Weighting or Imputations? The Example of Nonresponses for Daily Trips in the French NPTS

JIMMY ARMOOGUM
Institut National de la Statistique et des Etudes Economiques
Institut National de Recherche sur les Transports et Leur Sécurité

JEAN-LOUP MADRE
Institut National de Recherche sur les Transports et Leur Sécurité

ABSTRACT

This paper reports on methods used to correct nonresponse for daily mobility in the French National Personal Transportation Surveys. A two-stage technique was used for *unit* nonresponse: 1) post-stratification according to the households' characteristics related to response behavior; and 2) correction for sampling error by calibration on margins. Imputation procedures (e.g., deductive, regression-based, hot-deck) were also used to correct *item* nonresponse. These methods maintained the consistent relationships among the main variables describing trips. The paper also addresses how the specific circumstances of this case (e.g., sample drawn from the census, no computer assistance during the interviews) led to the choice of methods.

INTRODUCTION

All sample surveys contain incomplete data, even if great care is taken before and during data collection. Two fundamental types of nonresponse may occur:

Jimmy Armoogum, (INSEE-INRETS), INSEE-UMS, Timbre F410, 18 Bd. A. Pinard, 75675 Paris, Cedex 14, France. Email: jimmy.armoogum@dg75-f410.insee.atlas.fr.

1. *unit* nonresponse, when no information is collected for a household or an individual (e.g., not at home, unable to answer);
2. *item* nonresponse, when most of the questions for a unit are answered, but for some respondents, either no answer is given or the answer is clearly wrong and must be deleted.

Missing data for items can occur when an interviewer fails to ask a question, the respondent is not able or refuses to provide an answer, or the interviewer fails to record correctly the answer provided.

There is no a priori justification for assuming that people who respond have the same characteristics as those who do not. Thus, in computing estimates from the available data collected, we may face biases whose size and direction of error are unknown. In this paper, we show how nonresponse problems were addressed for daily trips in the French National Personal Transportation Survey (Madre and Maffre 1994).

There are two main strategies for handling nonresponse: 1) re-weighting by increasing certain expansion factors, which is commonly used for unit nonresponse; and 2) imputation, replacing the missing item by a value consistent with the respondent sample, which is generally used for item nonresponse. There are also intermediate cases, for instance, weighting for omitted trips. We will discuss advantages and disadvantages of each method.

THE SAMPLE DESIGN AND DATA COLLECTION

From a sample of 20,002 dwellings drawn from the census of 1990 and from the list of new residences built since that date, 20,053 address cards were prepared. The increase in households is due to "burst" lodging (dwellings that have been divided into two or more separate residences since the last census). The sample was spread over eight waves from May 1993 to April 1994 in order to neutralize the seasonal effects, which are important for personal trips. One individual was chosen (the probability of being chosen was equal for everyone in the household) among the eligible individuals (individuals six years and older,[1] present at the time of the survey, and able to answer) of each household. The chosen individual was interviewed face-to-face and asked to describe all trips he or she made the day before and the previous weekend. All motorized households had to complete a car diary, in which they reported all trips made by one of their vehicles, chosen at random, during the span of one week. Generally, the car diary was completed after the interview on daily mobility, which did not allow immediate cross-checking of individual car trips, but only the computation of global statistics from both data sources on the same sample of households. Information collected with those survey instruments is described in a later section.

During each of the eight waves, the surveyor interviewed a given set of households living in the same area. The interviews were spread over the six-week period of the wave, but the day of interview was not assigned a priori. As a result, it was necessary to correct for temporal representativeness (especially for the days of the week) in the weighting procedure.

Although the majority of residences in our first sample were the main residence of a household, this was not always the case: among the 20,053 dwellings visited, 2,666 (13.3%) were out of scope (vacant housing, or second or occasional homes). Among the 17,387 selected households in scope, 3,174 (18.3%) of them refused to respond to the survey.

CORRECTION FOR UNIT NONRESPONSE

For each residence drawn from the 1990 census, there is useful information concerning the probability that a household will respond to the survey. The relationship between the household characteristics and the probability of response is called the response mechanism. We estimated a logit model to describe the response mechanism. Although the household living in a selected dwelling could be different from the one that lived there in 1990, we assumed they were the same, since the survey was conducted only three years after the census.

