithmic scale for the distance. This show how important the exact cut-off value is for the interpretation of the results, because a cut at the lower end of the distance will have huge impact on the overall number of long-distance journeys; e.g. in the KITE survey 10.2% of all journeys were cut off due to lower distances then the defined value.

Not reasonable is the ascent in the frequency in the very low distance groups, where here it would be expected, that there would be a decrease with distance. This must be an artefact from the survey methodology where in the lower distance band the interviewees do not give all the journeys, because they believe the distance would be too low. This is another argument for choosing a lower boundary (75 km) in the survey than the actual definition (>100 km) is.

Figure 6 Distance distribution of journeys



5 Modes of transport

The resulting modal split of the KITE survey is compared in Table 9 to the INVERMO survey. Switzerland is known for a much higher share of public transport than other European countries. The modal split of the Czech Republic is very similar to the results of INVERMO from Germany for the shares of car journeys. The Coach is a transport mean, which is not very common in Switzerland and mostly used for longer trips of foreign born residents to visit their families in the country of their family's origin. Overall the share of air travel is higher than in INVERMO, which can be a result of the price drop in this segment in the last years.

Table 9 Comparison of long-distance journeys per person and year and modal split

	INVERMO	KITE			
Modal Split [%]		Switzerland	Czech Republic	Portugal	Total KITE
Car	74	50.98	74.88	67.00	68.07
Coach	5	3.92	3.33	11.21	8.33
Train	11	28.43	9.51	4.53	11.61
Air	8	14.31	12.28	15.62	10.39
Public Transport Total	24	46.66	25.07	31.06	30.33
Other	2	-	-	1.64	-

The modal split is highly dependent on distance, especially if we consider air travel, which is not an alternative for lower distances and it is the only generally reasonable mode for really long distances. The other modes are competitive to each other in comparable distance bands.

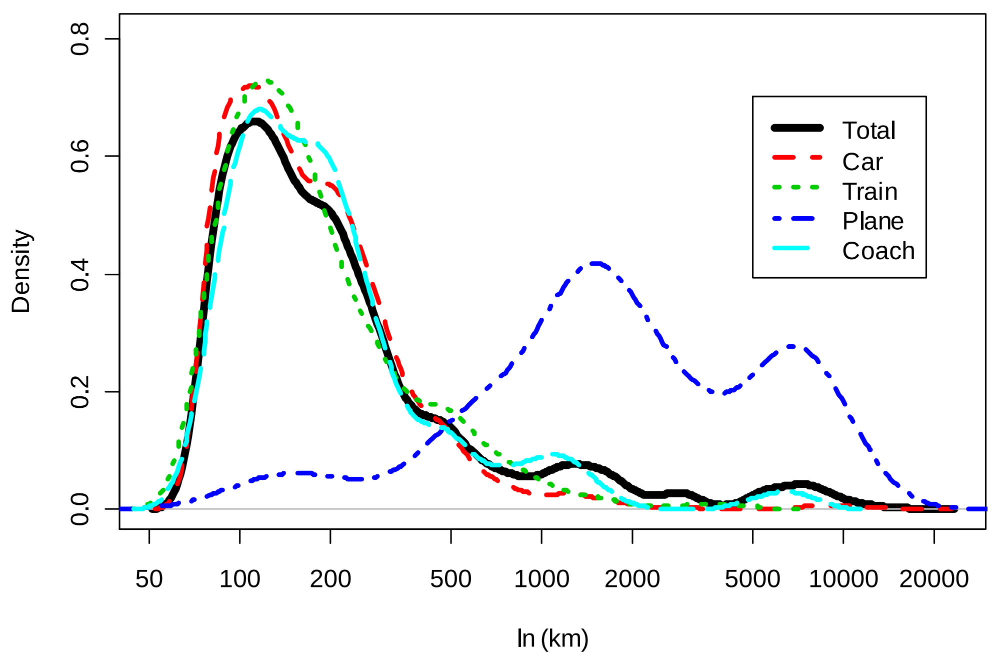
Table 10 shows the distances of the different modes by country. As already stated above, the mean distance for the plane is much higher compared to the other modes. What we also see is that the Czech Republic has overall a much lower mean distance than Switzerland and Portugal. This is an artefact of the survey methodology which was already described in section 3.2. The mean distance for the coach is in Switzerland much higher than in the other countries, because the coach is used only for longer journeys e.g. to Spain, Portugal, Croatia or Serbia.

