The Urban Panel as a Longitudinal Data Source

R. D. WORRALL, Department of Civil Engineering, Northwestern University

This paper examines some of the problems posed by the lack of an adequate, longitudinal data set describing the locational preferences, daily activity sets and daily travel patterns of urban households, and proposes the use of a permanent response panel—analogous to the consumer panels employed in market research—as a means of collecting such information.

Consideration is given to the need for both long- and short-run data, with particular reference to (a) model structure and calibration, (b) forecasting, and (c) before-and-after studies. A brief description is given of a case study, employing the panel technique, in which data on daily travel behavior were collected for a sample of 50 households for a period of one month. Finally, a scheme is proposed for the continuous monitoring of urban travel demand, integrated within the framework of a continuing urban transportation study.

•THE development of a satisfactory land-use model is constrained by a wide range of considerations. One of the most common of these is the availability of an adequate data set. Existing data systems, particularly those developed prior to the current quantitative focus of model building activity, tend to meet only some of the analyst's requirements. At the macro-level, for example, he is often faced with a dirth of accurate economic and time-series data. At the micro-level there is usually an equivalent lack of detailed, disaggregate information, specified at the level of precision which he requires and couched in a longitudinal rather than a cross-sectional format.

This paper considers the operational feasibility of a specialized data system—the urban panel—designed to provide detailed, time-series information on the daily activity patterns, locational and behavioral preferences and trip-making characteristics of a controlled sample of individual households. The technique is based on the concept of the continuous household response panel, analogous to the consumer panels frequently employed in public opinion and market research. Its potential application to household behavior patterns has important implications both for the structure of future land-use models and for the design and operation of future data systems.

In the present case, emphasis is placed on the application of panel techniques to the study of urban travel behavior. This emphasis is merely a matter of convenience, reflecting the particular background of the author. The paper might equally well have focused on one of the other, parallel topics of interest to the urban model builder.

MONITORING URBAN TRAVEL

It is generally acknowledged that the structure of urban travel is subject to a continuous process of change. Constant modifications in the existing transportation environment—the network of roads, transit lines, and the distribution of potential destination points—coupled with parallel changes in the behavioral preferences and activity sets of the households making up the metropolitan population all exert a continuously changing influence on the daily demand for travel. Some of these changes—or "shocks" as they may be termed—have lasting, long-term effects (e.g., the opening of a new expressway or transit line); others yield only transitory consequences (e.g., a sudden snow storm

Paper sponsored by Committee on Land Use Evaluation and presented at the 46th Annual Meeting.

or the breakdown of the family car). In either case their detection and evaluation, under the generic term "induced and diverted traffic," is of considerable interest to the transportation analyst.

Conventional before and after and home-interview surveys provide two possible, but rather unsatisfactory, approaches. Both techniques provide the planner with only a limited, and rather ill-defined, information set based on conditions which pertained to one or perhaps two isolated points in time. They yield no really significant information on the structure of changes in travel behavior. Similarly, continuing volume-count or ridership surveys, though exhibiting the necessary temporal characteristics, do not contain sufficiently detailed information to permit a truly penetrating analysis to be performed.

A fourth line of approach, and that which will be enlarged on here, is the use of a continuous household response panel, designed to continuously monitor changes in travel behavior in a manner analogous to the panel techniques used in public opinion sampling and market analysis. A panel consists essentially of a set of respondent units whose behavior is observed over a continuous term. Assuming knowledge of the sample and population characteristics, the results of a panel study provide the analyst with a valid basis for the detection and analysis of time-dependent behavior [this interpretation of "panel" differs from that used by Davis (1)]. In the present case, the respondent units are households (or equivalent traffic generators), and the behavior under study is the day-to-day changes which occur in the established patterns of urban travel.

In its conventional format, the panel survey may be diagrammed as:

Group	Selection Criteria	Before Measurement	Natural Exposure to Treatment	After Measurement
Test	Nil	T_B	Yes	T_A
Control	Nil	C_B	No	C_A

There is an obvious equivalence between this two-wave design and a simple, non-randomized before and after survey with test and control groups. In this context, "wave" refers to a set of test-control observations made at a single point in time. A two-wave panel, thus, involves two cross sections in time, an n-wave panel involves n successive observations. The distinction should be drawn, however, between the artificially inserted treatments of such a before and after survey, and the natural occurrence of a "treatment" in a panel study completely divorced from the control of the analyst.

The two-wave panel illustrated may be extended to the more common n-wave panel simply by extending the sets of controlled C and test T measurements horizontally, resulting in a format superficially equivalent to a multiple time-series design, but with the provision that successive natural treatments may occur between any two successive time waves.

