PAPER

Methodological Research for an European Survey of Long-Distance Travel

K. W. AXHAUSEN
ETH IVT, Zürich

ABSTRACT

Long-distance travel has in comparison with its share of trip making (1 percent of all trips) a disproportionate impact in terms of miles traveled, costs and subjective importance. The lack of a reliable and compatible public data base for this type of travel at the European level has held back policy making in a number of areas: development of the *Transeuropean Networks*, airport investment and regulation/deregulation of railroads and airlines.

The paper will concentrate on two areas: examples from current long-distance travel surveys in the member states of the European Union (EU) and a discussion of the results of some recent methodological research on surveys of long-distance travel funded by the European Commission.

A variety of recent national travel surveys (e.g., Austria, France, Sweden, Denmark, and United Kingdom) have included long-distance travel, either in its regular questioning or in specialized extensions to the main survey. The paper will describe the methodologies used and present some key results in terms of amount of travel, modes used and distances covered.

The large differences in the survey methodologies applied and the resulting difficulties in collating the results into a coherent European whole have convinced the European Commission to support research into the development of a uniform survey methodology for long-distance travel, in the first instance. Two major activities were undertaken: a series of surveys undertaken by national statistical agencies and a pair of substantial research projects within the 4th Framework Research Programme. The paper will describe the research undertaken in both streams of work and discuss important substantive and methodological results.

The final section of the paper will present the current state of discussion about a possible EU-wide survey of long-distance travel behavior.

INTRODUCTION

Long-distance trip making (journeys to destinations beyond 100 km from home or current base) has a very small share of all journeys (about 0.5 percent), but represents a much larger share of the total miles traveled (about 20 percent) and therefore of the commercial and environmental impacts of travel. The U.S. figures are rather comparable considering the different spatial structures in the United States [according to the American Travel Survey (ATS) 1995 with a 100-mi minimum distance threshold about 0.5 percent of all trips covering roughly 25 percent of the person-miles traveled].

The ongoing changes in European transport policy have highlighted the need for better

information about long-distance travel and also the lack of adequate publicly owned information about it. While there are various national official and commercial sources their survey techniques are too disparate to allow an easy compilation of these sources. Still, there are a number of policy areas that urgently require uniform data at the European level for their decision making:

- Development of *Transeuropean Networks*: a system of infrastructures intended to support the uniform European market by closing the remaining gaps between the national infrastructures and networks;
- Large-scale infrastructure projects of European importance, here in particular the airport expansions planned at this time (e.g., Heathrow Terminal 5, Schipol expansion, Paris Charles de Gaulle, etc.);
- Monitoring the social and economic cohesion policies of the European Union (EU) and the member states; and
- Monitoring the deregulated transport markets, in particular the airline and the railway industries.

The European Commission has recognized this gap and has supported two initiatives, which are trying to address it by preparing the methodological ground for an European-wide survey of long-distance travel: the research project Methods for European Survey of Travel Behaviour (MEST) and series of pilot surveys coordinated by Eurostat. This paper will present in its first part some of the current results from the ongoing national work, while the second larger part will highlight some of the more important methodological results of the two methodological initiatives. The remainder of this introduction will briefly discuss the specific methodological difficulties of long-distance travel surveys to set the scene.

There are methodological issues, which arise specifically in the design of long-distance surveys, in particular in contrast with surveys of daily mobility. In surveys of daily mobility certain questions do not need to be asked; they seem either self-evident or professionally agreed upon (Richardson et al., 1995). In particular, the study object is clear: to capture all movements of the respondent for a day or more, excluding only movements within large facilities, such as shopping precincts or factories. This basic question is open to discussion in the case of long-distance travel. The division between movements relevant to the long-distance travel and the related decision making and the irrelevant local movements need to be defined, as it is impossible to ask the respondents to report all movements undertaken during a multiday long-distance trip (Axhausen, 1997). The difficulty for the design is to avoid both complex discussions of definitions and to find a natural description, which invites the respondent to report the relevant movements, while minimizing the difference between those reported and those that should have been reported.

At the core of the design problem is the very limitation of long-distance travel surveys to journeys with a minimum distance, e.g., 100 km. The movements become rare events—in the case of the 1995 ATS, four journeys per year over 100 mi [Bureau of Transportation Statistics (BTS), 1997]—requiring long reporting periods to increase the chances that the respondent can report at least one journey and that the contact is not wasted in terms of capturing no information about long-distance travel. Counterbalancing this is the problem of recalling events that might have happened weeks ago, in some detail, which limits the reporting period to a range of 4 to 8 (12) weeks, given the relatively low salience of routine

long-distance travel for many above average frequency travelers (Armoogum and Madre, 1997).

The duration of the reporting period interacts as a design variable with the basic unit of reporting chosen: stage, trips or journey (*1*). 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 4-week reporting period might be compatible with stages, while a 12-week reporting period only with journeys.

In postal questionnaires this issue is compounded by the issue of the reference trip or journey to provide for in the design: 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 freedom in the choice of the basic unit and its definition. Related to this is the question of how frequent travelers 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, while in the second one would like to offer shortcuts to avoid the tedium of repeating the description of very similar movements. While both things can be achieved in Computer-Assisted Telephone Interviewing (CATI)/Computer-Assisted Personal Interviewing (CAPI) contexts, they are not possible on paper forms without inviting other respondents to use them as shortcuts. In addition, one is interested in the details of those journeys by frequent travelers, if one has doubts about the identity of those repeated journeys.