Table 10 Distance by country and mode [km]

	Total		Switzerland		Czech Republic		Portugal	
Mode	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total	528	1390	954	2167	282	873	692	1409
Car	267	764	429	1144	229	767	261	459
Train	248	359	235	214	242	475	321	323
Plane	2’949	2’781	3’249	3’367	2’391	2’590	2’903	2’420
Coach	363	857	1’277	2’021	289	776	302	285

Figure 7 shows the distance distribution by mode. It is clearly visible, that the plane is only competing in a distance band from about 300 km to 1000 km crow-fly distance. Also visible is the peak at around 10’000 km which indicates the overseas flights.

Figure 7 Distance distribution and mode



The share of travellers using public transport to the airport for taking the plane is very different in Switzerland to the other countries. In Switzerland 36% of the stages to the airport are done with the public transport (in deliverable D4 figures for Switzerland from the Swiss national travel survey indicates even a higher share of 46%), where in the Czech Republic and Portugal only 18.7% respectively 20.7% of the stages are undertaken by public transport. The share of intermodal journeys is in total with 30% a little bit higher than the share for using public transport to the airport. But also again as expected is the share in Switzerland higher than in the two other countries.

6 Mode choice in long distance travel and value of travel time

6.1 Background and introduction

As already mentioned in Chapter 2, the second part of the KITE survey was a written self-completion stated preference survey based on one of the reported long distance journeys from the telephone interview. The main focus in this stated preference survey is to analyze the different parts of the utility function, which describes the preferences of the travellers undertaking long-distance journeys. These preferences show the requirements of the users and their requirements towards a more sustainable use of transport means, e.g. under which circumstances they would change the transport mean and use public transport instead of the car. The use of a transport mode is of course dependent on the available infrastructure in the different countries and regions etc., which was shown in the previous chapters, if we look at the differences between different countries. The available infrastructure is not analyzed in this survey, but the results give parameters, which indicate under what kind of infrastructure change the population would change their transport mode.

6.1.1 Generating the experiments

To generate the experimental design for the SP-questionnaires the software Ngene (Rose *et al.*, 2008) was used. This software makes it possible to generate efficient experimental designs and therefore have small numbers of experiments by interviewee without losing goodness of fit in the models estimated with the data.

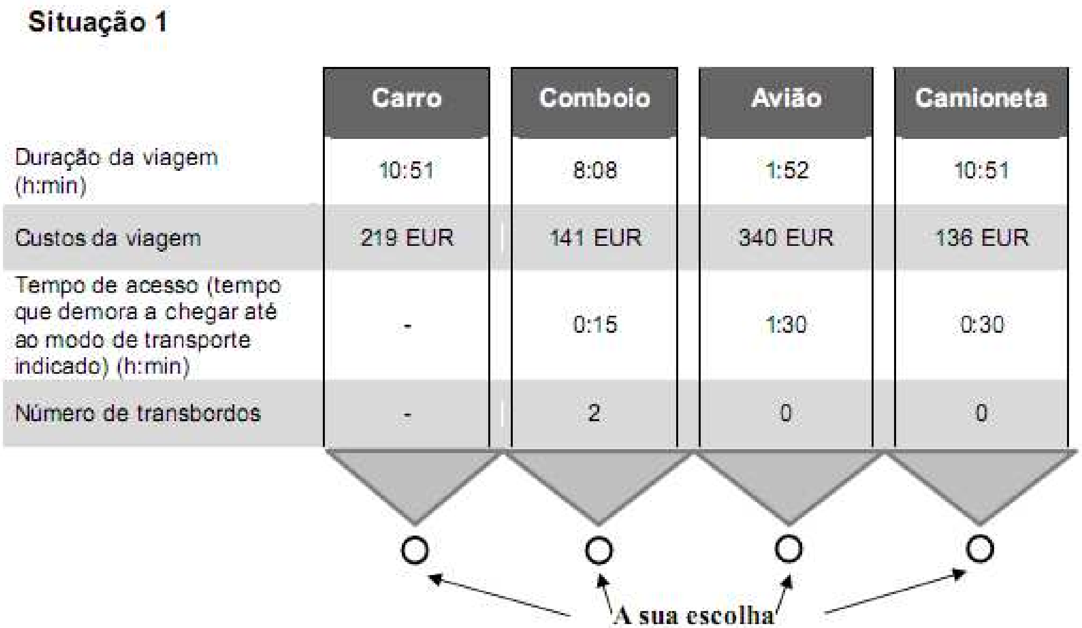
The characteristics of the SP-experiments used to describe the different alternatives are available in Table 11.

