


# Workload, response rate and data yield

## Experiments with long-distance diaries

**Journal Article****Author(s):**

Axhausen, Kay W.  Köll, Helmut; Bader, Michael; Herry, Max

**Publication date:**

1997-01

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000024842>

**Rights / license:**

In Copyright - Non-Commercial Use Permitted

**Originally published in:**

Transportation Research Record 1593, <https://doi.org/10.3141/1593-05>

**Workload, response rate and data yield: experiments with long-distance diaries**

Presented at the 76th Annual Transportation Research Board Meeting

**KW Axhausen**

Institut für Straßenbau und  
Verkehrsplanung  
Leopold-Franzens-Universität  
Technikerstr. 13  
A - 6020 Innsbruck

Tel.: +43-512-507 6902  
Fax.: +43-512-507 2906  
EMail: k.w.axhausen@uibk.ac.at

**H Köll, M Bader**

Ingenieurbüro Köll  
A - 6060 Ampass

**M Herry**

Büro Dr. Herry  
A - 1040 Wien

December 1996

Paper presented at the 76th Annual Transportation Research Board Meeting

**Workload, response rate and data yield: Experiments with long-distance diaries**

KW Axhausen  
Institut für Straßenbau und Verkehrsplanung  
Leopold-Franzens-Universität  
A - 6020 Innsbruck

H Köll, M Bader  
Ingenieurbüro Köll  
A - 6060 Ampass

M Herry  
Büro Dr. Herry  
A - 1040 Wien

**Abstract**

Workload, response rate, data yield and data quality of travel diaries are interacting variables. It has long been suspected that it is impossible to maximise all variables at the same time. Still, empirical work trying to improve our understanding of the trade-offs between them has been very rare. This paper reports results from experiments with long-distance diaries, which aim to clarify some of the possible relationships.

The object of experimentation are surveys of long-distance travel behaviour, which are currently of particular interest in Europe and elsewhere. The development of the tourism industry, deregulation of the long-distance modes and infrastructure concerns require improved data about long-distance travel, both in the sense of improved inventories as well as in the sense of improved behavioural understanding.

The experiments undertaken here varied the workload of the respondents by varying the number of items to be reported about any long-distance journey, the duration of the survey period and the temporal orientation of the survey.

The results indicate that the response rate and the data yield, i.e. the number of reported journeys and stages, change systematically with changes in the experimental variables (reduced response rates for prospective surveys; reduced number of reported journeys and stages for retrospective surveys). Detailed results for these trade-offs are given in the paper. This trade-off forces the designer of such surveys to choose carefully and to invest time and effort in correcting for the potential biases resulting from this systematic behaviour.

**Keywords**

Travel diary - Long-distance - Austria - Experiments - Respondent work load - Data yield - Data quality

## INTRODUCTION

The designer of a travel diary survey has to balance a number of competing aims: he has to cover a certain set of data items to allow the development of pre-specified models or the estimation of particular pre-specified statistics; he has to maximise response rate to minimize possible non-response biases; he has to obtain complete, valid and error-free data. It is understood that trade-offs are necessary as the respondents will not necessarily be willing to provide the time and effort required to respond fully to the survey. The survey has to motivate the respondents and to enable the respondents to give a complete and error-free reply. Surprisingly there is very little published empirical work, which would allow the designer of a new survey to optimize his designs a-priori. This paper is a contribution to this literature reporting results on the interaction between respondent work load and the data yield of a long-distance travel diary instrument.

The workload of the respondent in a travel diary is the product of the number of journeys undertaken during the survey period times the number of items requested for each journey plus the rather less variable "overhead" of questions about the household, person and vehicles. It is assumed that the willingness of the respondent to participate depends on the workload perceived by respondent and on the motivation of the respondent created by the topic of the study or any supporting material which is included in the survey or provided by direct contact with the respondent. It is clear, that the perceived workload will be very high for highly mobile persons, but that they might be especially motivated due to their large exposure to the transport system. It is now well-known, that special efforts are required to correct the biases resulting from the non-response of both highly mobile and non-mobile persons (Ettema, Timmermans and van Veghel, 1996).

In addition to the response "Not to respond at all", the respondent is able to reduce his workload by omitting complete journeys, tours or trips within journeys or stages within trips. It is therefore necessary to give special attention to this data yield as it might indicate differential biases due to this specific underreporting.