The design outlined may be modified in a variety of different ways. It need not, in fact, take the form of a formal panel design at all. It could equally well be based on successive random sampling from within a single, well defined selection-frame. Alternatively, it might utilize one of a variety of partial-overlap designs (2). Similarly, the panel units themselves may be defined variously as elements of a purposive quota sample, a stratified probability sample or a multistage cluster sample. The period and intensity of the actual monitoring process may vary, from the extreme of a comprehensive daily record to a regular, once-a-week analysis or a randomized coverage with average frequency of 14 or 28 days.

The permutation of possible designs is almost endless. So is the list of devices which may be used to circumvent the operational problems which arise in such a study (e.g., the decay of an initial panel due to migration and maturation within the sampling

frame). No attempt will be made here to treat all of these considerations comprehensively. Rather, attention will be directed toward the design of one or more feasible survey formats; feasible, that is, with respect to cost, complexity and the fundamental requirements of the travel analyst.

SAMPLE DESIGN—SELECTION OF RESPONDENT HOUSEHOLDS

Two major decisions must be made at the outset of any monitoring design: the desired intensity of the sample coverage and the desirability of sample stratification. Consider first some of the problems of sample coverage.

In a conventional home-interview study, a sample size of between 5 percent and 20 percent is aimed for, depending on the size of the urban area (the larger percent pertains to areas of less than 50,000 population, the smaller to areas of over 1,000,000). An equivalent rate would be impracticable and even undesirable in a continuing panel analysis. Two alternatives present themselves: the selection of a much smaller, wholly randomized sample, dispersed uniformly throughout the urban area, or the use of a small number of high intensity, cluster samples in which respondents are concentrated in only one or two different locations. The benefits and disadvantages of the two designs are relatively well defined; however, the choice between them is not quite so clear-cut.

A randomized sample, provided its size is adequate, will generally yield the more efficient estimate of elementary unit variance and the sounder basis for generalized induction. The cluster sample is cheaper to operate (per element) and provides for closer control over the external environment (2). The efficiency of the cluster statistics may if necessary be improved (at an increase in cost) by a process of multistage sampling.

In general, it is the author's feeling that the former, randomized design is likely to provide the more efficient basis for a continuous monitoring operation. Presumably, the analyst is concerned more with the identification of a general temporal effect, pertinent over the entire metropolitan area, than with the behavior of particular clusters within that area. Furthermore, the degree of effective environmental control attainable within a cluster sample, unless the survey is designed for a very explicit locational purpose (e.g., the evaluation of induction and diversion effects consequent on the opening of a new expressway), is likely to be minimal. A more effective general control is provided by the randomization process implicit in the non-cluster design.

Accepting the dispersed, randomized design in principle, one is then faced with the question of its desirability or stratification.

Superficially, stratification appears to be eminently desirable, preferably in terms of stratified random sampling or, if this proves too costly, in terms of purposive, quota sampling. It has been shown (3) that the intensity and, to a lesser extent, the temporal variability of daily travel to and from the home vary with the characteristics of a single household. Theoretically, therefore, in the sense that efficient stratification results in more efficient estimates of sample parameters and also permits more flexible and perhaps more appropriate subsample analysis, a stronger panel design should result from making use of this additional information.

Unfortunately, however, this is not the case. An efficient stratification scheme is predicated on the identification of a limited number of pertinent variables, all of which relate closely to the phenomena under study and all of which are based on an available and reasonable recent information set.

In the case of travel analysis, a variety of parameters act to influence the behavior of a single household—family size, structure, income level, occupation, maturity—the importance of each varying with the precise behavior under analysis. The nature and diffusion of these characteristics generate two serious problems.

Information on many of the pertinent parameters is not readily available in conventional sampling frames. (The census, for example, contains neither disaggregate measures of household structure nor of stage in the family life cycle, both of which are of significant importance in determining the frequency of daily trip-making from the home.) Similarly, frame data generally used in stratification and quota design—age, sex and race

measurements—are not particularly pertinent to the study of household travel. This, therefore, raises a serious problem of frame selection and stratification control.

Even if an adequate sampling frame is identified, the diffusion of pertinent variables introduces a second and equally serious combinatorial problem. Consider, for example, six variables divided into three strata each. This yields a requirement for 3^6 or 729 separate stratification cells. Given that considerably more than one observation is required per cell, the total sample size rapidly reaches astronomical proportions. A variety of extremely complex, stratified probability designs are proposed in the literature (2) to overcome this latter problem. None of these, however, is particularly pertinent to the present discussion. Conventional, simplistic stratification is virtually meaningless.