The same questions reoccur in the design of the set of questions about each reported unit: number, complexity of the items, complexity of the available precoded 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 crowding the page, an impression one should.

Given the long reporting periods usually employed in surveys of long-distance travel, the attraction is large to make the survey prospective, i.e., to inform the respondent before the reporting period starts that she or he will be asked to participate in a survey of that period. In surveys of daily mobility it is usual to send the survey form the day before and ask the respondents to fill out the form the following evening. This has been shown to increase substantially the number of movements reported and the related detail in comparison with retrospective surveys, where the respondent is approached cold and asked to recall the last day or a period ending with the day of contact. In long-distance surveys the long reporting period reduces this advantage by asking the respondent to commit in advance an unknown amount of time to the survey task, as the respondent cannot exactly know in advance how much he will travel. This uncertainty in comparison with a retrospective survey, for which the respondent can assess the workload, seems to reduce the response rate (Axhausen et al., 1997). Still, those who do respond provide more and better quality data.

In European surveys the survey protocol has to allow for the very different attitudes and experiences with surveys across Europe, as well as for the different legal requirement and limitations. It is clear from the experiences so far that a protocol based on a single method of approach and retrieval, be it postal, telephone, or personal, will not work. The protocol will have to mix these methods to combine their strengths in reaching and motivating different group of potential respondents. Telephone can be used to motivate respondents, as well as to retrieve information from people unable or unwilling to read and write or who

have language difficulties. Written forms allow the respondents to work at their leisure and to reach persons who do not answer the phone or are seldom at home. Personal interviews can be used for those requiring the personal presence of an interviewer to support them in the task.

The disadvantage of a mix of methods is that it is unclear, whether the data can be combined without special treatment, as it is currently done. The experiences with combining stated preference and revealed preference data in the modelling of choices has shown that at least their different variance structures have to be incorporated into the model formulation (Hensher and Bradley, 1993). Equivalent models have not yet been developed for travel data resulting from multimethod travel diary surveys. This brief discussion has highlighted some of the special difficulties inherent in conducting surveys of long-distance

Survey object	Detailed definition of the survey object.
Approach	Basic reporting unit (stage, trip, journey).
Minimum distance	Amount and style of measurement (crow-fly, network distance).
Minimum duration	If any, amount.
Destination area	Definition of what constitutes a location for the survey.
Reference location	Definition of location, from which a journey/tour can start.
Geographic range of exclusion	If any, areas within which movements need not be reported.
Temporal range of exclusion	If any, definition of time, during which movements need not be reported.
Other exclusion	If any, definitions.
Details	Number of items, item selection, number of precoded options treatment of other and "don't know" (all levels: movement, person, household, vehicle); wording.
Temporal direction	Orientation of the survey (prospective, retrospective).
Reporting period	Duration of the reporting period.
Protocol	Sequence of planned contacts with the sample/respondents, their tasks, form, associated materials, and their temporal spacing (pre-contacts, screening, announcements, reminders, survey distribution, other contacts, data retrieval method); logic of sequence; rules of scheduling; treatment of non-respondents.
Incentives	If any, form and value.
Design of postal instruments	Number of separate forms, distribution of items between forms, sequence of items, design of form (size, orientation and structure of page, fonts, color, density of items, illustrations, etc.), explanatory materials, if any (amount, presentation, style, etc.), support materials for the respondent.
Design of oral instruments	Sequencing of questions, provision of support materials (address list, timetables, maps, etc.), reuse of prior information.

travel, where the complexity of the subject, the resulting response burdens and the data needs have to be balanced, so that we obtain valid and useful data at a reasonable cost. The list below summarizes the key dimensions to be decided upon in the overall design.

EUROPEAN LONG-DISTANCE TRAVEL: SOME RESULTS

The ATS and the National Personal Transportation Survey (NPTS) provide the United States with consistent information about the travel behavior of its residents. No comparable data sources exist for the EU. While most member states of the EU have conducted recent national surveys of daily mobility behavior (the exceptions are Spain, Italy, Greece, Ireland, and Luxembourg), dedicated recent national surveys of long-distance travel are the exception (Austria, France, Denmark, Sweden, and, last year, Portugal, not yet reported). This section will provide some insights into European long-distance travel based on the available reports from these countries (Axhausen, 1998; Armoogum and Madre, 1997; Axhausen and Youssefzadeh, 1998; Herry and Sammer, 1999; and MEST-Consortium, 1998).

Survey Approaches

The five countries considered here show marked differences in the approaches used, which makes their results difficult to compare. Reporting periods range from 2 weeks to 3 months; minimum distances from 50 to 100 km and the contact might be purely postal or CAPI face-to-face. The reporting is equally uneven and this summary draws heavily on unpublished conference and workshop presentations.

The postal 1995 Austrian National Travel Survey consisted of a survey of daily mobility and of a survey of long-distance mobility (all trips to destinations beyond 50 km from home for last 2 weeks). This long-distance survey was only distributed to about a third of all households sampled, with a significantly lower response for this element. The survey was retrospective and followed in its graphic design the general KONTIV-philosophy of a multicolumn layout.

The France 1994 NPTS included a range of different instruments: a survey of daily mobility, a logbook for car use, and two long-distance trip-based surveys (minimum distance 80 km)—one with a 3-month retrospective reporting period (face-to-face interview) and one with a 3-month prospective reporting period using a self-administered diary. The survey was partially self-administered, partially face-to-face.