Table 11 Characteristics of the SP-Experiments

Alternative	Parameter	Characteristics
Car	Travel time	-20%, 0%, +20% of the observed travel time
	Costs	-20%, 0%, +20% of the imputed travel costs
Train, Plane and Bus	Travel time	-20%, 0%, +20% of the observed travel time
	Costs	-20%, 0%, +20% of the imputed travel costs
	Access time	5, 10, 15 minutes for the train 90, 120, 150 minutes for the plane 30, 60, 90 minutes for the coach
	Change (transfer)	0, 1, 2 times

Based on one of the reported journeys, the journeys' characteristics for the different modes were calculated (travel times from the IVT Air Network, IVT Road Network and IVT TransEuropean Train Model; cost from automatic internet requests, manually corrected). With these observed/imputed values and the given characteristics from the experimental design, the different choice situations for the SP-questionnaires were produced. An example is shown in Figure 8.

Figure 8 Example of a mode choice situation in Portuguese


Situação 1

	Carro	Comboio	Avião	Camioneta
Duração da viagem (h:min)	10:51	8:08	1:52	10:51
Custos da viagem	219 EUR	141 EUR	340 EUR	136 EUR
Tempo de acesso (tempo que demora a chegar até ao modo de transporte indicado) (h:min)	-	0:15	1:30	0:30
Número de transbordos	-	2	0	0

6.1.2 Theory of the choice models used

To analyze the behaviour and the reactions to the changes of the characteristics explained in the previous section, discrete choice models (Train, 2003) and in particular the Multinomial Logit-Model (MNL) is used. The estimates reported here were calculated with the software Biogeme (Bierlaire, 2003, 2009).

A Logit Model describes the utility of an alternative with a utility function, which is composed by a deterministic and a random (error) tem. The utility is generally described as following:

$$U_{jq} = V_{jq} + \varepsilon_{jq}$$

V_{jq}	deterministic term
ε_{jq}	error term

V_{jq} is a combination of the different attributes of the available alternative, the decision situation see Figure 7 for an example).

Based on these assumptions, the probability to choose a certain alternative can be calculated by:

$$P_j = \frac{e^{V_j}}{\sum_i e^{V_i}}$$

as the modelled probability for the choice of alternative *j*. To estimate the parameters of the utility function, the *Maximum-Likelihood*-method is used.

6.2 Model formulation

Based on a linear model for all relevant attributes of the utility function, non-linear interactions between those are formulated for all data and for the individual countries. The application of such interaction terms are explained in Mackie *et al.* (2003). The interactions used in the models estimated in this deliverable have the following formulation:

$f(y,x) = \beta_x \cdot \left(\frac{y}{\overline{y}}\right)^{\lambda_{y,x}} \cdot x$, where:

x	= observed variable, e.g. travel time, travel costs, …
β_x	= linear parameter of the observed variable *x*
y	= observed value for a different variable, e.g. distance,…
$\overline{y}$	= reference value of the variable *y*
$\lambda_{y,x}$	= elasticity of the dependence of the utility of the value of variable *y*

The choice of the reference value of $\overline{y}$ is arbitrary and has no influence on the estimated parameters or the model's goodness of fit. In the models the mean value of the parameters is used as a reference value, so the β_x parameters give directly, without consideration of $\lambda_{y,x}$,the utility of one unit of the variable *x* of their mean.