Nonresponse Correction: Post-Stratification

The main factors explaining unit nonresponse are listed below, from the most important to least important ones.

[1] Unlike the previous survey (1981 to 1982), children under six years old did not describe their mobility.

1. People living in rural areas or in small towns (<20,000 inhabitants) had a lower rate of nonresponse than those living in the conurbation of Paris. We distinguish three classes: 1) rural + small urban areas (<20,000 inhabitants) with a response rate of 86%; 2) medium-size + large urban areas (20,000 to 2 million inhabitants), with a response rate of 81%; and 3) the Paris urban area (10 million inhabitants), with a response rate of 74%.
2. Single persons were less likely to respond than households with many persons. We identified three categories: 1) households of one person, with a response rate of 72%; 2) households composed of two persons, with a response rate of 81%; and 3) households composed of more than two persons, with a response rate of 87%.
3. Motorized households had a higher response rate than those with no car. We identified 3 classes: 1) nonmotorized households, with a response rate of 72%; 2) households with one car, with a response rate of 82%; and 3) multivehicle households, with a response rate of 87%).
4. Households whose head was over 60 years old had a 78% response rate; those with a younger head had an 84% response rate. We chose only two age groups, because under 60 the response rates seem almost constant across age groups.

By cross-classifying these variables, we obtained 54 classes, which form the framework for post-stratification. The response rates ranged from 55% for an individual who is single, living in the Paris conurbation, with no car, and who is over 60 years old (230 people in this class), to 90% for three or more persons living together in rural areas or small towns, with two or more cars, and whose household's head is under 60 years old (2,358 people in this class). We implemented the post-stratification by multiplying the reciprocal of the household's selection probability with the reciprocal of the individual's selection probability and with the reciprocal of the response rate of the individual class:

$$\text{Postratification weight} = \frac{1}{\text{Household's selection probability}} \times \frac{1}{\text{Individual's selection probability}} \times \frac{1}{\text{Response rate of the individual class}}$$

Sampling Error Correction: Calibration on Margins

After reducing the error due to nonresponse by the post-stratification, we found that the margins in the sample differed from those of the largest household survey conducted by INSEE (the French National Institute of Statistics and Economic Studies), an employment survey in which 80,000 households were interviewed in 1993–94. That survey is considered to be a mini-census.[2] We corrected these differences by a calibration on margins. This stage is essential to ensure a representative sample allowing comparison with other data sources (e.g., other INSEE surveys). Calibration on margins is done by iterative proportional fitting, a methodology developed by Deming and Stephan in the early 1940s. We used INSEE-developed software called CALMAR for calibration on margins (Sautory 1993).

Calibration on margins must be based on variables that explain (or are correlated with) transport behavior, and for which the total is accurately known. We took advantage of this stage to compute two temporal variables—"the day of the week" and "the period of the year"—in order to neutralize the temporal effects. Therefore, the variables used to calibrate on margins for the person describing daily trips are the following (see table 1):

- the social category of the individual;
- age and gender;
- the size of the household;
- the zone of residence: three concentric zones (city center, and inner and outer suburbs) for four different urban area sizes;
- the day of the week (one day before the visit of the interviewer) for which daily trips are described (so each day of the week is equally represented); and
- the period of the survey (the year was divided into eight waves).

[2] Obviously, the employment survey is subject to sampling error, but it is also more accurate than the NPTS's sample (with only 14,000 households). The survey methodology was exactly the same in both cases (face-to-face interview), which leads us to conclude that the only source of difference is sampling error.

TABLE 1 Margins in the Sample and in the Population for Persons Interviewed on Daily Mobility