The Swedish National Travel Survey is a continuous CATI survey, which has been administered by Statistics Sweden for last 5 years. It includes a survey of daily mobility and a survey of long-distance travel (all trips over 100 km for 1 month and all trips over 300 km for 3 months).

The 1996-1997 Danish long-distance survey was an add-on to the ongoing continuous CATI survey of daily mobility, which is conducted by Statistics Denmark. The minimum distance was 100 km and the reporting period was the last month and the survey was based on the concept of the stage.

A large-scale Portuguese Eurostat pilot was conducted in 1996 as a face-to-face CAPI survey in the north of Portugal by the national statistical institute. The minimum distance was 100 km and the reporting period were the 3 months of April, May, and June.

Main Results

The purpose of this section is to highlight the large amount of diversity in the experiences of the European with regards to long-distance travel. Table 1 summarizes some central results about the frequency of travel, purposes, main modes, and distances covered. For the Portuguese from the relatively poor north of the country long-distance travel is a rare event (0.02 journeys/person per week and about 1 journey/person per year), while the Northern European residents have between 0.10 and 0.14 journeys/week (between 5 to 7 journeys/year). Overall business travel, visits to friends and relatives, and holidays have roughly equal shares highlighting the importance of private long-distance travel (only France has a significantly lower share of business travel). The car dominates the modal choice in all surveys, but the shares of coach, train, and air depend on the distances traveled and the availability of alternatives. The train is more prominent in Austria and France and air travel is more dominant in Denmark and Sweden, while in relatively poor north Portugal the coach is the most prominent public transport mode. The distances traveled reflect the normal distance decay overlaid with the effect of holiday travel by air to destinations at the European coasts and even further.

While the results show some comparability, the very different underlying methodologies give reason to worry about the appropriateness to merge the results. In addition the different base units of the enquiry (stage or trips) would make such a merger only possible at the level of the journey implying a substantial loss of detail.

SOME METHODOLOGICAL RESULTS

The methodological work undertaken in recent years in Europe has focused on three main questions:

- Is it possible to define a protocol and design, which is acceptable across all member states of the EU?
- What impacts have the different elements of the protocol and design on the response behavior of the respondents?
- What impacts have the different elements of the protocol and design on the data yield, i.e., the number of journeys and their parts?

This section will discuss the results for the last two questions, while the first will be discussed in the concluding section.

Protocol, Design, and Response Interaction

It is well known that the protocol and the design of a survey influence the response rate by selectively excluding certain types of possible respondents completely or by selectively encouraging them to participate. While the research on surveys of daily mobility has been ongoing for many years, equivalent research on long-distance surveys has been rare. The increased complexity of long-distance surveys makes this actually surprising. Relatively more research is needed here to reach firm conclusions.

A total of 45 different surveys in terms of their protocol and design were conducted

within the MEST project and the Eurostat pilot initiative. They ranged from mostly small-scale pilots (samples between 150 to 400 sample units) to large national exercises. Table 2 provides an overview of studies undertaken during the Eurostat pilot (Axhausen, 1998). The response behavior varies considerably between the surveys. The best response rates are achieved by mixed-mode surveys, involving both paper and telephone elements, and CATI only surveys, but those are conducted in Scandinavia only, which makes it difficult to generalize these experiences.

The various pilot surveys of the MEST project allow a more detailed analysis of the interactions between the different elements of the design and the protocol and the resulting response rate due to the consistency between the surveys save for those dimensions varied by the experiments. The dependent variable, response rate, will be analyzed in three forms:

- Postal response (where relevant): Share of sampled persons replying in writing.
- Telephone response: Share of sample answering on the phone, who had not answered in writing (full interview in case of CATI surveys; non-response interviews in case of postal surveys).
- Total response: sum of the above.

TABLE 1 Some Results from Recent European Surveys of Long-Distance Travel

Characteristic	Survey				
	Austrian NTS	**Danish LDS**	**French NPTS**	**Portuguese Pilot**	**Swedish RVU**
Minimum distance	50 km	100 km	100 km	100 km	100 km
Journeys/person per week	0.5	0.14	0.10	0.02	0.14
Purpose (%)					
Work	36	40	15	37	24
Visit to friends, etc.	28	33	37	19	—
Holidays	16	17	35	44	—
Leisure	—	—	—	—	68
Other personal	20	10	14	n/a	8
Main mode used (%)					
Car	78	62	79	69	69
Train	17	14	11	5	9
Air	–	12	5	5	10
Other public transport	4	—	—	14	7
Other	2	12	5	7	5
Distance (%)					
100-200 km	—	—	—	41	16
200-400 km	—	—	41[a]	44[b]	18
400-800 km	—	—	27	10[b]	22
800+ km	—	—	32	5	44

[a] Private journeys only; 100-200 km not shown and included in 200-400 km

[b] Estimate

The three waves of the MEST pilot addressed the following issues:

Wave I

- Effect of temporal orientation (prospective/retrospective).
- Effect of data collection method (self-completion/telephone).
- Effect of respondent workload ("small" question set/"large" question set).
- Cultural effects (Sweden/Portugal).

Wave II

- Effects of trip versus stage reporting.
- Effect of page versus column presentation of survey.
- Effect of survey duration (4 and 8 weeks).
- Cultural effects (United Kingdom, Portugal, France, and Sweden).