There is currently an increasing policy interest in the amount and type of long-distance travel undertaken in Europe. This reflects the growing economic importance of tourism widely understood and the growing concern about the environmental impacts of traffic, in particular by road and air. Additionally, the current policy initiatives of the European Union in the context of the development of the "Transeuropean Networks" (Deiss, 1996) require data for planning and evaluation. The work reported here is part of both a national and European initiative to improve the methods for the collection of long-distance data. The national background is the development of a national infrastructure and policy programme for all transport networks in Austria (Kovacic, 1996). It was felt, that for the future development of that programme improved methods for the collection of long-distance travel data are a necessity. At the same time EUROSTAT, the statistical office of the European Union, is coordinating studies to test a common set of definitions for long-distance travel diary surveys (EUROSTAT, 1995a).

The paper is structured in the following way: the next section will discuss the methodological approach, while the following section will present the survey instruments developed. The results of the initial analysis will be the focus of the main section followed by conclusions, which will also outline the further work planned.

## APPROACH

The task of the experiments was to study the relationships between respondent work load, response rate and data yield and quality. The work load of the respondents is a function of the following variables under the control of the survey designer:

- *Object definition* is the exact definition of the object(s), which the respondents have to recollect and describe in the survey. While there is a large degree of standardization of this definition for urban/regional travel diaries, this is not the case for long distance surveys (Youssefzadeh and Axhausen, 1996). The definition has to cover the following main points:
  - the minimum distance for the furthest destination, which is required for inclusion as a long distance journey
  - the structure of the reporting, i.e. whether the journey has to be reported as stages, trips or only as complete journeys and their definitions in turn.
  - treatment of sidetrips on the way to and from the destination
  - treatment of movements at the destination
- *Number and detail of items* involves the specific items queried for each journey, trip or stage of the journeys, but also the detail of the coding involved.
- *Temporal orientation* concerns the choice of a retrospective or a prospective format. Prospective means in this context, that the respondent receives the survey form before the start of the survey period and replies as the time passes. In a retrospective survey the work is concentrated in one episode, while a prospective survey allows the respondent to distribute the work involved over multiple episodes. This in turn should influence both response rate and data yield.
- *Survey period* is the duration of the reporting period. This is especially crucial if the survey is retrospective, as memory effects have to be considered

The study was framed as a contribution to the on-going EUROSTAT-pilots (EUROSTAT, 1995a). This provided to a large extent the object-definition, as participation in the pilots required the adoption of a common object definition: 100 km minimum, stage-based description of the journeys, inclusion of all side trips, but exclusion of all trips at the destination. The common definition also provided a minimum set of items and their coding detail.

To test the interactions it was decided to vary the other factors in a complete  $2^3$  factorial design. The attributes chosen were:

- *Temporal orientation*: prospective and retrospective

- *Survey period*: 4 weeks and 8 weeks to provide enough spread, without either extending it to a duration where recall problems could be expected<sup>1</sup> or to be so short, as to have too many respondents without any long-distance journey.
- *Items and their level detail*: in addition to the set defined by EUROSTAT, a second more extensive set was developed, which covered each item in more coding detail, as well as adding further items reflecting concerns raised by other EU-policy initiatives (EUROSTAT, 1995b).

To avoid boundary problems a reduced minimum distance of 75km was chosen.

Due to resource constraints it was decided to execute the main survey as a mail out/mail back supplemented by a telephone based non-response interview, as well as telephone interviews with respondents about their experience.

The mail out/mail back survey consisted out the following elements:

<i>Prospective surveys</i>	<i>Retrospective surveys</i>
Announcement letter	Announcement letter
Survey distribution	Survey distribution
2 letters during the survey period reminding the respondents	
Reminder letter	Reminder letter
Redistribution as retrospective survey	Redistribution

This survey protocol is in-line with current practice in urban travel diary survey work and is required to assure reasonable response rates (Brög, 1996).

The non-response interviews began two weeks after the redistribution of the surveys, while the telephone interview of the respondents started immediately after the receipt of the surveys. Up to five contact attempts were undertaken before the address was classified as non-reachable. The NR-interviews covered basic socio-demographics including income, the amount of long-distance travel including the basic details of date, destination and travel party and the reasons for non-participation, while the response-interviews covered income, corrections to the coded journeys and the assessment of the survey by the respondent.

The sample was developed from a random sample of 1500 young to middle-aged adult males in the city of Innsbruck provided by an address dealer. Other, official, sampling frames were not available due to legal restrictions. Excluded were persons without a telephone number in the latest CD-ROM version of the official telephone book, as well as persons who could not be found at the address given by the address dealer, as these could not be guaranteed to be the same person. Of the 1297 names remaining, 200 were used for a pretest not reported here and 1080 for the main survey.