Variable	Margins in the sample after post-stratification (%)	Margins in the population (%)
Social category of the person		
Farmer	1.8	1.6
Craftsman/tradesman	3.5	3.3
Senior executive	6.3	5.6
Intermediary	9.9	9.3
Employees	14.3	13.6
Blue collars	12.9	12.9
Retired/students	17.8	18.1
Unemployed	20.4	22.5
Children (6 to 15 years old)	13.1	13.1
Gender and age		
Males:		
from 6 to 24 years old	13.8	14.6
from 25 to 34 years old	7.7	8.1
from 35 to 49 years old	11.6	11.6
from 50 to 64 years old	8.0	7.9
over 65 years old	6.3	6.4
Females:		
from 6 to 24 years old	13.5	13.9
from 25 to 34 years old	8.8	8.1
from 35 to 49 years old	12.6	11.6
from 50 to 64 years old	8.7	8.2
over 65 years old	9.0	9.6
Number of persons in the household		
1 person	11.9	12.0
2 persons	27.0	26.9
3 persons	19.9	19.7
4 persons	22.2	23.0
5 persons or more	19.0	18.4
Zone of residence		
Rural area living on farm	3.8	3.4
Small urban areas (<50,000 inhabitants)		
Central city	4.6	5.4
Inner suburbs	1.6	1.7
Outer suburbs	6.6	7.2
Medium-size urban areas (50,000 to 300,000 inhabitants)		
Central city	10.1	9.4
Inner suburbs	6.4	6.3
Outer suburbs	16.5	14.3
Large urban areas (> 300,000 inhabitants)		
Central city	10.1	10.2
Inner suburbs	11.9	12.4
Outer suburbs	9.4	10.4
Paris urban area		
City of Paris	3.7	3.9
Inner suburbs	12.2	12.7
Outer suburbs	3.1	2.7
Day		
Monday	20.4	20.0
Tuesday	19.5	20.0
Wednesday	18.0	20.0
Thursday	15.5	20.0
Friday	26.6	20.0
Wave		
1st (from 3 May to 14 June 1993)	12.4	11.6
2nd (from 14 June to 9 Aug. 1993)	12.0	15.4
3rd (from 9 Aug. to 11 Oct. 1993)	12.9	17.3
4th (from 11 Oct. to 15 Nov. 1993)	12.4	9.6
5th (from 15 Nov. 1993 to 3 Jan. 1994)	12.2	13.5
6th (from 3 Jan. to 14 Feb. 1994)	10.4	11.5
7th (from 14 Feb. to 21 March 1994)	10.4	9.6
8th (from 21 March to 30 April 1994)	13.4	11.5

Sources: INSEE-INRETS French NPTS 1993–94 and French Employ Survey 1993–94.

Discussion

Australian data has shown that within small homogeneous population groups the travel behavior of nonrespondents does not differ significantly from the behavior of respondents (Ampt and Polak 1996). Thus, post-stratification according to crossed categories with homogeneous response rates is essential. Unfortunately, the information used for calibrating on margins is slightly different from the sample base. There is no information on newly built dwellings in the census, and no information on car ownership in the employment survey used for calibration. Thus, the second stage changes the margins obtained after post-stratification, and is not satisfactory. Following the methods implemented in Austria (Sammer and Fallast 1996), we are now investigating a single-stage procedure.

For reasons of comparability and efficiency, our daily trips questionnaire was presented in a manner similar to urban survey questionnaires. On the other hand, some of the methods described here might be applied to other types of surveys. This is surely the case for calibration on margins. The size of the conurbation is the best explanatory factor of unit nonresponse, but geographic post-stratification is not sufficient to get a good fit to the sample and an expansion consistent with other data sources. For instance, calibration of age groups could be useful for demographic modeling (Armoogum et al. 1994, 1995). However, as contradictions could appear between the two steps of the procedure we have used, INSEE is now studying a single-step procedure that calibrates on margins according to variables explaining both the nonresponse mechanism and travel behavior.

CORRECTION OF ITEM NONRESPONSE

Correcting for item nonresponse has two objectives:

1. obtaining not only unbiased estimates of averages, but also keeping the distribution of each variable as "natural" as possible; and
2. checking and maintaining the consistency of relationships between the different variables that describe a trip (e.g., origin, destination, distance, time, mean of transport).