Wave III

- Acceptability of a draft benchmark survey.
- Test of specific minor alterations to the draft benchmark survey (United Kingdom and Sweden).
- Effect of temporal orientation (prospective/retrospective) (France).
- Cultural effects (United Kingdom, Portugal, France, and Sweden).

The protocols of the postal-based surveys always involved a telephone interview of the persons not replying in writing. These non-respondent interviews covered the same elements as the written instrument, if in slightly less detail.

The results of the linear regressions for the three dependent variables are presented below for all surveys. The analysis across all surveys with the smallest variable set has consistently the best fit, as measured by the adjusted R^2. The estimated equations are jointly highly significant as measured by the *F*-values. Weighting the regression with the size of the sample did not change the conclusions drawn from the results. They are therefore not reported here.

In the analysis of the postal returns the dummy "postal survey" acts as an additional constant. Willingness to participate in writing is equally low in Sweden, France, and Portugal. Only in the United Kingdom there is a significantly higher willingness to respond in writing, even having adjusted for the fact, that in the second wave the U.K. sample was drawn from a panel of respondents to a prior survey. The new contractor in the third wave in Sweden had a substantial positive impact. The MEST pilots profited here from the trust in the official Statistical Central Bureau (SCB), which acted as our contractor in the third wave.

The sample screening in the first wave in Portugal had also a significant positive effect, but gives rise to a general worry about sample self-selection.

The changes between the first and the second and the first and the third wave in both design and protocol had no recognizable impact on the written response, in spite of the seriousness of the changes. Roughly a quarter of the sample was willing to participate in

each country all other things being equal. It seems there is a core of willing respondents, who participate out of interest, civic sense, boredom, or curiosity.

The telephone returns, either CATI only returns or returns to the complex non-response interviews is higher overall (43 percent on average across all surveys and countries). A written element sent prior to the telephone contact has a significant negative impact on the willingness to respond on the phone, which was significantly further reduced by the more complex design of the second MEST wave and even more so by the protocol of the third wave (see below for further discussion).

In addition, there are substantial country effects. While the Portuguese are rather willing to respond on the phone, this willingness is significantly lower in France and even more so in Sweden and the United Kingdom, although the official SCB was able to more then equalize this effect in Sweden in the third wave.

The overall response, averaged over all surveys, was satisfactory with 66 percent. The massive negative effect of the third wave protocol, resulting from the reduction in the telephone responsiveness, is a disappointment, but further analysis might shed a light on the reasons for this effect.

The results here have to be treated with care given their aggregate base and the small sample of surveys. Still certain main points are clear.

Country Effects

The effects of the countries are strong and generally significant. Ceteris paribus respondents in Portugal, Sweden, and France are less willing to reply in writing, while the French, Swedish, and English are less willing to reply on the phone. The first result could be the effect of an orally oriented culture in Portugal, which does not prioritize the written world. Still, the unwillingness to reply on the phone in the northern European countries could be the result of the substantial amount of telephone interviewing and sales, which reduces the general willingness to participate. Portugal might therefore catch up as time goes by.

Sampling Effects

The self-selection effect, either through the use of panel or screening, is visible only for the written reply, as one would expect. Its use can only be recommended, if it does not bias the sample drawn. Further research is required here.

Contractor Effects

The country effects above confound country and contractor effects for France, Portugal, and the United Kingdom. The interpretation above is based on the assumption that the firm employed are representative for private sector survey firms in these countries and there is no reason not to believe this.

The case of Sweden is instructive, as we contrast a private-sector firm with the official statistical office acting as a consultant. Given the scale and skill differentials between a small private firm and the national statistical office it is difficult to judge what

TABLE 2 Overview of Eurostate Pilot Surveys

Country	Survey	Approach	Temporal orientation	Duration	Min. distance	Survey period	Contact	Non-response interviews	Sample unit	Reporting by	Sample size	Sampling area	Sample type
Austria	Fessel+ IFES	Trip-based	Retrospective	Last 14 days	50 km	Fall 95	Written/pers collection	None	HH	All over 6	6036	Nationwide	Stratified by Bezirk
	Sammer +Herry	Trip-based	Retrospective	Last 14 days	50 km	Fall 95	Mail-back	Telephone	HH	All over 6	1200	Nationwide	Stratified by Bezirk
	Pilots	Stage-based	Retro- and prospective	Last 4 and 8 weeks	75 km	Spring 96	Mail-back	Telephone	HH	Traveler	1080	Local	Random
Denmark		Stage-based (drilling down)	Retrospective	Last month	100 km	11/96-10/97	CATI	None	Person	Persons (16-74 yrs)	21600	Nationwide	Random
France	CATI	Trip-based	Retrospective	1 or 3 mo.	100 km	1/97-3/97	CATI	None	HH	1 person (6+ yrs)	500	Regional	2 methods
	Postal	Trip-based	Retrospective	1 or 3 mo.	100 km	1/97-3/97	Mail-back	None	HH	1 person (6+ yrs)	500	Regional	2 methods
Germany	CATI	Trip-based	Retrospective	2 mo.	100 km	Spring 1996	CATI	None	Person	1 person (14+ yrs)	130	Nationwide	Random
	CATI/ postal	Trip-based	Retrospective	2 mo.	100 km	Spring 1996	Mail-back/ CATI	None	Person	1 person (14+ yrs)	130	Nationwide	Random
	Postal	Trip-based	Retrospective	2 mo.	100 km	Spring 1996	Mail-back	None	HH	all members	250	Nationwide	Random
Italy	CATI	Stage-based	Retrospective	Last month	100 km	Spring 1997	CATI	Telephone	Person	Persons (18+ yrs)	7000	Nationwide	Stratified by region
	Postal	Trip-based	Retrospective	Last month	100 km	Spring 1997	Mail-back	Telephone	Person	Persons (18+ yrs)	1000	Local	Random
Portugal		Trip-based	Retrospective	4/96-6/96	100 km	7/96-8/96	CAPI	None	HH	all over 15 yrs	5694	Regional	Stratified by fregusia
Spain		Trip-based	Retrospective	2 months	100 km	1/97-4/97	Postal/ CAPI retrieval	None	HH	all over 15 yrs	1500	Regional	Stratified by district
Sweden		Trip-based (drilling down)	Retrospective	1 to 3 mos	100/300 km	1996	CATI	None	Person	Persons (6-84)	9882	Nationwide	Random