The survey period covered March to June 1996 with the last NR-interviews taking place in July.

---

<sup>1</sup> The initial results of the French NTPS had indicated strong recall problems for periods over 8 weeks.

## SURVEY INSTRUMENT

The survey instrument was designed from scratch, as no comparable survey instruments had been used before in Austria, or Europe. It was structured into four parts distributed across two instruments:

- *Household details* covering household details and the description of all members of the household.
- *Vehicle details* covering the description of the vehicles of the household.
- *Journey details* describing each journey as a whole (main purpose, main destination, names of accompanying household members, number of other accompanying persons)
- *Stage details* covering each stage (movement with one mode) and some aspects of the stay at the destination. The set of items and their coding level were varied to vary the work load.

The items included are shown in 1. These were grouped into a:

- *Household and vehicle form* (A3 portrait form folded to an A4 booklet<sup>2</sup>).
- *Journey form* (A4 booklet with space for eleven journeys of up to five stages). The front page gave the minimum instructions, including survey period, plus the number of a help line. The backpage asked for the last journey undertaken just before and the first planned after the survey period.

This was supplemented by a:

- *Cover letter* signed by the first author stressing the random nature of the sample and the confidentiality of the data received and describing the background of the study in rough terms.
- *Policy background flyer* (A4 folded to \_ A4) explaining the survey and the policy background in more detail.
- *Explanations and examples* (A4 booklet) giving more detailed definitions of the survey object and an overview calendar of the survey period on the front page and three examples of growing complexity inside.

The journey form was for all members of the household. Each member was asked to record his/her journeys in this one booklet and to mention all other members accompanying him/her. This design was chosen, as the pretest had shown that households resented the multiple entry of identical journeys.

The design of the journey form and the distribution of the journey and stage information was difficult and a number of alternatives was tested in informal focus groups. The final version adapted the column design familiar from the KONTIV-design (Axhausen, 1995). The six available columns were allocated to one column of journey-level questions and five columns for a stage each. Heavy arrows were needed to guide the respondents from the bottom of the column to the start of the next and even heavier arrows were required to direct respondents to continue a journey overleaf (See 1). The purpose *Wohnen* (return home) was added as a stage purpose to remind the respondents that their journey had ended and to avoid spurious answers on the last stage.

---

<sup>2</sup> A3 and A4 are standard paper sizes (42,0 cm \* 29,7 cm and 29,7 cm \* 21,0 cm respectively)

Given the sampled population the lettering was chosen at a size, which would be unsuitable for a general population survey. In this case, in particular with the same number of questions and codes, a different format would have to be chosen. A possibility would then be to dedicate a whole page to one stage (see also Richardson, Ampt and Meyburg, 1995).

To minimize the initial exposure to definitions, explanations and examples the design distributes these over three locations. This allows the "browsers" to go straight to the form without being detracted or discouraged by these elements, while the "readers" can study them at length (White and Collins, 1996)<sup>3</sup>. The journey form states only the bare essentials (minimum distance, survey period, coverage of out- and return-trip and of stages). The front page of the explanation booklet gives detailed definitions in as easy a language as possible, but still not covering all possible complexities and exceptions. The examples range from a simple car-only shopping journey, to a more complex one-day business trip involving multiple modes and a multi-day, multiple mode holiday journey. They cover some of the complexities in their annotations of the examples. This three-tier concept allows the respondents to match their more detailed questions against more detailed material, but might give some respondents the impression of inconsistency, as the less detailed instructions hide complexities revealed only in the more complex ones. Further testing is required here.

The announcement letter and the return envelope were franked with a stamp, while all other communications were franked by machine.

The respondents were offered a small incentive in the form of a ÖS 20 (US\$ 1.90) donation to a charity of their choice from a list of six, which offered institutions with environmental, third-world and social, global and local foci.

## **RESULTS**

The results presented here focus on response rates and data yield in terms of their descriptive analysis and of suitable disaggregate modelling. The substantive content of the results and the description of the respondents will be kept to minimum.

### **Sample and journey characteristics**

2 gives the income distribution of the households, while the household members mentioned in the returned household forms are described in 3. The results match the expectations for the sample drawn. No differences between the experimental groups could be detected. Characteristics of the journeys are given in 4

---

<sup>3</sup> About a quarter of all respondents judged the booklet and other materials as useful, whereas only 3% claimed to have been confused by them.



