REFERENCES

1. M. Wermuth. Effects of Survey Methods and Measurement Techniques on the Accuracy of Household Travel-Behavior Survey. In New Horizons in Travel-Behavior Research (P.R. Stopher, A.H. Meyburg, and W. Brög, eds.), Lexington Books, D.C. Heath and Co., Lexington, MA, 1981.
2. W. Brög and A.H. Meyburg. Nonresponse Problem in Travel Surveys: An Empirical Investigation. TRB, Transportation Research Record 775, 1980, pp. 34-38.
3. Sozialforschung Brög. Kontinuierliche Erhebung zum Verkehrsverhalten, KONTIV 76 [Continuous Survey of Travel Behavior, KONTIV 76]. Ministry of Transport, Bonn, West Germany, 1976.
4. SOCIALDATA GmbH. Verkehrsuntersuchung Spandau--Auswirkungen des U-Bahn Baues auf das Verkehrsverhalten [Traffic Study Spandau--Effects of Subway Construction on Travel Behavior]. Senator for Construction and Housing, West Berlin, West Germany, 1979.
5. R. Höttler. Ausschöpfung von Stichproben und die Non-Response Problematik in Theorie und Praxis [Response Rates and Problems of Non-Reponse in Theory and Practice]. Paper presented at the 5th Workshop of the German Transport Science Society, Grainau-Eibsee, West Germany, Nov. 1979.
6. B.A. Moolman. A Fresh Look at Non-Sampling Errors in Survey Research. Paper presented at the 5th Workshop of the German Transport Science Society, Grainau-Eibsee, West Germany, Nov. 1979.
7. SOCIALDATA GmbH. Marktanalyse des Verhaltens im Personenfernverkehr [Market Analysis of Intercity Travel Behavior]. Ministry of Transport, Bonn, West Germany, 1979.
8. Studienkreis für Tourismus e.V. Urlaubsreisen 1979, Ergenbnisse der Reiseanalyse [Vacation Travel 1979, Results of a Travel Analysis]. Starnberg, West Germany, 1980.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Validity Problems in Empirical Analyses of Non-Home-Activity Patterns

ARNIM H. MEYBURG AND WERNER BRÖG

Validity problems of empirical data have been neglected to a large extent in the transportation planning field. This paper illustrates the impact that choice of survey method has on the validity of the data. It shows that the recorded data depend directly on the method selected for obtaining them. An uncritical application of survey methods is not justifiable and, in fact, can lead to incorrect survey results. Basic research in the area of empirical survey methods is long overdue. An international exchange of experiences in this regard is considered most beneficial, as illustrated by this paper. The exchange of information and insights is often hampered because the survey methods used for specific investigations tend to be inadequately documented. This deficiency makes subsequent assessment of data validity very difficult, if not impossible. Furthermore, the use of such data without consideration of the underlying survey method is dangerous. The paper cites examples where the results of analyses can be manipulated by means of different survey methods. Greater efforts should be made to integrate data collection with the research effort performed on the basis of these data.

Many transportation planners, engineers, and modelers have, for all too long, ignored the quality of the basic input to their research efforts, namely the data. Since much of this research is of an empirical nature, the data are obtained through empirical surveys. This paper is intended to add to the efforts concerned with survey methodology for empirical analyses of travel activities. It presents a number of examples that show how the survey method and design can influence the results of an investigation.

We recognize that only a limited set of examples can be shown in the context of a paper of this nature. Further, the basis of comparison is yet another survey, albeit one that is generally recognized as representative of the state of the art in survey methodology [e.g., the national travel survey (KONTIV) (1)]. We need to accept the quality of that survey and the validity of its results in order to believe in the results of the research presented in this paper.