TABLE 3 Aggregate Analysis of the Response Rates: All Surveys

VARIABLE	DEPENDENT VARIABLE					
	Postal returns		Telephone returns		Complete returns	
	Parameter	Significance	Parameter	Significance	Parameter	Significance
Constant	-9.27	7.4	**	67.2	**	
Postal survey	25.2	**	-12.0	**	12.0	**
France	4.9	-8.3	**	-2.0		
Sweden	7.3	-19.7	**	-11.7	**	
United Kingdom	16.6	**	-22.9	**	-5.7	
Sweden: 3rd Wave contractor	30.4	**	29.9	**	60.5	**
Recruitment from set of prior respondents	9.1	-5.2	4.8			
Screening of sample by telephone interview	10.2	*	-8.4	*	1.8	
Summer holidays during survey period	-4.5	-0.9	-5.5			
MEST 2nd Wave	0.4	-7.6	*	-6.9		
MEST 3rd Wave	-1.4	-30.6	**	-31.1	**	
F	20.6	**	31.7	**	25.6	
Adj. R2	0.89	0.93	0.91			
N	26	26	26			

** Alpha = 0.10; alpha = 0.05

has contributed to the substantially larger response obtained by SCB. More empirical work is required to establish the advantage of using an official survey firm in comparison with a private survey firm of similar size.

Design Effects

With the exception of the negative effect of prospective surveys on postal response, no consistent design effects could be identified at this level. The designer seems therefore relatively free in his or her choices.

The design effects are more visible and stronger in the telephone response. The respondents, who did not participate in the written element of the survey, are less willing to participate in the telephone element of a task, they must have perceived as difficult. The significant negative effect of a complexity variable tested before and the negative, but not significant, Wave II design variable indicates this.

This conclusion is supported by the negative and significant effect of the Wave II dummy in the analysis of the telephone response across all surveys. This result is disappointing, but may be not surprising. The design changes reflected the responses of the participants in the cognitive laboratories undertaken by MEST (Axhausen and Youssefzadeh, 1998), which are likely to belong to the group of respondents willing to participate in the written element—no effect there of the design changes. Still, the design seems to have been perceived as complex by a significant share of the respondent and lowered their willingness to participate at least in the telephone element of the survey.

The disappointment of Wave III is based on the unsatisfactory participation in the telephone element. The written element was no worse then before.

The contrast between the results obtained by the private firms in Portugal, France, and the United Kingdom and the results obtained by Statistics Sweden raise the issue to what extent the protocol of the third wave is at fault or the performance of the firms. Given the general competence of the firms employed, it is unlikely that the main fault can lie with the firms. In addition, there were special circumstances in each case. In France the survey was conducted by a firm also heavily involved in electoral opinion polling and it was felt by the firm that their name reduced the willingness of the sampled to participate. In the United Kingdom, the survey was conducted in a economically polarized area with both the poorer and the richer respondents overrepresented; both groups are known to be less willing to participate then middle-class respondents. In Portugal the negative effect of the written material, perceived by many as complex, was felt strongly. In Sweden SCB might have been particularly keen in this, their first involvement in MEST, whereas the other firms might have been professional, but not keen in their second or third wave of the project.

These aggregate results across all surveys are consistent with the more detailed disaggregate modelling of the Austrian Eurostat pilots, which could only identify one design effect on response behavior: a positive effect of retrospective surveys on response (Axhausen et al., 1997).

Data Yield and Survey Design

The second important interaction in the design of a survey is the data yield, the number of journeys reported by each respondent, and the detail with which it is reported, e.g., the

number of stages or the number of trips for each journey. The analysis of the Austrian Eurostat pilots had revealed a negative impact of retrospective surveys on the number of journeys reported and a positive one of complexity (Axhausen et al., 1997). This analysis will be extended here by looking at the written responses of the MEST second wave pilots and the results of the French Eurostat pilots.

The analysis of the MEST data-based on the written returns of Wave II and Wave III reported below looks at two facets of the data:

- The number of journeys reported modeled by the negative binomial model, including no journeys reported (Table 4); and
- The presence of at least one reported journey modelled by the probit model (Table 5) for each of the waves and both data sets combined.

Other than the obviously significant survey duration, only one of the design variable is significant for the model of the number of reported journeys: the Wave II indicator. This is difficult to interpret, as it is also a dummy for the seasonal difference between Waves II and III (winter versus early summer). In the models of reporting at least one trip, the page-based design shows a significant positive effect. The retrospective protocols dominated in this sample, so it is difficult to interpret the lack of significance of this variable.