Standard Imputation Methods

The main imputation methods for item nonresponse are the following:

1. *Deductive imputation* refers to those cases where a missing value can be obtained through a logical conclusion. The deduction is based on responses given to other items on the questionnaire. A common example in travel diaries is travel distance, which can be checked and calculated from the location of the origin and destination of a trip.
2. *Overall mean imputation* consists of the replacement of all missing values for a given item by the respondent mean for that item. Unless the number of nonresponses is negligible, this procedure may lead to severely understated variance estimates and to invalid confidence intervals.
3. *Class mean imputation* partitions the unit response set into imputation classes such that elements in the same class are considered similar. This classification uses auxiliary variables. There will be some distortion of the "natural" distribution of values, but the bias is less severe than with overall mean imputation.
4. *Hot-deck and cold-deck imputations* replace missing responses with values selected from other respondents in the current survey in the hot-deck method; cold-deck procedures use sources other than the current survey. A number of hot-deck procedures have been proposed, including random overall imputation, random imputation within classes, sequential hot-deck imputation, and hierarchical hot-deck imputation.
5. *Regression imputation* uses respondent data to estimate a regression equation where the variable for which one or more imputations is needed is the dependent variable and other available variables serve as explanatory variables.

Validation and Correction of Daily Trip Data

Trips were described in a weekly stage diary in a previous 1981–82 survey, by interviews on the previous day and the last weekend in 1993–94, and in a weekly car diary for both surveys. The main characteristics of the trips are:

1. *origin and destination*, coded by French municipality and by NUTS3 (regions with about 500,000 inhabitants) for neighboring countries in the last survey;
2. *length*, as estimated by interviewed persons, calculated as the difference on the odometer at the origin and destination in car diaries;
3. *duration*, computed as the difference between arrival and departure times;
4. *transport mode* (up to four different modes in the case of a multimodal trip); and
5. *trip purpose*.

There are obvious relationships among these variables. Some locations are described in the general part of the questionnaire (e.g., the residence and the regular work place). Trip length must be consistent with the distance between the origin and destination (trip length must be greater than crow-flight[3] distance with a margin of 5 km, unless the origin and destination are located in two neighboring municipalities). Door-to-door mean speed (calculated as the ratio of trip length to trip duration) must stay within reasonable limits (see table 2). For car trips, for instance, door-to-door mean speed must fall between 2 km/h and the maximum authorized speed on motorways, which is 130 km/h in France.

Interview on Daily Mobility: 1993–94

Like most surveys, there were almost no item nonresponses on origin and destination locations. Only 10 out of 100,000 trips could not be coded. Thus, we have used crow-flight distances to fill item nonresponses on trip length (1,300 cases) or to replace responses leading to an unreasonable mean speed (400 cases). Generally, the crow-flight distance is multiplied by a circuity coefficient specific to each mode (e.g., 1.3 for private car).

In order to estimate missing or questionable values for duration, we used a regression technique, calibrating the relationship between mean speed and trip distance. For motorcycles and cars, this equation is:

$$\text{SPEED} = 1.4 + 14.6 \log(\text{DIST}+1)$$

For the 1993–94 car diary, where additional information about destinations in "city-centers" was available, four different estimates of this equation were made on correctly described trips:

- if origin and destination were in a city center:

$$\text{SPEED} = \underset{(9.3)}{1.54} + \underset{(185.7)}{15.25} \log(\text{DIST}+1.3) \quad R^2 = 0.474$$

- if origin or destination were in a city center:

$$\text{SPEED} = \underset{(14.9)}{2.46} + \underset{(219.5)}{15.72} \log(\text{DIST}+1.3) \quad R^2 = 0.467$$

- if origin and destination were not in a city center:

$$\text{SPEED} = \underset{(31.0)}{4.39} + \underset{(246.6)}{15.64} \log(\text{DIST}+1.3) \quad R^2 = 0.445$$

- if information was missing for origin or destination:

$$\text{SPEED} = \underset{(4.7)}{1.74} + \underset{(102.0)}{15.90} \log(\text{DIST}+1.3) \quad R^2 = 0.511.$$

Because of congestion, the average speed is lower in denser areas, and increases significantly less with trip distance when origin and destination are both situated in city centers. In 1981–82, the previous form of this question concerned the use of a motorway during the trip. This information did not provide significantly different equations of speed as a function of trip distance.

Because walking trips usually have their origin and destination in the same municipality, crow-flight distance between municipalities cannot be used to compute trip distance. For this mode, we have assumed that the mean speed is 3 km/h, either to estimate trip length (500 cases) or to fill the few missing data on duration.

Using these techniques, we succeeded in getting totally consistent data on locations, distance, duration, mean speed, and mode. There are very few missing values left : 2 on trip distance, 6 on trip duration, plus 11 cases where trip duration was given, but arrival and departure times remain unknown (see table 3).