Even if the argument is made that the different survey results obtained through two surveys do not prove the correctness (or lack thereof) of one or the other survey results, the disturbing fact remains that different results about mobility were obtained when the study objectives were identical and only the survey method differed. This result alone is worth keeping in mind. The transportation planner or modeler is well advised to pay careful attention to the procedure used to generate the inputs that are used in any modeling effort or in the generation of simple mobility statistics. The validity (correctness) of the data will also determine the validity of any model or statistical results, which in turn might be used as the basis for policy and investment decisions.

VALIDITY VERSUS REPRESENTATIVENESS

The paper concentrates on the validity aspects of survey results because problems of validity of empirical data have been neglected to a large extent in the transportation planning field. The issue of validity of survey data transcends that of representativeness. Representativeness addresses the question of whether we have enough data points for each of our strata or cells. The concept of validity is aimed at questions of whether the data obtained are valid (correct or relevant) or whether they are an accident ascribable to a particular method of data collection. We attempt to show in this paper that empirical results based on survey data typically contain substantial errors that result in severe misrepresentation of reality.

The error sources addressed in this research lie exclusively in the survey method and design employed to generate these data. A fair assessment of present survey and research practice in transportation probably is that these distortions of reality go largely undetected. This is in part due to the researcher's unawareness of any problems of bias in the data and to a great extent to the preoccupation with the modeling and analysis phase of the typical research undertaking. Since data collection is such a tedious and costly element of any empirical research effort, it is somewhat understandable, yet not justifiable, that some researchers are unwilling to look back when a data base has been generated that can be analyzed. Any additional effort that questions the validity of the information seems to detract from the real research work and progress.

VALIDITY PROBLEMS OF EMPIRICAL SURVEY RESULTS

The following sections contain a number of examples of the impact of survey technique and survey instrument on the validity of the results of studies of travel behavior, irrespective of the quality and structure of any models used in the analysis phase. The items illustrated here do not represent an exhaustive set of influence factors, but they are intended to provide an array sufficiently broad enough to show the severity of the problem. The issue of nonresponse, which is certainly an element of validity problems, is not dealt with in detail in this paper because it is discussed extensively elsewhere (2) and in our other paper in this Record.

Use of Perceived Versus Actual Values

Investigation of travel distances on the basis of information from survey respondents is very difficult, due to the difference between actual and perceived values. A Dutch study (3) shows an overall overestimate of approximately 10 percent based on perceived (reported) distances [Table 1 (3)]. Of course, these results also vary by mode (ranging from 0.8 percent for transit to 15.7 percent for moped) and with the actual distance traveled. Obviously, the use of perceived distances as input to other investigations, such as determination of travel speed, produces an error at the input stage. The problem becomes even more pronounced if airline distances are used.

In a separate investigation in West Germany that dealt with the effects of using reported rather than actual (measured) values [a study performed to determine price elasticities for travel by transit (4)], similar discrepancies were detected. For reported distances by the mode actually used, a close match between the Dutch and the German results can be observed for automobile (+10.7 percent versus +10.2 percent); however, the difference is more significant for public transit (+0.8 percent versus +4.5 percent). Of course, keep in mind that in the latter study only estimates by users of the mode under consideration were employed [Table 2 (4)].

A comparison of reported and actual travel times by automobile and transit shows significant overestimates [Table 3 (4)]. This is particularly true for automobile drivers in the judging of travel times by transit--61 percent overestimated transit travel time by more than 20 percent. Even transit riders overestimated their travel times on the average by 10.4 percent (Table 3).

The picture is much less dramatic for estimates of travel time by automobile, where drivers overestimate their times by 8.4 percent and transit riders judge automobile travel time on the average to be 4.0 percent more than the actual values (Table 3). When evaluating this information, keep the stratification of misestimates by percentage (provided in the table) in mind because the averages tend to veil a number of interesting details. For example, the largest percentage misestimate (28 percent) of automobile travel time lies in the 21 percent-plus stratum of transit riders.