Figure 2 Example page journey form

---



Table 5 Items covered

Area	Items
Household	Location of main residence
Person	First name Year of birth Sex Marital status Nationality Education achieved Current job description Selfemployment Current working hours/week Working hour flexibility Car driving licence ownership Public transport discount ticket ownership Frequent flyer program membership Car sharing club membership
Vehicle	Kind Type Marque cc Year of production Type of fuel Catalytical converter Mileage in the last twelve month Current odometer reading Main user Owner Leasing status
Journey	Departure time and date Place of departure Main purpose (You wouldn't have travelled without this activity) Main destination First names of all household members travelling Number of other accompanying persons
Stage	Mode (small + large set) Class (train and air) (large set) Travel paid by (large set) Destination (small + large set) Arrival time and date (small + large set) Purpose (small + large set) (multiple purposes possible) Type of overnight accommodation, if any (large set) Accommodation paid by (large set) Departure time and date (small + large set)

Table 6 Persons: socio-demographic characteristics (responding households only)

Characteristic	Share [%]
–	
Sex	
Male	54.7
Female	45.3
Age	
under 18	31.7
18-30	9.9
31-40	41.8
41 and over	16.7
Nationality	
Austrian	98.0
Other	2.0
Marital status	
Married	54.7
Single, Divorced, Widowed	45.3
Highest qualification	
Not completed yet	27.6
Compulsory schooling	11.0
Apprenticeship	18.3
Baccalaureate	13.5
Technical college	13.9
University	15.7
Working status	
Not working	11.0
In education	35.3
Working	53.7
Working hours	
Part time	22.7
Full time	77.3
Working pattern	
Fixed	61.4
Flexible	30.9
Shift	7.7
Selfemployed	8.0
Driving licence (of adults)	90.4
ÖBB Half-Price pass	14.1
ÖBB Family pass	6.2
ÖBB <i>Kilometerbank</i> (mileage prepurchased at a discount)	1.5
Regional six monthly season	3.2
Regional annual season	4.1
Member in the AUA frequent flyer programme	0.7
Member in other frequent flyer programmes	0.9
Membership in car sharing programme	0.4

Table 7 Characteristics of the journeys (within survey period only)

Characteristic	Share [%]
—	
Destination	
Austria	66.0
Germany	12.9
Italy	16.5
Other	4.6
Departure day	
Monday	15.2
Tuesday	12.6
Wednesday	12.6
Thursday	12.9
Friday	20.0
Saturday	17.3
Sunday	9.5
Return day	
Monday	14.2
Tuesday	12.6
Wednesday	11.4
Thursday	13.0
Friday	12.9
Saturday	10.8
Sunday	24.9
Main purpose	
Business	35.5
Education	6.5
Escorting	1.3
Leisure	21.4
Private business	1.2
Private service	0.8
Shopping	2.9
Visiting	24.7
Work	4.0
Other	1.8
Main modes	
Air	6.6
Regional public transport	0.3
Cars etc.	73.6
Long distance bus	1.6
Long distance train	17.1
Non-motorized modes	0.2
Sea	0.7
Mean duration of stay	3.6 days
Mean stage durations	9.64 hours
Mean size of party	2.63 persons
Mean number of household members	1.66 persons
Mean number of non-household members	0.98 persons

Table 8 Distribution of household income [ÖS/month]

Income class [%]	Households responding	Households interviewed only
—		
Known	63.7	80.0
20000 and below	34.0	28.9
20000 to 40000	56.9	59.0
40000 and above	9.1	12.1
Unknown/refused	36.3	20.0
N	328	195

## Response

9 details the overview of the response behaviour for the eight experiments. Sample loss is the number of survey packs returned by the post or new tenants as "Addressee unknown" or "Wrong address". The response rate is the share of those remaining after the sample loss returning usable forms. The rate of unreachables is the share of households which we tried to reach by phone for either type of telephone interview and which could not be reached within five attempts. The rate of non-response (NR) interviews is the share of households eligible for a NR-interview, which did participate (i.e. excluding households which we could not reach by phone). The rate of response-interviews is defined analogously (10). 11 gives the same information, but in absolute numbers.