[3] Defined as: crow-flight distance = $[(X_o - X_d)^2 + (Y_o - Y_d)^2]^{0.5}$; where (X_o, Y_o) are the origin's coordinates and (X_d, Y_d) are the destination's coordinates.

TABLE 2 Controlling Data by Mode

Mode	Speed (in km/h) Minimum	Medium	Maximum	Trip/crow-flight distance[1]
Walking	1	3	10	0.3
Bicycle	1	8	30	1.0
Motorcycles	2	15	[2]130	1.2
Car, truck, taxi	2	24	[3]130	1.3
Bus	2	12	110	1.4
Rail urban transport	3	12	75	1.1
Train	10	54	[4]150	1.2
Aircraft	100	400	1,000	1.1
Seacraft	1	10	75	1.1

[1] Crow-flight distance is between different municipalities. Thus, this coefficient is low for short-distance modes (especially for walking, bicycle, and urban transport), since some of those trips only cross the boundary between two neighboring municipalities. This coefficient is smaller for long trips (e.g., by air) than for medium-distance trips.
[2] Only 70 km/h for mopeds.
[3] We have admitted a few verified exceptions up to 140 km/h door-to-door.
[4] Up to 250 km/h for the TGV (high-speed train).

Sources: INSEE-INRETS 1981–82 and 1993–94 NPTS.

TABLE 3 Validation and Correction of Daily Trip Data

	1993–94 survey[1]		1981–82 weekly diary[2]		
				After correction	
Mode	Original file	After correction	Original file	Stages	Trips
Unknown origin	12	10	1,052	53	47
	0.0%	0.0%	1.5%	0.1%	0.1%
Unknown destination	10	7	914	50	44
	0.0%	0.0%	1.3%	0.1%	0.1%
Distance unknown	1,812	2	2,732	96	81
	1.9%	0.0%	3.8%	0.1%	0.1%
Distance over 5 km less than	299	0	1,327	0	0
crow-flight distance	0.3%	0.0%	1.8%	0.0%	0.0%
Unknown duration	59	6	769	168	162
	0.1%	0.0%	1.1%	0.2%	0.2%
Speed too fast	194	0	108	0	0
	0.2%	0.0%	0.1%	0.0%	0.0%
Speed too slow	292	0	1,469	0	0
	0.3%	0.0%	2.0%	0.0%	0.0%
Unknown transport mode	74	73	203	59	48
	0.1%	0.1%	0.3%	0.1%	0.1%

[1] Previous day and last weekend trips in 1993–94.
[2] Week-long stage diary in 1981–82 was converted into a trip diary for comparison with the 1993–94 survey (see the two columns at the right side).

Sources: INSEE-INRETS National Transportation Surveys.

The Trips Diary: 1981–82

For daily trips in the 1993–94 NPTS, hot-deck imputation was not appropriate, because trips were described for typical days (Saturday, Sunday, and a weekday). In 1981–82, similar trips were more frequent, as they were reported in a weekly diary. Thus, hot-deck inside a diary could be used in order to fill nonresponses or to make data consistent.

After matching origin-destination and trip distance, hot-decks were run to fill nonresponses, first on transport mode and then on trip duration. The criteria used to find a correctly described trip similar to one with inconsistent or missing information are: 1) geography (origin and destination in the same municipalities), and 2) trip purpose to provide mode or trip distance to provide duration. The results were not as satisfactory as those of the 1993–94 survey: out of 66,000 trips, 81 missing values were left on trip length and 162 on trip duration.

Car Diaries

In 1981–82, as in 1993–94, the driver had to copy the odometer at the beginning and the end of each trip. This information is highly structured (mileage must increase throughout the diary giving an objective measurement of trip distance), but there are occasional missing odometer readings for trip ends. In order to fill them, we first tried a hot-deck method structured by origin-destination and duration. If this was not successful, we computed mileage proportional to trip duration or to crow-flight distance, while ensuring that the mean speed stayed within reasonable limits. Finally, we filled nonresponses on trip duration with a hot-deck run on geographical and distance criteria. At the end, there were no missing values left for mileage or trip duration, but departure and arrival times were still missing for 105 trips out of 58,000 in 1981–82, and for 2,485 out of 200,000 in 1993–94. This satisfactory result for distance and duration is partly due to the fact that we skipped not only the diaries where the interviewer mentioned underreporting (about 5% of them), but also those where the information necessary for imputations was missing on at least one trip (less than 1% of diaries).