For travel cost by automobile and transit [Table 4 (4)] we observed the well-known phenomenon that automobile travel costs tend to be underestimated by both automobile users and nonusers. The absolute misestimation is most pronounced for automobile drivers (62.4 percent) and the largest underestimation is in the 21 percent-plus stratum (57 percent). Naturally, transit riders report their transit fare accurately because the out-of-pocket costs for that mode are obvious and easy to recall.

Influence of Elapsed Time on Reported Trip Volume

Another serious influence on the results of mobility

Table 1. Difference between reported and actual travel distance.

Travel Mode	No. of Trips	Total Distance (km) Reported	Total Distance (km) Measured	Difference (%)
Car	1342	12 352	11 162	10.7
Truck or delivery van	40	860	814	5.6
Moped	100	464	401	15.7
Bicycle	843	1 583	1 467	7.9
Walk	1004	626	594	5.3
Public transit	104	735	729	0.8

Table 2. Reported travel distance by mode used for all trip purposes.

Estimate	Automobile Driver (n = 930) (%)	Transit Rider (n = 2327) (%)
Overestimation		
By 21 percent and more	22	17
By 11-20 percent	8	13
By 1-10 percent	3	6
Total	33	36
Correct	42	28
Underestimation		
By 1-10 percent	7	7
By 11-20 percent	12	18
By 21 percent and more	6	11
Total	25	36
Average misestimation	+10.2	+4.5
Absolute misestimation	|18.5|	|18.2|

Table 3. Reported travel time for all trip purposes.

Estimate	Perceived Travel Time by Transit (%): Automobile Driver (n = 400)	Perceived Travel Time by Transit (%): Transit Rider (n = 2380)	Perceived Travel Time by Automobile (%): Automobile Driver (n = 891)	Perceived Travel Time by Automobile (%): Transit Rider (n = 1306)
Overestimation				
By 21 percent and more	61	27	28	20
By 11-20 percent	10	16	11	11
By 1-10 percent	5	16	7	2
Total	76	59	46	33
Correct	9	9	7	14
Underestimation				
By 1-10 percent	3	14	10	9
By 11-20 percent	4	9	22	16
By 21 percent and more	8	9	15	28
Total	15	32	47	53
Average misestimation	+28.9	+10.4	+8.4	+4.0
Absolute misestimation	|36.5|	|20.6|	|25.8|	|31.2|

Table 4. Reported travel cost for all trip purposes.

	Perceived Travel Cost by Automobile (%)		Perceived Travel Cost by Transit (%)	
Estimate	Automobile Driver (n = 870)	Transit Rider (n = 841)	Automobile Driver (n = 331)	Transit Rider (n = 2397)
Overestimation				
By 21 percent and more	21	7	21	1
By 11-20 percent	4	5	15	1
By 1-10 percent	1	2	1	0
Total	26	14	37	2
Correct	5	22	38	93
Underestimation				
By 1-10 percent	5	4	2	2
By 11-20 percent	7	4	18	0
By 21 percent and more	57	56	5	3
Total	69	64	25	5
Average misestimation	-3.1	-25.4	+12.4	+0.2
Absolute misestimation	|62.4|	|39.7|	|22.2|	|2.4|

Table 5. Influence of elapsed time on reported trip volume.

	Travel Mode				
Elapsed Time Before Survey	Total	Automobile	Train	Air	Other
Intercity Vacation Trips During 1 Year					
First quarter, travel directly before survey	100.0	100.0	100.0	100.0	100.0
Second quarter, 3-6 months	100.0	96.8	100.0	100.0	100.0
Third quarter, 6-9 months	95.7	90.3	100.0	100.0	100.0
Fourth quarter, 9-12 months	87.0	83.9	85.7	100.0	100.0
Other Personal Intercity Trips During 3 Months					
1 month	100.0	100.0	100.0	100.0	100.0
2 months	92.1	91.2	100.0	100.0	100.0
3 months	80.3	79.4	100.0	50.0	100.0

studies can be traced to memory lapses that obviously increase with the length of time for which travel activities are to be reported by the respondent. An example from the investigation of intercity vacation and personal travel is selected to illustrate this point [Table 5 (5)]. For intercity vacation trips the recollection of trips decreases by up to 14.3 percent for train travel and 18.1 percent for automobile travel undertaken more than nine months prior to the reporting date. The average underestimate of travel by all modes is 4.3 percent after a six-month and 13.0 percent after a nine-month time lapse. Also note that the more significant underreporting occurs for the more common modes, namely automobile and train, and air and other constitute more memorable (less frequent and costlier) events and result in accurate reports of vacation trip making.