The general utility functions used in the model are given by:

For car:

$$V_{car} = \beta_{car} \cdot Constc + \beta_{Incomeunder2000} \cdot Incomeunder2000 + \beta_{Income2000to5999} \cdot Income2000to59999$$
$$+\beta_{\cos ts,i} f(Distance, Costs, i) \cdot Costs_{car} + \beta_{traveltime_{car},i} f(Distance, Traveltime, i) \cdot Traveltime_{car}$$

For train:

$$V_{train} = \beta_{train} \cdot Constt + \beta_{Incomeunder2000} \cdot Incomeunder2000 + \beta_{Income2000to5999} \cdot Income2000to59999$$
$$+\beta_{railwaycard} \cdot Railwaycard + \beta_{acctimet_{rain}} \cdot Accesstime_{train} + \beta_{change} \cdot Change_{train}$$
$$+\beta_{\cos ts,i} f(Distance, Costs, i) \cdot Costs_{train} + \beta_{traveltime_{train},i} f(Distance, Traveltime, i) \cdot Traveltime_{train}$$

For plane:

$$V_{plane} = \beta_{plane} \cdot Constp + \beta_{Incomeunder2000} \cdot Incomeunder2000 + \beta_{Income2000to5999} \cdot Income2000to59999$$
$$+\beta_{frequentflyer} \cdot Frequentflyercard + \beta_{acctime_{plane}} \cdot Accesstime_{plane} + \beta_{change} \cdot Change_{plane}$$
$$+\beta_{costs,i} f(Distance, Costs, i) \cdot Costs_{plane} + \beta_{traveltime_{plane},i} f(Distance, Traveltime, i) \cdot Traveltime_{plane}$$

For coach:

$$V_{coach} = \beta_{Incomeunder2000} \cdot Incomeunder2000 + \beta_{Income2000to5999} \cdot Income2000to59999$$
$$+\beta_{acctime_{coach}} \cdot Accesstime_{coach} + \beta_{change} \cdot Change_{coach}$$
$$+\beta_{costs,i} f(Distance, Costs, i) \cdot Costs_{coach} + \beta_{traveltime_{coach},i} f(Distance, Traveltime, i) \cdot Traveltime_{coach}$$

The observed values are the following:

$Incomeunder2000$ = 1, if the interviewee's income is under 2000 EURO, else 0,

$Income2000to5999$ = 1, if the interviewee's income is between 2000 and 5999 EURO, else 0,

$Railwaycard$ = 1, if the interviewee is a owner of a railway card (monthly, yearly), else 0,

$Frequentflyercard$ = 1, if the interviewee is a owner of a frequent flyer card, else 0,

$Accesstime_{Train}$ = Access time attribute of the alternative train in minutes,

$Accesstime_{Plane}$ = Access time attribute of the alternative plane in minutes,

$Accesstime_{Coach}$ = Access time attribute of the alternative coach in minutes,

$Change_{Train}$ = Changes attribute of the alternative train,

$Change_{Plane}$ = Changes attribute of the alternative plane,

$Change_{Coach}$ = Changes attribute of the alternative coach,

$Distance$ = Great-circle distance between origin and destination in km,

$Costs_{Car}$ = Cost attribute of the alternative car in EURO,

$Costs_{Train}$ = Cost attribute of the alternative train in EURO,

$Costs_{Plane}$ = Cost attribute of the alternative plane in EURO,

$Costs_{Coach}$ = Cost attribute of the alternative coach in EURO,

$Traveltime_{car}$ = Travel time attribute of the alternative car in minutes,

$Traveltime_{train}$ = Travel time attribute of the alternative train in minutes,

$Traveltime_{plane}$ = Travel time attribute of the alternative plane in minutes and

$Traveltime_{coach}$ = Travel time attribute of the alternative coach in minutes.

6.2.1 Linear models

All the models are estimated for the individual countries and for the whole data set, i.e. the combined data from all three countries. The first models are estimated with a linear formulation of the utility function for the relevant attributes. With a second formulation also the socio-demographic variables are incorporated into the utility function. Interactions are not incorporated in the first step as these models serve to gain first insights into the relevant attributes and to see if the results are arguable and consistent with expectations and theory. This means that in the model formulation in the previous section all the □-parameters, which are a part of the functions of the travel costs and travel time are set to zero.

6.2.2 Non-linear models

In a second step the nonlinear part of the models are added, which means that the □-parameter is no longer fixed and set to zero, but is variable. In these models the impacts of travel costs and travel time are estimated as dependent on the distance.

6.3 Results

The results are divided into different levels of details, e.g. for just a few parameters the data is split up by country. For the estimates of alternative specific parameters of the transport modes, the models are only estimated for all countries together. This is because of the number of data, which is available, which limits the number of parameters, which can be estimated on a significant level.