Table 12 Response behaviour

Experiment Orientation	Survey Stage period item set		Response rate [%]	Unreach- able by phone [%]		Share of	Share of
				NR-Inter- views [%]	R-inter- views [%]		
—							
Prospective	4 weeks	Small	25.6	20.5	46.5	100.0	
Prospective	4 weeks	Large	33.3	26.7	42.4	96.6	
Prospective	8 weeks	Small	26.7	22.5	40.6	100.0	
Prospective	8 weeks	Large	27.6	29.1	31.2	100.0	
Retrospective	4 weeks	Small	39.0	25.2	37.5	91.7	
Retrospective	4 weeks	Large	37.2	16.5	50.8	97.4	
Retrospective	8 weeks	Small	38.3	24.2	26.4	97.8	
Retrospective	8 weeks	Large	38.0	18.6	36.5	100.0	

Table 13 Response behaviour

Return	Telephone contact	Prospective				Retrospective			
		4 weeks		8 weeks		4 weeks		8 weeks	
		Small	Large	Small	Large	Small	Large	Small	Large
—									
Sample loss		18	15	15	8	12	14	7	6
No	Not reached	16	21	18	31	17	13	24	15
No	Yes	29	23	21	18	20	30	13	21
No	Refused	36	25	33	35	32	29	33	37
Empty	Not reached	0	0	1	0	2	0	2	2
Empty	Yes	4	2	7	1	1	2	1	2
Empty	Refused	2	9	8	7	3	2	6	3
Yes	Not reached	8	11	8	6	12	7	5	7
Yes	Yes	22	28	24	29	33	37	43	42
Yes	Refused	0	1	0	0	3	1	1	0

The response rate of 33.2%, on average across the experiments, is in the range expected in Austria, but the total response (return of the form or a full NR-interview covering journeys made) is 52.7% of all respondents (excluding the sample loss). Using linear regression to analyse the aggregate rate of response only the temporal orientation had a significant effect on the response rate. Prospective surveys have a response rate 9.8% (t-statistic 5.12; adjusted  $r^2 = 0.78$  with only the three main effects) lower than retrospective surveys; but see below for a disaggregate analysis. One can only speculate about the reasons for this drop: it is possible that the households are just less willing or less

able to organize the filling-in of the survey form over such a long period; that the expected workload and its possible variability is perceived as too high a risk in comparison to the retrospective surveys, where the respondents can assess the likely workload better.

The share of households unreachable by phone is independent of the characteristics of the survey, as would be expected (grand mean 22.9%).

The regression analysis of the share of NR-interviews indicates, that the duration of the survey period has a marginally significant effect (-10.6% for eight weeks in comparison to four weeks;  $p = 0.09$ ). It seems obvious that respondents will be less likely to participate, if the interview can be expected to be longer due to the longer survey period, while the temporal orientation of the original written instrument should have no influence at this point.

The availability of the data from the non-response interviews allows a more detailed, disaggregate analysis of the non-response behaviour, as we know through these interviews some of the relevant information, which might have predisposed them to non-response. In particular, we know about the long-distance travel behaviour of the initial non-respondents and it is reasonable to assume that persons with many long-distance journeys will not participate. The analysis performed here follows Polak and Ampt (1996), which adapted an idea from Kitamura and Bovy (1987) for this context. The central idea of this analysis is to test for the impact of the amount of long-distance travel on the response probability in conjunction with the other socio-demographic variables available. Two steps are involved: the first step reported below (see Section ) relates the socio-demographics to the reported number of journeys, while the second step reported here relates the socio-demographics, the predicted number of journeys and the observed residual to the probability/willingness to return the survey form.

The probability, that a household will return the form, is estimated as a probit model (See 14) using once the number of household journeys and in the other case the number of person journeys<sup>4</sup> as the indicator variable for the travel behaviour of the household. The model using the person journeys is to be preferred on the ground of better statistical fit, but the conclusions are nearly the same in any case.

The parameter value of the temporal orientation confirms that the recipients of retrospective surveys are more likely to respond. While the effect of the complexity of the stage questionnaire had not been significant at the aggregate level, it is so at the disaggregate level. Surprisingly, an increase in complexity increases the likelihood of participation all other factors being equal.

---

<sup>4</sup> A journey involving two members of the household would be counted as one household journey, but as two person journeys.