In general, the reporting of other personal trips is even less reliable than that of vacation trips. After only three months, 19.7 percent of all trips are lost due to memory lapses, with the automobile and air modes being the main factors. One implication seems to be that vacation trips are more memorable and, therefore, more easily recollected.

Panel Effects on Reported Mobility

Data obtained through the use of a survey panel generally are considered to be a reliable source of information input for research studies. Aside from the fact that time-series data can be obtained by this method, it also represents a certain level of efficiency in sampling. The standard virtue of this survey technique is that a more or less consistent set of sample elements is available.

Unfortunately, a number of disadvantages are also associated with panel surveys, aside from the problem of setting up a willing set of respondents. As is illustrated in Table 6 (6), which represents the results of travel activity and trip frequency reports for a three-year panel survey performed in Munich, West Germany, substantial decreases in reported mobility can be observed over the three-year reporting period. This apparent decrease in overall activity and trip frequency represents a special hidden form of nonresponse influence. The respondents who are basically willing to participate in the panel survey, more and more frequently return their second- and third-stage questionnaire with the remark that they did not partake in any out-of-home activities during the survey day. Consequently, the share of immobiles (i.e., respondents who claim not to have performed any trips at all) increased by 6 percent from year to year, or more generally from phase to phase, and the average trip frequency of the mobiles (i.e., those who report out-of-home activities) remained relatively stable. Due to the fact that survey panels tend to measure artifacts of the methods rather than results, survey researchers are becoming more hesitant to use panels.

Comparison of Oral Versus Written Responses

Recent methodological research into survey methods has established that mail-back and interview surveys will produce different results for the identical reported phenomenon (7, and Brog and Meyburg in this Record). Underreporting and poor reporting tend to be the rule for oral responses. Table 7 illustrates the substantially different level of accuracy for length of time of travel generated by the two survey techniques. The average deviation from the correct (objective) travel time is -11 percent and +36 percent for automobile and transit, respectively, when an oral survey is used; however, deviations of only +3 percent and +12 percent are registered for written surveys.

How fundamentally oral interviews influence the result of the survey can also be seen in the parameter of "number of activities listed." For this purpose, diaries were kept for a week in two random samples. The first group, after receiving adequate instructions, filled out each day's activities by themselves. In the second group, the persons were orally questioned on the first day, and then filled out days two through seven by themselves. The surveys differed from each other only in this procedural method on the first day (i.e., case 1 = written survey, case 2 = oral and written survey). If the number of listed activities in the first case is set equal to 100, we get the results shown in column 2 of the table below (7).

	Index of Activities	
Day	Case 1 (n = 1162)	Case 2 (n = 882)
One	100	85
Two	100	99
Three-seven	100	100

Although the answers on days two through seven are comparable, oral interviewing lowers the result noticeably. Responsible are the following factors:

1. Unconscious mistakes on the part of the interviewees, who are forced to completely remember something within a short period of time;

Table 6. Panel effects on reported mobility.