The results of the estimates are presented in tables 11 to 13.

The tables contain the values of the parameters as well as the t-statistics. The t-statistics is the ratio between the parameter estimate and its standard error. It measures if the parameter estimate is significantly different from 0. A t-statistic greater than 1.96 is significant on the 95%-level. The adjusted rho-square gives the proportion of the variance in the data, which can be explained by the model and is a measurement for its goodness of fit.

The first model is a linear model and can be used to check the plausibility of the model and the data with which the parameters are estimated (see Table 12). The estimated parameters have the expected sign. The access time and the number of changes have in addition to travel time and travel costs a negative influence on the probanbility to choose the particular alternative. The access time parameter is higher than that of travel time, which is expected.

The goodness of fit is with a range between 0.38 and 0.41 for a linear model already high.

Table 12 Results of the relevant characteristics with non-specific parameters for transport means

	All Data		Switzerland		Czech Republic		Portugal	
Parameter	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
β_{car}	2.060	14.79	1.350	5.13	2.240	13.04	2.580	2.33
β_{train}	1.740	13.59	1.720	7.22	1.640	10.58	3.340	3.19
β_{plane}	4.170	16.07	3.240	8.77	6.140	10.60	5.080	3.91
$\beta_{acctime}$	-0.022	-10.26	-0.019	-5.89	-0.026	-8.32	-0.018	-2.02
β_{change}	-0.174	-4.63	-0.220	-2.90	-0.164	-3.61	-0.344	-2.01
β_{costs}	-0.014	-17.56	-0.014	-11.06	-0.020	-9.75	-0.011	-3.33
$\beta_{traveltime}$	-0.010	-20.51	-0.011	-13.89	-0.010	-14.71	-0.008	-2.96
$Adj.\rho^2$ ¦ N	0.394	3108	0.378	916	0.415	2044	0.398	148
Number of parameters: 7								

To evaluate the differences between the modes for long distance travel a further model was estimated where the travel time and the access time parameters were estimated as specific for each mode whereas the travel costs and the number of changes are kept constant over the different transport means, which lets us estimate the differences between the modes with not that many more parameters. The results are shown in Table 13.

Table 13 Results of the relevant characteristics with mode specific parameters

	Car		Train		Plane		Coach	
Parameter	Value	t	Value	t	Value	t	Value	t
Constant	2.320	6.66	1.980	5.60	4.330	9.89	-	-
$\beta_{acctime}$	-	-	-0.100	-8.24	-0.016	-7.77	-0.057	-7.06
β_{change}	-	-	-0.073	-1.91	-0.073	-1.91	-0.073	-1.91
β_{costs}	-0.015	-19.05	-0.015	-19.05	-0.015	-19.05	-0.015	-19.05
$\beta_{traveltime}$	-0.013	-20.05	-0.011	-19.01	-0.026	-13.02	-0.009	-11.69
$Adj.\rho^2 =$	0.420	N=	3108					
Number of parameters: 12								

The differences between the modes are visible and not only in the constant, which shows the preferences of the users apart from the relevant characteristics. As expected travel time is evaluated more negative by the users as it normally takes longer to travel with a certain transport mode. So is the travel time for the plane not as highly negative as it is for a coach. With the additional mode specific parameters the goodness of fit has slightly improved to 0.42.

In a further step also the socio-economic characteristics of the interviewees were tested to see, if they have an influence on mode choice. Also the mobility tool ownership was included and tested in the models. Further inertia of mode choice was tested, which means to test, if the use of a certain mode of the interviewees for their real journey has an influence on the chosen alternative in the stated preference experiments. The inertia as well as the most sociodemographic characteristics, except income, were not significant in the model, but for mobility tool ownership of a frequent flyer card and a railway card subscription. The latter was only significant in Switzerland (beta = 0.282, t=2.52); as such a railway card is not common in Portugal and the Czech Republic.

Further interaction terms are introduced into the model, where income in interaction with travel cost was tested as well as distance with travel cost and travel time. The income interaction parameters, were not significant, but the distance interactions. This means, that the influence on the valuation of the travel time and the costs by the users is dependent on distance. The result of the non-linear model with socio-demographic characteristics of the interviewees is presented in Table 14.