Among the sociodemographic variables, the variables describing both income levels and overall level of mobility dominate as a group. Interesting, although obvious on reflection, is the usefulness of the number of telecommunication links as an indicator of mobility. Models were estimated with the four contributing numbers separately (phones, mobiles, fax lines, Internet connections), but they had significantly less explanatory power. Their joint effect is the relevant one.

The results reported here do not explicitly confirm the earlier results, but not contradict them either. Essential is the relative lack of strong design-based impacts on data yield. Page-based forms have a positive effect, but only in the restricted sense of encouraging people to report at least one journey.

The French Eurostat-pilots provide a second opportunity to analyze in detail the interaction between survey design and protocol and the data yield (Axhausen, 1998). The main experimental design variables employed in this study were:

- Sample selection: The person responding was either selected at random from the telephone book or selected at random from a list of persons who had indicated earlier to the survey firm that they would be willing to participate in future research.
- Contact approach: The person was either contacted by phone or by mail.
- Style of reporting: The persons were asked to report either 3 months retrospectively or three times 1 month retrospectively.

The telephone survey was conducted as a CATI survey, while the postal survey employed a newly designed form. The study area was the Rhone-Alpes region in France.

TABLE 4 Data Yield: Number of Journeys Reported (Written Replies)

Variable	Wave					
	Wave II		WaveIII		Waves II and III	
Negative Binomial Model	**Parameter**	**Significance**	**Parameter**	**Significance**	**Parameter**	**Significance**
Constant	–2.25	**	–3.13	**	–4.32	**
Wave 2	—	1.36	**			
Page- versus column-based	0.26	–0.24				
Trip- vs. stage-based	0.33	–0.04				
Survey duration (weeks)	0.27	**	–0.19	**		
Prospective versus retrospective	–0.22	0.28				
United Kingdom	0.51	1.98	**	1.21	**	
France	0.07	1.96	**	1.06	**	
Portugal	–0.43	–0.19	–0.43			
Number of telecom-munication links[1]	0.03	0.24	**	0.16	**	
Second home owners	–0.12	0.38	**	0.38		
Number of household vehicles	0.30	**	0.10	0.21	**	
Female	–0.10	–0.26	–0.15			
Born 1920s	–0.45	0.11	–0.02			
Born 1930s	–0.19	0.24	0.18			
Born 1940s	–0.36	0.25	0.09			
Born 1950s	0.02	–0.19	0.06			
Born 1960s	–0.13	0.16	0.09			

continued on next page

TABLE 4 (*continued*) Data Yield: Number of Journeys Reported (Written Replies)

Married	0.12	–0.08	0.10			
Disabled	–0.39	–0.77	–0.47			
University degree	0.40	0.26	0.34	**		
In education	0.66	**	0.69	**	0.72	**
Part-time working	–1.69	0.84	**	0.54		
Full-time working	0.04	0.67	**	0.63	**	
Self-employed	–0.32	–0.08	–0.28			
Frequent flyer card	–0.09	0.16	–0.34			
Rail discount card	0.15	0.40	0.10			
Both cards	—	0.43	—			
Overdispers-ion parameter [I]	1.09	**	0.91	**	1.11	**
<(0)	–563.9	–622.2	–1202.3			
<(Poisson)	–494.7	–483.3	–1010.9			
<(Negative binomial)	–417.1	–443.3	–876.4			
Y2	0.26	0.29	0.27			
N	281	407	688			

* = significant at I = 0.10; ** significant at I = 0.05

[1] Number of phones, mobile phones, fax lines, and Internet connections in the household.

TABLE 5 Data Yield: Report of at Least One Journey (Written Replies)

Variable	Wave					
	Wave II		WaveIII		Waves II and III	
Probit Model	Parameter	Significance	Parameter	Significance	Parameter	Significance
Constant	–2.30	**	–2.58	**	–3.70	**
Wave 2	—	1.37	**			
Page- vs. column-based	0.48	**	–0.47	**		
Trip- versus stage-based	0.08	—	0.11			
Survey duration (weeks)	0.18	**	–0.15	**		
Prospective versus retrospective	–0.31	0.11				
United Kingdom	0.66	**	1.29	**	0.95	**
France	0.54	*	1.62	**	1.11	**
Portugal	–0.02	–0.11	–0.09			
Number of telecommuni-cation links[1]	0.14	*	0.21	**	0.16	**
Second home owners	0.22	0.38	**	0.35	**	
Number of household vehicles	0.28	**	0.12	0.21	**	
University degree	0.92	**	0.44	**	0.49	
Part-time working	–0.48	0.51	*	0.39		
Full-time working	0.40	0.41	**	0.48	**	
<(0)	–194.3	–270.4	–472.2			
<(β)	-173.0	-197.3	-376.1			
Y2	0.11	0.27	0.21			
N	281	407	688			

* = significant at I = 0.10; ** = significant at I = 0.05

[1] Number of phones, mobile phones, fax lines, and Internet connections in household.

The data was analyzed using the weights provided by the research firm employing again the negative binomial regression to model the number of journeys reported by the respondents presents the results of the regression (*1*).