Item	Spring 1977 Total (n = 1938)	Spring 1977 City (n = 1152)	Spring 1977 Region (n = 786)	Spring 1978 Total (n = 1938)	Spring 1978 City (n = 1152)	Spring 1978 Region (n = 786)	Spring 1979 Total (n = 1938)	Spring 1979 City (n = 1152)	Spring 1979 Region (n = 786)
Travel activity									
Did not leave home on survey day (%)	12	10	14	18	15	21	25	21	31
Left home on survey day (%)	88	99	86	82	85	79	75	79	69
Trip frequency									
No trips (%)	12	10	14	18	15	21	25	21	31
One trip (%)	0	0	1	1	1	1	1	2	1
Two trips (%)	42	40	47	39	38	40	34	36	30
Three trips (%)	8	10	5	7	8	5	7	8	5
Four trips (%)	20	22	18	20	21	19	18	19	16
Five trips (%)	7	8	5	6	7	5	5	5	5
Six trips (%)	8	8	8	7	7	6	6	6	7
Seven or more trips (%)	3	3	3	2	2	3	4	3	4
Average mobility									
Average trips per person for all survey days	2.91	3.02	2.74	2.70	2.78	2.58	2.50	2.57	2.38
Average trips per person for survey days that had trips	3.29	3.36	3.18	3.27	3.28	3.26	3.35	3.27	3.47

Note: Percentages may not total 100 due to rounding.

Table 7. Deviations of oral and written responses with respect to length of travel.

Estimate	Length of Travel: Oral Survey, Automobile (n = 800)	Oral Survey, Public Transportation (n = 520)	Written Survey, Automobile (n = 1100)	Written Survey, Public Transportation (n = 538)
Correct, within the permissible limit[a] (%)	72	79	93	88
Wrong, not within the permissible limit[a] (%)	28	21	7	12
Average deviation[b]	89	136	103	112

[a]Permissible limit is ±25 percent. [b]Correct (objective) time = 100.

2. Conscious mistakes on the part of the interviewees, who are unwilling to give a stranger certain information; or

3. Influence of the interviewer, who attempts to complete the interview as quickly as possible.

The diary filled out in the presence of an interviewer measures 15 percent fewer trips than the one completed with more leisure by the interviewee alone. We may conclude that interview surveys will produce substantially reduced mobility levels due to survey characteristics alone. To that element other factors have to be added, such as nonresponse influences (see our other paper in this Record).

Influence of Survey Instrument Layout on Reported Mobility

The KONTIV travel survey (1) conducted in the Federal Republic of Germany is generally considered an excellent example of the application of state-of-the-art survey design, implementation, and instrument layout. Therefore, a number of comparisons have been performed to test other results against those generated by KONTIV. The comparison of the KONTIV results with another survey of the travel behavior of senior citizens in Germany showed some interesting differences with respect to the mobility characteristics of that segment of the population. These differences could be traced to the somewhat inferior design of the survey instrument of the second survey.

That survey instrument has the following characteristics: The beginning of the survey form contains a filled-in example that leaves the respondent unsure as to whether return trips are to be reported. Also, the trip sequence in the example does not fit chronologically. Another problem is that the form has to be flipped over after the third trip. As a result of these problems, the survey shows many more survey days with an uneven number of trips, as can be seen in the table below (1,8).

No. of Trips Reported	KONTIV Survey Days (%) (n = 13 710)	EMNID Survey Days (%) (n = 6411)
One	2	17
Two	52	47
Three	8	16
Four	23	14
Five	5	3
Six	6	2
Seven and more	4	1

In the EMNID survey the number of days when four or more trips were taken is 18 percent smaller than in the KONTIV results (largely due to turning the page). The total mobility is reported to be 23 percent lower (2.52 versus 3.10 trips/person on the survey day) and this may be due largely to poor design of the survey instrument.

Substantial research on the effects of different survey instrument layouts and survey administration has been performed at the Socialdata Institute for

Empirical Social Science Research in Munich. A voluminous paper would be required to detail all research findings on these topics. The pitfalls a survey designer might encounter range from the obvious and trivial to the very subtle. The simple example of column versus row arrangement of the survey instrument illustrates how lack of methodological insight can lead to incorrect results. The KONTIV survey (1) and a comparative experimental study by Socialdata showed that a mobility difference per household of 9 percent was observed when a column arrangement was used in conjunction with simple check-off response possibilities. Also, the response rate proved to be 12 percent larger with the column layout.