Table 14 Non-linear model with socio-demographic characteristics

	Car		Train		Plane		Coach	
Parameter	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
Constant	3.290	7.03	3.360	6.92	5.330	9.43	-	-
$\beta_{acctime}$	-	-	-0.086	-6.92	-0.018	-8.50	-0.058	-7.34
β_{change}	-	-	-0.080	-2.09	-0.080	-2.09	-0.080	-2.09
β_{costs}	-0.015	-18.28	-0.015	-18.28	-0.015	-18.28	-0.015	-18.28
$\beta_{traveltime}$	-0.015	-22.11	-0.014	-16.25	-0.028	-12.65	-0.007	-5.79
$\beta_{Incomeunder2000}$	-	-	-0.272	-2.24	-0.272	-2.24	-0.272	-2.24
$\beta_{Income2000to5999}$	-	-	0.049	0.51	0.049	0.51	0.049	0.51
$\beta_{frequentflyer}$	-	-	-	-	0.607	2.91	-	-
$\lambda_{\text{Distance,Costs}}$	0.296	5.06	0.296	5.06	0.296	5.06	0.296	5.06
$\lambda_{\text{Distance,Traveltime}}$	-0.203	-2.38	-0.322	-2.88	-0.662	-7.15	0.0479	0.17
$Adj.\rho^2 =$	0.430	N=	3108					
Number of parameters: 21								

The interaction between distance and costs as well between distance and travel time (transport mode specific) show the expected sign. The linear parameters for all the relevant attributes are in the same range as in the linear model; which means that the parameter estimates are stable. In addition, the model was estimated using different optimization algorithms to avoid having a local minimum with one of the algorithm. But as the parameters were the same for the different algorithms, the global minimum seems to be found. The income is only significant for the lowest income group against the highest (6000 EUROS and more) and shows that those people use more often public transport instead of the car.

The inclusion of the elasticity terms has improved the goodness of fit slightly. The □^{2} increases to 0.43.

6.3.1 Values of travel time and trade-offs

The parameter ratios and their elasticities are beside the goodness of fit statistics and the parameter values of special interest. These ratios are called trade-offs and they show the willingness to trade in a unit of one variable for one of another variable. The parameter ratios in choice models can be interpreted as such trade-offs.

The most common used trade-offs in transport are the ones, which include a cost parameter. These trade-offs are interpreted as a willingness to pay. For example, the ratio between travel time and travel costs are the willingness to pay for saving one minute of travel time. In this deliverable we always use one hour as the reference. This ratio is well known as the value of travel time savings (VTTS).

In Table 15 presents the trade-off values by country based on Table 12. The first row shows the value of travel time-savings. The second one the price people are willing to pay for change the transport mode less often, the third one is the value of travel time savings for the access and the last one is the willingness of trade in one change for x minutes of travel time.

Table 15 Trade-offs by country and transport mean (Linear models)

Trade-off	Unit	All Data	Switzerland	Czech Republic	Portugal
$\beta_{\text{travel time}} / \beta_{costs}$	[EUR/h]	42.08	46.42	29.67	41.24
$\beta_{change} / \beta_{costs}$	[EUR]	12.08	16.06	8.20	31.27
$\beta_{acctime} / \beta_{costs}$	[EUR/h]	89.58	81.46	77.40	98.73
$\beta_{traveltime} / \beta_{change}$	[min]	17.23	20.75	16.58	45.50
Trade-off	**Unit**	**Coach**	**Car**	**Plane**	**Train**
$\beta_{\text{travel time}} / \beta_{costs}$	[EUR/h]	36.29	54.21	105.93	46.34