The results confirm, but also expand the earlier results. Only one of the three experimental design variables has a significant main effect (recruitment), but there are also significant interaction effects, which give new insights. The recruitment from the panel improved response rates, but does significantly reduce the data yield as a main effect, but in certain combination with a 3-month reporting period this effect is equalized. The other positive interaction occurs in the combination of a fresh random sample and a monthly reporting period, at least for those continuing in the study, as the response was lower for those recruited for three monthly reports. The other significant interaction is between reporting frequency and style of contact. Here combinations of either a 3-month postal report or a monthly telephone report improved data yield, while a 3-month telephone report or monthly written reports reduced it.

The results show again the importance of the style of recruitment, but also of the style of contact. The significant interactions indicate that certain combinations might be successful, but that others are not in spite of positive main effects of the individual design variables.

The methodological results highlight the complexity of the choices to be made, especially the French results. Still, there is a clear preference emerging for the following choices:

- *Retrospective* surveys, as in the context of the long-distance travel the respondents prefer the known workload involved in a retrospective survey in comparison to the open commitment in a true prospective survey. Interviews with prospective respondents in the third wave of MEST showed, that these have nearly exclusively recorded their trips after the end of reporting period. The observed gains in data yield must therefore be due to sample selection and the increased commitment of such prospective respondents.
- *Complexity* in terms of size and detail of the item set needs to reach a certain level to give a postal survey credibility with the literate part of the sample. At this point the gains from this group outweigh the losses from the functionally illiterate groups, for which a complex written survey is a stumbling block, reducing their willingness to participate later on the phone. The optimal level of complexity, which satisfies the literate, while minimizing the irritation of the functionally illiterate is not yet known.
- Within the limitations of the various MEST pilots the *physical design* of the forms had little impact on the response behavior or the data yield. There is an obvious preference for generous page layouts and clear guidance through the form.
- A *trip-based* approach to the reporting of the journey details. But the respondents seem willing to report stage information, when the question is formulated as one about the route taken (Which route did you take? Report the main roads and junctions; stations or airports, especially those where you changed train or plane).

The strong country and contractor effects raise important questions for a European survey of long-distance travel. Should the national statistical offices, as still the most credible research firms, be asked to undertake such a survey, in spite of the difficulties to harmonize the survey

protocol and design between independent institutions with such strong methodological preferences of their own, or should independent market research firms be asked, promising a greater control over the protocol, but potentially fewer respondents.

There is a clear need for further methodological research to address remaining methodological issues, such as the optimal choice of the complexity of the survey object and of the question set, the trade-off between the duration of the survey period and the minimum distance, the division of tasks between a roster of journeys with few details and the detailed journey description form,

The methodological work performed in MEST and in the Eurostat Pilot surveys in addition to the other 4th Framework projects defining a European Transport Information System have highlighted the need for a publicly funded large-scale European survey of long-distance travel. Without such a survey European researchers can neither the input data for the modeling and monitoring nor gain the practical experience required for the further development of such a tool. The survey would also provide the framework for the further technological development of travel surveys (TEST, 1999).

This large-scale survey cannot and should not replace the equally large-scale national surveys. It is intended both as a benchmark against which the national efforts could be scaled and compared across and as an independent source of information at the European level (in contrast to the detailed national-level surveys).

While there is a clear professional consensus about the necessity for such a survey, the politics and funding of such a survey is very problematic. The European Commission has, as a rule, not commissioned its own survey work, but has satisfied its needs through the aggregation of national surveys, which were mandated through legislation. These directives leave the national statistical offices wide scope in the implementation of mandated surveys. This is acceptable in many areas, but this framework is too loose for travel and tourism surveys. The initial experiences with the new directive on tourism statistics and travel are not promising in the view of the author. No directive exists for statistics on personal travel. While some member states have their own, often well established, surveys in the field a fair number do not. For general political reasons at this time a directive enforcing the harmonisation of the existing travel surveys is not a possibility. The Commission has to find a budget to break with the tradition of not commissioning surveys on its own to fill this glaring data gap in its new policy area: transport and transeuropean networks.

There is currently hope that the Commission will use the opportunity of the 5th Framework Research programme starting in 1999–2000 to fund such a survey, in spite of the difficulties inherent in using this programme for such purposes. Still, nothing is known until the first call for proposals in March. The main difficulties is that the budget allocated is likely, based on past experiences, going to be too small to fund the continuous multiyear effort, which is required. It might provide for the first year and most of the set-up costs. This implies that any consortium wanting to undertake this work will need further external funding from governments, firms, administrations, etc. The rules of the research programme create difficulties for finding this money by providing only 3 months between the call for proposals and the submission of the tender. In the past, they created further difficulties by asking survey firms to find cofinancing for the survey work exaggerating the problems of finding the required budgets. There have been informal indications of the removal of this requirement. In spite of these difficulties, a possible call for proposals would be a great opportunity for research and, in turn, policy analysis.