Reweighting for Underreporting of Short Trips or Underestimation of Short Distances

In the last NPTS, a selected person in the household had to describe the trips he or she made during the day before the interview and during the last weekend. As the last Saturday could be as much as one week earlier, we suspect that imperfect memory could affect the responses. The car diary collected in the same survey gives a more homogeneous image through the course of the week. Table 4 compares the results from these two survey instruments.

For weekdays, the two survey instruments give similar data for car trips. Because car diaries cannot be completed by persons absent too long from home, information on additional long-distance trips was obtained by interview (e.g., the return trip from holidays). If we limit the scope to trips within an 80 km crow-flight distance from the residence of the household, however, total travel (in vehicles-kilometers) is almost the same. There are 2% fewer trips collected in the car diary, but their average length is a little higher (9.9 km in the car diary vs. 9.7 km for car drivers in daily trips). Because of large sample sizes, this small difference is significant at a level of .05 and denotes a slightly different understanding of the notion of trip when the driver completes the diary alone, without the assistance of the interviewer (short stops may be omitted).

The previous weekend was too far in the past to ensure accurate memory of trips taken. Underestimation occurred about 30% of the time for very short trips (under 2 km). For longer trips, those on Sunday were a little less underreported than Saturday trips, probably because they were more recent. Thus, we used the figures in the two right columns of table 4 as correction coefficients for all motorized weekend trips. The figures offset the bias on average, but we are not sure that they

TABLE 4 Total Number of Car Driver Trips: Comparison Car Diary and Daily Trips

Trip distance	Weekday	Saturday	Sunday
<2 km	1.01	1.29	1.32
2–11 km	0.97	1.21	1.16
12–44 km	0.98	1.19	1.12
>44 km	0.91	1.06	1.00

Source: INSEE-INRETS 1993–94 NPTS.

correctly show the distributions, since they add the omitted trips to respondents who have described some and not to those who have declared none. In fact, if we compare the distribution of weekend trips for the persons interviewed on Monday with those obtained from later interviews, the proportion of zero trips explains less than 10% of the difference in average mobility (up to one-third for trips under 2 km). Thus, this reweighting method, which compensates, on average, for the underreporting of short weekend trips, does not seem to introduce a large bias in trip distributions (see tables 5 and 6). Moreover, this comparison shows almost the same rates of underreporting according to trip length as those obtained from the comparison with the car diary.

Comparison of the trip interviews and car diaries also allowed us to investigate drivers' perception's of distances. Controlled by the odometer, the car diaries estimated trip distance well. If we compare trips by class of crow-flight distance between origin and destination, we notice that long-distance trip lengths are a little overestimated. Moreover, there is a substantial underestimation of distance for trips whose origin and destination are in the same municipality; this underestimation is also observed for travel time, but it is less significant (see table 7). The underestimation of trip distance for car driver trips cannot be generalized to all modes. If we use the same coefficient of correction, many walking and cycling trips become too fast. Thus, in order to maintain consistency between time and distance variables, we could not implement a uniform correction for the underestimation of local trip distances.

TABLE 5 Frequency of Trips According to the Day of Interview (in percent)

Day	0 trip	1–2 trips	3–4 trips	>4 trips	Total
Total	23.9	25.2	22.1	28.8	100.0
Mon.	21.6	23.3	23.1	32.0	100.0
Tues.–Wed.	22.9	25.1	21.8	30.2	100.0
Thu.–Sat.	25.5	25.8	22.1	26.6	100.0

TABLE 6 Frequency of Short Trips by Day of Interview (in percent)

Day	0 trip	1–2 trips	>2 trips	Total
Total	84.7	11.0	4.3	100.0
Mon.	81.6	12.3	6.1	100.0
Tues.–Wed.	84.8	11.0	4.2	100.0
Thu.–Sat.	85.6	10.5	3.9	100.0

Note: Short trips are under 2 km.