Table 15 Response probability: disaggregate analysis (all households responding)

Variable	Household journeys		Person journeys	
	Parameter	t-statistic	Parameter	t-statistic
Retrospective	0.487	2.21 *	0.516	1.96 *
Weeks of survey period	0.104	1.97 *	0.105	1.49
Large set of movement items	0.368	2.38 *	0.349	2.29 *
Redistribution required	-1.517	-4.69 *	-1.563	-4.21 *
Income (in classes)	-0.349	-1.77 +	-0.306	-1.65 +
2 adult household	0.226	1.02	0.148	0.49
3 adult household	0.576	1.42	0.488	1.23
4 adult household	0.764	0.95	0.630	0.67
Estimated number of journeys	0.195	0.71	0.111	0.45
Observed residual	0.055	1.18	0.042	1.60
N		357		357
$\rho^2$		0.17		0.32
_(0)		-242.8		-242.8
_(B)		-182.8		-182.4
Correct predictions [%]		76.2		74.8

Due to the rather restricted set of questions of the non-response interviews, the socio-demographic description of the households is poor. Household size has no significant impact on participation, while income has a significant positive effect. Neither a large number of predictable journeys, nor the observed residual increases the participation probability significantly. The results indicate, that active household might appreciate the opportunity to participate in such surveys and that a complex form indicates to them the seriousness of the study and of the analysis.

### Data yield

The response rate is only one part of the equation. The second part is the amount of information obtained from the respondents. 16 summarizes the data yields as a function of different socio-demographic characteristics. The pattern of differences in the travel behaviour are as expected. The overall amount is in the range reported from other studies and countries (see for example ÖSTAT, 1995).

The effect of the experimental variables on the data yield is subject to a disaggregate analysis at the level of the household, the person and the journey. Using the negative-binomial model, an extension of the Poisson model for count data (Greene, 1995), the analysis describes the number of journey, respectively stages as a function of the characteristics of the experimental design, and of the household, person and journey, respectively.

Table 17 Travel participation by person characteristics (within survey period only; households replying only)

Characteristic	Unweighted Journeys/week and person
–	
All	0.13
Age	
17 and under	0.08
18 to 30	0.20
31 to 40	0.19
41 and over	0.21
Sex	
Male	0.18
Female	0.11
Working status	
Not working	0.09
In education	0.07
Working	0.19
Car licence	
Yes	0.20
No	0.09
Highest educational qualification	
None	0.09
Compulsory schooling only	0.10
Apprenticeship	0.14
Baccalaureate	0.24
Technical college	0.16
University	0.28

18 shows the results for the household level analysis. Judging by the explanatory power the person journey analysis is to be preferred. Concentrating on this analysis, the results show, that a retrospective surveys leads underreporting of person journeys in comparison with prospective surveys. Equally the more complex survey form reduces the number of journeys reported, although neither effect is significant for the person journey analysis. Income has a significant positive impact on the amount of travel, while the household size impacts the amount of travel negatively, although none of the parameters is significant. The significant overdispersion parameter indicates that the negative binomial analysis is to be preferred in comparison with the Poisson model.

Repeating the analysis at the person level (including only those persons reported on the survey forms) allows a richer socio-demographic description of the process (19). The analysis is presented for all persons and for the males adults, as they have the dominant share of long-distance travel.

Table 20 Household: disaggregate analysis of travel behaviour (all households reporting)

Variable	Number of journeys in the survey period			
	Household journeys		Person journeys	
<i>Negative binomial model</i>	Parameter	t-statistic	Parameter	t-statistic
Retrospective	-0.221	-1.15	-0.254	-0.90
Weeks of survey period	0.072	2.25 *	0.092	2.31 *
Large set of movement items	-0.145	-0.97	0.008	0.04
Redistribution required	-0.77	-4.31 *	-0.627	-2.83 *
Income (in classes)	0.303	3.19 *	0.187	1.49
2 adult household	-0.216	-0.97	0.381	1.23
3 adult household	-0.449	-1.06	0.146	0.29
4 adult household	-0.221	-0.37	0.625	0.59
Overdispersion parameter $\alpha$	0.746	4.83 *	1.589	7.60 *
N		357		357
$\rho^2$		0.17		0.32
_(0)		-643.4		-989.1
_( $\beta$ - Poisson)		-574.4		-871.9
_( $\beta$ - binomial)		-521.9		-673.7

Again, the retrospective surveys has a significantly lower data yield in comparison to the prospective survey. The complexity shows no significant effect. The socio-demographic variables reveal the expected patterns of effects on travel behaviour. Interesting to note is that at this initial stage of the diffusion of frequent flyer passes, they have no significant effect yet; probably indicating a substantial number of people just experimenting with them.

The final element of the yield is the number of stages reported per journey. The disaggregate analysis (See 21) confirms that a retrospective survey reduces data yield significantly, as does the complexity of the description. An influence of the survey duration is not detectable.