CONCLUSIONS

We have presented selected results of past and ongoing research in the area of empirical survey methods. It is hoped that this paper can make a contribution to an increased level of awareness and knowledge of the dangers of uncritical use of data in travel behavior research.

This paper indicates that methodological experiments are necessary to improve the generated base information. Use of perceived versus actual values, influence of elapsed time, panel effects, nonresponse, oral versus written responses, and survey instrument layout have been shown to severely affect the validity of the survey data. These error sources can severely undermine the relevance and validity of any models and modeling results based on invalid data. One should keep in mind, though, that these factors constitute only a subset of factors that can influence survey and, therefore, modeling results.

It is quite common in our research community to use data collected for other than our own purposes or collected by some other organization that provides insufficient or no knowledge as to the survey method, design, administration, or questionnaire layout characteristics that were employed in the generation of the data set. In short, frequently the analyst or modeler is completely removed from the data source.

Clearly, given the various influencing factors on the validity of survey data, this fact produces some doubts as to the validity of many modeling results. In the course of preparing this paper, we encountered substantial evidence of a certain amount of obliviousness by some researchers as to how the data they used were generated. This problem becomes particularly serious when data of unknown or questionable validity are used in modeling efforts that result in claims of providing new insights into travel behavior. In our opinion, there is substantial reason to question the validity of these claims. It is quite evident that research results can be subject to manipulation by means of choice and execution of a survey procedure. For example, the results of ignoring the influence of nonresponse were illustrated by us for a number of cases (2, and in our other paper in this Record). We might also point out that sometimes one encounters a certain level of puzzlement and bemusement when one tries to get the researcher to detail the methods used to obtain the data that form the fundamental source of claims to new research insights. It is more than careless to conduct surveys without proper documentation of all details of the data collection method used or to use data without knowledge about the source and the survey method.

ACKNOWLEDGMENT

We gratefully acknowledge the useful inputs by Alfred Nickesen, Stephen Ritchie, and Orlando Strambi of Cornell University, who reviewed an earlier draft of this paper.

REFERENCES

1. Sozialforschung Brög. Kontinuierliche Erhebung zum Verkehrsverhalten, KONTIV 75,76,77 [Continuous Survey of Travel Behavior, 1975, 1976, 1977]. Ministry of Transport, Bonn, West Germany, 1977.
2. W. Brög and A.H. Meyburg. Nonresponse Problem in Travel Surveys: An Empirical Investigation. TRB, Transportation Research Record 775, 1980, pp. 34-38.
3. The Extent to Which People Can Estimate Their Trip Distances. Institute for Road Safety Research, Voorburg, Netherlands, Feb. 1978.
4. W. Brög and N. Meyer. From Price Elasticity to Price Susceptability--A Reality-Oriented Model for Influencing Transportation Fees According to Demand. Socialdata GmbH, Munich, West Germany, 1978.
5. Socialdata GmbH. Marktanalyse des Verhaltens im Personenfernverkehr [Market Analysis of Intercity Travel Behavior]. Ministry of Transport, Bonn, West Germany, 1979.
6. KONTIMUC: Survey of Transit Use in Metropolitan Munich. Socialdata, GmbH, Munich, West Germany, 1977.
7. W. Brög and K.H. Neumann. The Interviewee as a Human Being--A Critical Contribution to the Discussion of Methods in Empirical Social Research. Paper presented at the 44th European Society for Opinion Surveys and Market Research Seminar on Ways and New Ways of Data Collection, Jouy-en-Josas, France, Nov. 1977.
8. EMNID Institute. Type, Duration, Frequency as Well as Purpose and Destination of Travel Activities by People of Age 60 and Older in Selected Regions of the Federal Republic of Germany (in German). Bielefeld, West Germany, Sept. 1976.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.