In filling item nonresponses and verifying the consistency of data, geographical information plays a key role. That is why we have systematically used origin and destination in hot-decks. This information is accurately recalled by interviewed persons, but has to be geographically encoded during data processing. Manual coding is done only for difficult cases, since most municipality names in Europe can be automatically identified and coded (Flavigny and Madre 1994). Coding at a more detailed level is still a problem, except in some large urban areas (e.g., Montreal or Paris (Chapleau 1997)). In the case of car diaries, data are also strongly structured by the odometer. The comparison between different kinds of survey instruments allows us to assess memory effects and to detect substantial biases in the perception of short distances in travel diaries. Reweighting procedures are not always successful in correcting these biases, however. Thus, in the future, the need to collect data on trip distance will probably decrease, since this essential parameter of transport behavior can be calculated by traffic assignment algorithms, if the knowledge of locations (origin and destination) is sufficiently precise.

CONCLUSIONS

To some extent, the methods presented in this paper are specific to the context and characteristics of the NPTS. The analysis of the nonresponse mechanism for post-stratification relies on the availability of an exhaustive and up-to-date sampling base. Working with the National Institute of Statistics and Economics Studies, we had the opportunity, in 1993–94, to draw the sample from the relatively recent 1990 census. In some countries, this is not possible because of privacy and confidentiality concerns.

Some amount of household information is needed to compute imputations; implementation weighting procedures do not have this requirement. Therefore, weighting is the appropriate

TABLE 7 Car Driver Local Trips[1] Seen Through Different Survey Instruments

	In the same			Origin and destination (O-D) in distant municipalities: <15 km			Origin and destination (O-D) in distant municipalities: >15 km			Total		
	DT	CD	DT/CD	DT	CD	DT/CD	DT	CD	DT/CD	DT	CD	CD/DT
Travel diary and car diary in 1981–82[2]												
Number of trips (millions)	79.0	172.0	1.04	164.0	155.0	1.06	29.0	29.0	100.00	372.0	356.0	1.04
Trip length (km)	2.8	3.7	0.76	8.3	9.1	0.90	37.6	39.1	0.96	7.9	8.9	0.88
Crow-flight distance (km)	0.0	0.0	—	5.9	6.0	0.99	28.5	28.2	1.01	4.8	4.9	0.99
Trip duration (mn)	9.7	10.1	0.96	17.0	17.0	1.00	42.8	44.1	0.97	15.5	15.8	0.98
Mean speed (km/h)	17.2	22.0	0.78	29.1	32.2	0.90	52.7	53.2	0.99	30.5	33.8	0.90
Daily trips and diary in 1993–94[3]												
Number of trips (millions)	193.0	199.0	0.97	230.0	216.0	1.06	56.0	55.0	1.02	479.0	470.0	1.02
Trip length (km)	2.6	3.4	0.77	8.8	9.3	0.95	37.4	36.2	1.03	9.7	9.9	0.97
Crow-flight distance (km)	0.0	0.0	—	6.3	6.4	1.00	28.5	28.8	0.99	6.4	6.3	1.01
Trip duration (mn)	8.8	9.6	0.92	16.4	16.7	0.98	41.4	40.2	1.03	16.2	16.4	0.99
Mean speed (km/h)	17.8	21.2	0.84	32.4	33.6	0.96	54.3	54.0	1.01	35.7	36.3	0.98

[1] As more long-distance trips were collected by interview than in a travel diary, we considered only local trips whose origin and destination were within 80 km from the residence, using a household car.
[2] DT collected in a weekly stage diary; CD refers to a weekly car diary.
[3] DT collected by interview on the previous day and on the last weekend (only single-mode trips; for multimodal trips, distance made by car and precise O-D are unknown). CD here refers to the same kind of weekly car diary; excluding trip purpose "to the station" (for comparison with single-mode trips).

Key: DT = daily trips; CD = car diary.

Sources: INSEE-INRETS 1981–82 and 1993–94 NPTS.

method for coping with unit nonresponse, while imputation is used to correct item nonresponse (Zmud and Arce 1997; Armoogum and Madre 1997). Of course, there are always intermediate cases, as illustrated by the example of omitted trips, in which the choice of method is not as clear.