TABLE 6 Data Yield: Number of Journeys Reported (French Eurostat Pilots)

Variable	All reporting	
Negative Binomial Model	**Significance**	**Parameter**
Constant	–0.22	*
Only reported in first month	–0.75	*
Only reported in first 2 months	–0.07	*
Postal contact	–0.01	
Recruited from panel of prior respondents	–0.11	*
Asked to report monthly	–0.02	
Interaction Term Recruitment * Contact	0.06	
Interaction Term Recruitment * Reporting Frequency	0.13	*
Interation Term Reporting Frequency * Contact[1]	0.10	*
Female	–0.07	
Under 21 years	–1.05	*
Over 40 years	0.04	
Full-time employed	0.42	*
Part-time employed	–0.21	*
Student	0.58	*
Head of family	0.23	*
Number of cars in family	0.31	*
Socioeconomic class (++)	0.32	*
Socioeconomic class (+)	0.32	*
Socioeconomic class (-)	–0.02	
Classification of municipality (rural)	–0.23	
Classification of municipality (semi-rural)	–0.50	*
Classification of municipality (small town)	–0.03	
Overdispersion parameter I	1.14	*
<(0)	–161087	
<(Poisson)	–141994	
<(Negative binomial)	–112610	
Y^2	0.21	
N	62665	

* significant at I = 0.10

[1] Number of phones, mobile phones, fax lines, and Internet connections in household.

The MEST consortium (MEST, 1999) has made some suggestions, how such a survey might be structured acknowledging that its own work and the available professional knowledge do not rule out a CATI/CAPI only survey as an alternative. Summarizing the consortium recommends the following protocol:

- Person sample (direct or equivalent two-stage sampling from households) (oversampling of persons with more long-distance trips) (from age 6; proxy reporting until age 14);
- Initial postal contact combined with a series of three reminders consisting of postal reminders and telephone calls, which can be changed into full CATI interviews, as requested by the respondent; no incentives; general follow-up telephone interview with all written responses;
- Rigorous treatment of respondent errors within the follow-up interviews, during the possible CATI interview, fully documented imputation of missing items, fully documented weighting for unit-non-response;
- Complete documentation of the data files in their various phases and of the contact history;
- Face-to-face interviews with a subset of the sample for non-response correction, respectively quality check of earlier responses;
- Continuous administration of the whole survey period (1-year minimum, 3 to 5 years recommended) (MEST, 1999, p. 3.-4) and design;
- Retrospective 8-week survey;
- Minimum 100 km crow-fly distance from current base (home etc.), but 75 km minimum distance stated for the respondents;
- Relevant household, person and vehicle details;
- The most recent journey (independent of reporting period) and all other journeys within the reporting period; and
- Trip-detail including a description of the route for the most recent journeys within the reporting period (up to three) (MEST, 1999, p. 4).

This approach is a compromise, which should give satisfactory response rates in all member states simultaneously, and which provides the information to estimate the total amount of long-distance travel and the details for the required choice modelling. The consortium knows that this protocol will not be the best-of-class in any one country or any one respect, but it provides a starting point for the empirical work required.

ACKNOWLEDGMENTS

The work reported here is based on the results of the EU-funded 4th Framework project "Methods for European Surveys of Travel Behaviour" involving the Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität Innsbruck, 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), TOI (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 and not of the CEC or the Consortium.

The author wishes to thank Mr. Michel Houee (French Ministry of Transport), who kindly made the French Eurostat Pilot data available for this first analysis.

The author wishes to thank all the members of the Consortium, who contributed to the results reported here. All errors are the author's.

NOTES

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).

2. The Netherlands and the United Kingdom have ongoing continuous surveys of daily mobility, which provide some insight into long-distance travel. The ongoing Belgian and Finnish national surveys of daily mobility will do the same. These results will not be reported here.

REFERENCES

Armoogum, J., and J. L. Madre. Accuracy of Data and Memory Effects in Home-Based Surveys of Travel Behaviour. Presented at the 76th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 1997.

Axhausen, K. W. Towards an European Benchmark Survey of Long-Distance Travel. Presented at the Meeting of the Concerted Action Committee, Brussels, Belgium, Dec. 1997.

Axhausen, K. W. The EUROSTAT Pilots of Long-Distance Travel Diaries: Summary of Final Reports. Report to the Österreichisches Statistisches Zentralamt, Wien and EUROSTAT, Luxemburg, Innsbruck, 1998.

Axhausen, K. W., H. Köll, M. Bader, and M. Herry. Workload, Response Rate and Data Yyield: Experiments with Long-Distance Diaries. In *Transportation Research Record 1593*, 1997, pp. 29–40.

Axhausen, K. W., and M. Youssefzadeh. Towards a New European Long-Distance Travel

Survey: Analysis of the MEST Pilots. Report to the CEC, DG VII, *MEST Deliverable*, D3&4, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens Innsbruck, Innsbruck, 1998.

1995 American Travel Survey: Profile. Bureau of Transportation Statistics, U.S. DOT, Washington, D.C., 1997.

Hensher, D., and M. Bradley. Using Steted Response Data to Enrich Revealed Preference Discrete Choice Models. *Marketing Letters*, Vol. 4, No. 2, pp. 139–152, 1993.

Herry, M., and G. Sammer. Österreichischer Bundesverkehrswegeplan: A3-H/2. Report to the Austrian Ministry of Science and Transport, Wien and Graz, 1999.

MEST-Consortium. Proceedings of the 2nd MEST Workshop, Working Paper, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens Innsbruck, Innsbruck, 1998.

MEST Consortium. Towards an European Survey of Travel Behaviour. Report to the CEC, DG VII, *MEST Deliverable*, D5, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens Innsbruck, Innsbruck, 1999.

Richardson, A. J., E. S. Ampt, and A. H. Meyburg. *Survey Methods for Transport Planning.* Eucalyptus Press, Melbourne, Australia, 1995.

TEST-Consortium. Final Recommendations, Report to the CEC, DG VII, *TEST Deliverable*, D8, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens Innsbruck, Innsbruck, 1998.