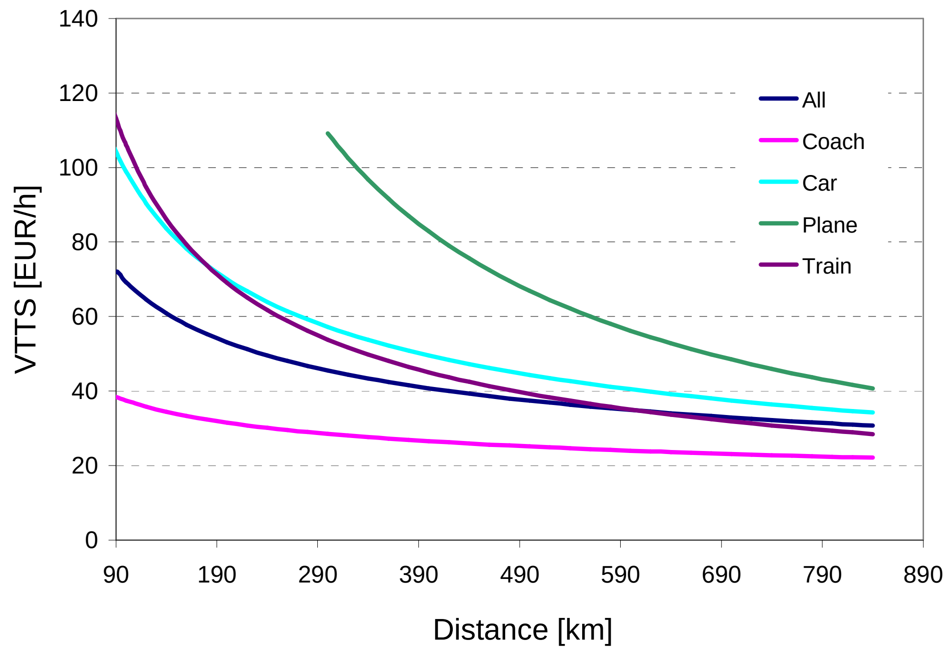
The trade-offs are over the countries more or less stable except that the ratios for the Czech Republic are smaller, which is expected in comparison to Switzerland for the trade-offs involving money, but not that much in comparison to Portugal.

The differences between the transport means are obvious and show the expected order for the ratios of the different transport mode. The lower the values are, the higher is the sensibility to price rises. Far higher are the ratios for the plane, as the plane is not as much in competition with other transport modes.

In contrast to the linear models trade-offs for the non linear models are not constant but dependent on distance. In Figure 9 plots the values of travel time-savings of the different transport modes against distance on the X-axis.

It can be seen, that the value of travel time-savings is much higher on shorter trips than on longer ones. This is the case, because the relative benefit from time saving is much higher the shorter the journeys are. The values for planes are only plotted from 300 km on, where the plane is used as a transport mode for normal trips.

Figure 9 Value of travel time savings by travel distance based on the non-linear models



6.3.2 Comparison of the results

For comparing the results presented in the last chapters, few values are available which presented the value of travel time savings also for trips with longer distances. In Switzerland a study about mobility pricing (Vrtic *et al.*, 2007) is one of these exceptions. The results of these studies are compared in Table 16. The difference is mostly due to the absense of plane and coach as an option in the Mobility pricing survey.

Table 16 Trade-offs in comparsion with other non long-distance travel studies

Trade-off	Unit	KITE	Mobility pricing
$\beta_{\text{travel time}} / \beta_{costs}$	[EUR/h]	44.7	19.2

6.3.3 Elasticities

The elasticity gives the relative change of a dependent variable, which's change results of an independent variable's change. In a MNL model this means the percentage change of the choice probability of an alternative when an attribute is changed by 1%.

The elasticity is described by:

$$\varepsilon_x = \frac{\partial P / P}{\partial x / x} = \frac{\partial P}{\partial x} \cdot \frac{x}{P},$$

ε_x = Demand elasticity relative to *x*

P = Choice probability

In a MNL model the elasticity is given by:

$$\varepsilon_x = \beta_x \cdot (1 - P) \cdot x.$$

The cross elasticity gives analogous to the normal elasticity the percentage change of the choice probability for a different alternative, which is given by:

$$\varepsilon_x = -\beta_x \cdot P \cdot x.$$

Table 17 shows the different elasticity's of the travel time and the travel costs and the cross elasticity of the cars' parameters on the other transport means.

This would mean e.g. that a 10% gain of the car costs for a trip, e.g. fuel costs would increase, its mode share would decrease in the short run by 11.5 %. The elasticities are larger then found in studies of daily travel, but are consistent with the experience of the low-budget airlines.

Table 17 Elasticity and cross elasticity of car on the other transport means (linear model) for the mean of the costs and travel time

Elasticity	Car	Train	Plane	Coach
Travel costs	-1.15	-0.93	-4.48	-0.97
Travel time	-5.02	-4.70	-2.44	-3.18
Cross elasticity of car on:		**Train**	**Plane**	**Coach**
Travel costs		0.28	0.26	0.02
Travel time		1.06	2.17	0.04

7 Conclusions

This deliverable has presented an initial analysis of the KITE survey methodology. The survey resulted in credible estimates of the number of long distance journeys, especially after a correction involving the information about the last long distance journey and the information about regular journeys. The response rates were in the expected range for a survey of this complexity.