We have also modified trip weights to correct for memory effects. This compensates for the trip length bias by increasing average mobility, but could distort trip distributions by adding travel distances when respondents declare trips. Imputation could be another solution to this problem (Polak and Han 1997), but we lacked information to implement it. Indeed, in order to be cautious, all our imputations have used either external information (e.g., deriving trip distance from crow-flight distance) or information concerning the same person or the same diary. In any case, there was some interaction between weighting and imputing for car diaries, since we skipped all diaries where the information needed for imputation was missing for at least one trip. Thus, they were considered as missing units and were corrected by weighting.

In the future, travel surveys will make greater use of computer-assisted survey methods. Automatic checking of the data as soon as they are collected, either face-to-face (CAPI) or by phone (CATI), will allow the immediate correction of many errors by asking more details of the respondent. Nonetheless, corrections a posteriori will still be necessary for self-completed questionnaires. New approaches, such as artificial intelligence and neural networks, are now being tested for a new European Program on survey methods (MEST 1996 and TEST 1997).

ACKNOWLEDGMENTS

The work reported here benefited from the scientific support of J.C. Deville (INSEE) and from the comments of Professor M. Lee Gosselin (University Laval in Québec) and P. Bonnel (Transport Economics Laboratory-ENTPE, Lyon). The French Department of Transportation (DRAST) funded the study. It also contains some of the results of the EU-funded 4th framework projects MEST and TEST, "Methods and Technology for European

Surveys of Travel Behavior," involving the Institut für Strassenbau und Verkehrsplanung, Leopold-Franzens-Universitt Innsbruck, Statistics Netherlands, Bro Herry (Wien), University of London Centre for Transport Studies, Imperial College (London), Deutsche Versuchsanstalt für Luft- und Raumfahrt (Köln), INRETS (Arcueil), Transportes, Inovao 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 authors and are not those of the CEC or the Consortium.

REFERENCES

Ampt, E. and J. Polak. 1996. An Analysis of Nonresponse in Travel Diary Surveys, presented at the 4th International Conference on Survey Methods in Transport, Steeple Aston, England.

Armoogum J., Y. Bussière, and J.L. Madre. 1994. Longitudinal Approach to Motorization: Long-Term Dynamics in Three Urban Regions, presented at the IATBR Conference, Valle Nevado, Chile.

____. 1995. Demographic Dynamics of Mobility in Urban Areas: The Paris and Grenoble Case, presented at the 7th World Conference on Transport Research, Sydney, Australia.

Armoogum, J. and J.L. Madre. 1997. Item Sampling, Weighting, and Nonresponse, presented at the International Conference on Transport Survey Quality and Innovation, Grainau, Germany.

Chapleau, R. 1997. Conducting Telephone Origin-Destination Household Surveys with an Integrated Informational Approach, presented at the International Conference on Transport Survey Quality and Innovation, Grainau, Germany.

Flavigny, P.O. and J.L. Madre. 1994. How To Get Geographical Data in Household Surveys, presented at the IATBR Conference, Valle Nevado, Chile.

Madre, J.L. and J. Maffre. 1994. The French National Personal Transportation Survey: The Last Dinosaur or the First of a New Generation? presented at the IATBR Conference, Valle Nevado, Chile.

MEST. 1996. Improved Methods for Weighting and Correcting of Travel Diaries. *4th Framework Programme.*

____. 1996. Long Distance Diaries Today: Initial Review and Critique. *4th Framework Programme.*

____. 1996. Possible Contents and Formats for Long-Distance-Travel Diaries. *4th Framework Programme.*

____. Sampling and Weighting Schemes for Travel Diaries: Review of Issues and Possibilities. *4th Framework Programme.*

Polak, J. and X.L. Han. 1997. Iterative Imputation Based Methods for Unit and Item Nonresponse in Travel Diary Surveys, presented at the IATBR Conference, Austin, Texas.

Sammer, G. and K. Fallast. 1996. A Consistent Simultaneous Data-Weighting Process for Traffic Behavior, presented at the 4th International Conference on Survey Methods in Transport, Steeple Aston, England.

Sautory, O. 1993. Redressements d'un Echantillon par Calage sur Marges, INSEE document de travail, no. F9310.

TEST. 1997. Technology Assessment. *4th Framework Programme.*

Zmud, J. and C. Arce. 1997. Item Nonresponse in Travel Surveys: Causes and Solutions, presented at the International Conference on Transport Survey Quality and Innovation, Grainau, Germany.