The disaggregate analysis corrects for the characteristics of the journey and of the person undertaking it. Only those modes/purposes were retained, which had a significant impact in comparison of the other modes/purposes. Education level had no impact, but being in education or working had a strong positive impact on the reported number of stages.

Table 22 Person: disaggregate analysis of travel behaviour (excluding non-response interview information)

Variable	Number of journeys in the survey period					
	All persons			Male adults		
<i>Negative binomial model</i>	Parameter	t-statistic		Parameter	t-statistic	
Retrospective	-0.242	-2.44	*	-0.237	-1.99	*
Weeks of survey period	0.130	5.74	*	0.143	5.24	*
Large set of movement items	0.005	0.06		-0.041	-0.37	
Redistribution required	-0.256	-2.20	*	-0.327	-2.30	*
Males	0.252	2.65	*	-	-	
Frequent flyer card	0.498	0.85		0.571	0.78	
Rail discount or season ticket	0.131	1.19		0.157	1.21	
Driving licence	0.222	1.26		0.147	0.70	
Single	-0.114	-0.92		-0.102	-0.74	
In education	-1.253	-5.26	*	-1.095	-3.80	*
Working	-0.946	-3.77	*	-0.721	-2.39	*
Compulsory schooling only	-0.039	-0.22		-0.071	-0.29	
University education	0.421	3.94	*	0.526	4.11	*
Overdispersion parameter $\alpha$	0.365	4.64	*	0.344	3.86	*
N		768			476	
$\rho^2$		0.14			0.15	
_(0)		-1151.6			-773.3	
_( $\beta$ - Poisson)		-1020.2			-675.6	
_( $\beta$ - binomial)		-996.1			-659.0	

The analysis so far has focused on the total number of movements reported, but has ignored how they were spread over the different survey periods. 3 shows the share of persons departing for a long-distance journey on any of the observed days by the temporal orientation of the survey they participated in. No fatigue effects are obvious. The difference in reported journeys between the retrospective and the prospective is concentrated in the weeks immediately at the start of the survey periods (see Easter).

The level of departures during the week, mostly business, remains rather constant, also indicating the lack of fatigue effects. Still, a more detailed statistical analysis is necessary before fatigue effects can be ruled out. The interactions between the different legal holidays (Easter Monday, 1. May, Ascension day, Monday after Pentecost), the school vacations/university term breaks, the different survey periods and the temporal orientation of the survey need more work.

Table 23 Stages: disaggregate analysis of travel behaviour (excluding non-response interview information)

Variable	Number of stages reported by journey					
	All persons			Involving adult males		
<i>Negative binomial model</i>	Parameter	t-statistic		Parameter	t-statistic	
Retrospective	-0.331	-5.10	*	-0.325	-4.54	*
Weeks of survey period	-0.003	-0.24		0.009	-0.63	
Large set of movement items	-0.118	-2.36	*	-0.125	-2.23	*
Redistribution required	0.086	1.10		0.066	0.73	
Males	-0.023	-0.46		-	-	
Single	0.093	1.63		0.125	2.06	*
Compulsory schooling only	-0.017	0.17		-0.130	-0.67	
University education	-0.064	-1.26		-0.146	-2.43	*
In education	1.151	8.64	*	0.987	6.16	*
Working	1.260	11.15	*	1.214	9.70	*
Journey abroad	0.166	2.98	*	0.156	2.48	*
By bus	0.616	5.00	*	0.591	4.59	*
By long-distance train	0.274	4.19	*	0.265	3.47	*
By sea	1.061	9.50	*	1.226	10.12	*
By air	0.932	14.85	*	0.919	13.02	*
For education	0.187	1.81	+	0.282	2.45	*
For private business	0.577	5.24	*	0.311	2.33	*
Overdispersion parameter $\alpha$	0.038	2.49	*	0.032	1.92	+
N		746			539	
$\rho^2$		0.14			0.15	
_(0)		-1670.2			-1215.6	
_( $\beta$ - Poisson)		-1440.4			-1035.8	
_( $\beta$ - binomial)		-1433.5			-1032.1	

#### CONCLUSIONS AND DIRECTIONS FOR FURTHER WORK

The work reported above shows that the survey designer faces trade-offs between the workload of the respondent, the response rate(s) and the data yield obtained from the survey: retrospective surveys increase the willingness of the respondents to participate, but the amount of travel and stage detail is reduced in comparison with a prospective survey. Increasing the duration of the survey period could induce fatigue, but will still increase the total number of journeys reported.

In addition, a longer survey period decreases the probability that the respondents will participate in NR-interviews, which would allow the collection of at least some information about the person and its travel.