The analysis of the stated preference survey revealed very strong elasticities for long-distance travel and high valuation for the avoidance of mode changes.

As the initial results from the analysis showed plausible results, there are still artefacts from the survey methodology which has to be analyzed further. The first one is the differences which are caused by the face-to-face interview compared to the CATI interview methodology. Dependent on that is also the lack of the numbers of long-distance journeys in the very low distance band. Numbers for the defined distance boundary over 100 km are more consistent if we look at the histogram in Figure 6, but still the analysis including the distance between 75 km and 100 km has shown, that such a lower boundary for the survey definition than for the actual long-distance definition is very useful.

The survey results will be analysed and compared to the numbers in deliverable D4 more precise in a further step, which was no more possible during the project time within KITE.

8 References

Armoogum, J. and J.L. Madre (1997) Accuracy of data and memory effects in home based surveys of travel behaviour, paper presented at the 76th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1997.

Axhausen, K.W. (1996) Befragungsmethoden im Fernverkehr: Neuere Erfahrungen, presentation at the DVWG Workshop zur Verkehrsstatistik, Heilbronn, September 1996.

Axhausen, K.W. (2003) Defining the scope of a long-distance survey, in K.W. Axhausen, J.-L. Madre, J.W. Polak and Ph.L. Toint (eds.) Capturing Long-Distance Travel, 7-26, Research Studies Press, Baldock.

Axhausen, K.W. (2007) Predicting response rate: A natural experiment, Arbeitsberichte Verkehrs- und Raumplanung, 434, IVT, ETH Zürich, Zürich.

Axhausen, K.W. and M. Youssefzadeh (2003) MEST survey work, in K.W. Axhausen, J.-L. Madre, J.W. Polak and Ph.L. Toint (eds.) Capturing Long-Distance Travel, 85-108, Research Studies Press, Baldock.

Axhausen, K.W., H. Köll, M. Bader and M. Herry (1997) Workload, response rate and data yield: experiments with long-distance diaries, paper presented at the 76th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1997.

Bierlaire, M. (2009) An introduction to BIOGEME 1.8, http://biogeme.epfl.ch.

BTS (1997) 1995 American Travel Survey: Profile United States, Bureau of Transportation Statistics, US Department of Transportation, Washington, D.C.

Chalasani, V.S., Ø. Engebretsen, J.M. Denstadli and K.W. Axhausen (2004) Precision of geocoded locations and network distance estimates, Journal of Transportation and Statistics, 8 (2) 1-15.

Denastadli, J.M. (1999) Travel behaviour 1998: Journeys of 100 km and more, Summary of TOI Report 466/1999, TOI, Oslo.

Hensher, D. and. M. Bradley (1993) Using Stated Response data to enrich revealed preference discrete choice models, Marketing Letters, 4 (2) 139-152.

Last, J, W. Manz and D. Zumkeller (2003) Heterogenität im Fernverkehr – wie wenige Reisen wie viel? Internationales Verkehrswesen, 55 (6) 267-273.

Mackie, P., M. Wardman, A.S. Fowkes, G. Whelan, J. Nellthorp und J.J. Bates (2003) Values of travel time savings in the UK, Institute for Transport Studies, University of Leeds und John Bates Services, Leed und Abingdon.

Rose, J.M., M.C.J. Bliemer, D.A. Hensher und A.T. Collins (2008) Designing efficient stated choice experiments in the presence of reference alternatives, Transportation Research B, 42 (4) 395-406.

Train, K. (2003) Discrete Choice Models with Simulation, Cambridge University Press, Cambridge.

Vrtic, M., N. Schüssler, A. Erath, K.W. Axhausen, E. Freijinger, J. Stojanivic, M. Bierlaire, R. Rudel, S. Scagnolari und R. Maggi (2007) Einbezug von Reisekosten bei der Modellierung des Mobilitätsverhaltens, Schlussbericht SVI 2005/004, Schriftenreihe, 1191, Bundesamt für Strassen, UVEK, Bern.