Given the quality advantages of prospective surveys confirmed here for long-distance surveys, although long known for urban surveys, future experimental work should focus on how to increase the response rate for such surveys in order to avoid any systematic bias due to respondent self-selection, although the modelling results indicate that we could correct for some of the selfselection (increased contact during the survey period via phone, improved motivational materials etc.). It is necessary to develop methods or designs, which reduce the apparent workload for the respondent to overcome any initial resistance to participating in the survey. The other major area of work has to be the comparison with telephone surveys and or methods combining reduced forms (memory joggers) with telephone retrieval.

A different avenue is a meta-analytic approach to the results reported in the literature, as it will be impossible to conduct all the necessary experiments as research studies. Unfortunately, this will require very substantial work, as for example, even response rates are not reported or reported wrongly.

#### ACKNOWLEDGEMENTS

The survey work reported here was undertaken under contract of the Austrian Ministerium für Wissenschaft, Verkehr und Kunst as part of the on-going development of the Bundesverkehrswegeplan (Federal Transport Infrastructure Plan). The authors wish to thank the project officers, Dr. Kovacic and Dr. Spiegel, for their helpful comments and suggestions.

Some of the analyses presented was carried out as part of the 4th Framework Transport Research programme project "Methods for European Surveys of Travel Behaviour" to guide the design of the survey work to be undertaken for this project. The project involves the Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität Innsbruck (coordinator), Statistics Netherlands, Büro Herry (Wien), University of London Centre for Transport Studies, Imperial College (London), Deutsche Versuchsanstalt für Luft- und Raumfahrt (Köln), INRETS (Arceuil), Transportes, Inovação e Sistemas (Lisbon), TØI (Oslo), Transport Studies Group Facultés Universitaires Notre Dame de la Paix (Namur), Socialdata (München), Statistics Sweden, TU Delft. All conclusions drawn are solely those of the author(s) and not of the CEC or the Consortium.

#### LITERATURE

- Axhausen, K.W. (1995) Travel diaries: an annotated catalogue, 2nd edition, *Arbeitsbericht*, **18**, Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität, Innsbruck.
- Brög, W. (1996) Das neue KONTIV-Design, presentation to FGSV-Committee AA 1.11, April 1996, Stuttgart.
- Deiss, R. (1996) Presentation of policy initiatives and data needs of DG VII, presentation at the 1. Meeting of the Concerted Action on Transport Information Systems, Brussels, May 1996.

- Ettema, D., H. Timmermans and L. van Veghel (1996) Effects of data collection methods in travel and activity research, report to the Dutch Ministry of Transport, European Institute of Retailing and Service Studies, TU Eindhoven, Eindhoven.
- EUROSTAT (1995a) Proposal for definitions and variables of a household survey for mobility, in *Minutes of the Task Force on Passenger Transport and Tourism Statistics*, Luxembourg, 31. 1.-1. 2. 1995.
- EUROSTAT (1995b) Implementation of the EUROSTAT methodology on basic tourism statistics: A practical manual, 2nd draft, EUROSTAT, Luxemburg.
- Greene, W.H. (1995) *Limdep Version 7.0*, Econometric Software, Bellport.





Figure 4 Share of persons departing [% of all persons reported]

---

---

- Kitamura, R. and P.H.L Bovy (1987) Analysis of attrition biases and trip reporting errors for panel data, *Transportation Research*, **21A** (4/5) 287-302.
- Kovacic, W. (1996) Stand des Bundesverkehrswegeplans, presentation, Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität, Innsbruck, June 1996.
- ÖSTAT (1995) Reisegewohnheiten der Österreicher im Jahre 1993, *Beiträge zur Österreichischen Statistik*, **1.184**, Österreichisches Statistisches Zentralamt, Wien.
- Polak, J.W. and E.S. Ampt (1996) An analysis of response wave and non-response effects in travel diary surveys, paper presented at 4th International Conference on Survey Methods in Transport, Steeple Aston, September 1996.
- Richardson, A.J., L. Ampt and A.H. Meyburg (1995) *Survey Methods for Transport Planning*, Eucalyptus Press, Melbourne.
- White, A. and D. Collins (1996) Redesigning the Census form using cognitive interviewing methods, *Survey Methods Centre Newsletter*, **16** (2) 14-20.
- Youssefzadeh, M. and K.W. Axhausen (1996) Long distance diaries today: Review and critique, *Deliverable*, **D1**, MEST-Project, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.