There remain two alternative lines of approach: either a purposive, nonrandomized design, subject to all of the consequent biases and external invalidities of non-experimental research, or the use of a simple, non-stratified but wholly randomized selection procedure. The optimal choice between these two cannot be generalized usefully. For some panel analyses, the purposive design is most appropriate (e.g., the study of reaction to a rapid transit innovation where the effect is postulated to be restricted to a limited subpopulation); in others, such as the generalized monitoring scheme proposed here (where it is intended to induce general conclusions from specific survey data), a wholly randomized design is indicated.

In each case, considerations of cost dictate a sample size no larger than 2,500 households, whereas requirements of statistical efficiency, as based on the variability of behavior observed in a 4-week case study of 104 households in Skokie, Illinois (see Appendix), suggest a minimal, non-supplemented sampling ratio of not less than $\frac{1}{4}$ percent to $\frac{1}{8}$ percent. These latter figures are theoretical estimates based on an assumption of no significant taxonomical structure. There is evidence to suggest that some such structure may well exist, based on an analysis of the case study data referred to previously. Should further research, based on the initial operation of a randomized panel sample, reveal the existence of a meaningful taxonomy of household types, then these may be used as a foundation for a more efficient stratification design and hence a reduction in the total sample size.

So far it has been tacitly assumed that the monitoring operation is to be performed only on household travel and not on a combination of residential, industrial and commercial trip-making. Some brief justification for this assumption is in order.

Residential based travel, particularly if allowance is made for the individual links in each home-home trip sequence, accounts for more than 80 percent of the total daily travel in an urban area. Wholly commercial and industrial trip-making adds up to only 20 percent or less of the total. One may logically argue, therefore, that any continuous monitoring of nonresidential traffic is likely to yield only an insignificant, and possibly rather expensive, marginal return. Furthermore, the extended study of a firm's distributory activities at the level of detail required here, makes for a considerable imposition on its traffic department. It is in fact virtually impracticable to collect such data on a continuing basis (4).

Should it be considered essential that data on nonresidential travel be made available to supplement that derived from the household sector, it is suggested that a separate subsample of industrial and manufacturing concerns be developed, which would then be subjected to a considerably less intensive and more specific surveillance. In particular, it is suggested that the onus of recording the necessary information be allowed to fall directly on the survey staff rather than be left to the firm's employees.

In view of the major logistical problems involved and the probability of only marginal returns, no further consideration will be given here to the monitoring of nonresidential travel.

CHOICE OF SAMPLING FRAME

A wide variety of alternate sampling frames are available to the urban survey designer, ranging from census records, city directory or voters' lists to utility company files and the results of recent specialized surveys. A detailed discussion of each of

these and the appropriate techniques to follow in their use would be inappropriate here. Extensive discussions are given in the methodological literature, particularly by Kish (2, Ch. 8, 9, 11) and Hansen et al (5). However, a number of brief, general comments may be made at this point.

Probably the most efficient sampling frame for a monitoring study is a recent home-interview traffic survey. Such a survey is generally based on an equal probability sample selection method or systematic sample drawn from a completely updated dwelling list. Its use also has the advantage both of tying in the survey directly with a previous data collection exercise and also providing some additional base data for sample control and possibly for sample stratification (the latter comment presumes the identification of the meaningful household taxonomy discussed earlier).

If a survey is unobtainable, then refer to one or the other of the standard frames mentioned. In this event, a variety of problems are likely to be encountered—missing or duplicate listings, ambiguous and rare elements, incompatible subsample sizes and so on. Each of these is treated in detail in the literature, for example, by the use of multistage "compact-segment" techniques (2, pp. 313-315).

Perhaps the most serious frame problem which is likely to arise in a monitoring design is that of sample maturation. Approximately one in every five American families changes its place of residence each year. This means that over a period of two years migration alone may reduce the original sample base by 35 percent. As Sobol (6) has shown, response decay may increase this figure still further, while similar effects may arise due to population growth, births and deaths, and marriages.

A number of gimmicks may be employed to reduce the bias introduced by such maturation processes. Kish (2, pp. 472-474) has suggested the use of dwelling units or compact segments—i.e., clusters of adjacent dwelling units whose composition but not periphery may change over a period of time—rather than families as the basic repetition unit. Sharp has suggested the matching of neighboring household pairs (7) and several writers have discussed techniques for continuously updating existing sampling frames. The final choice is up to the individual analyst. No one technique is uniquely preferable to any other, although one may perhaps argue in favor of Kish's compact-segment technique in the case of a repetitive random sample design and his fixed dwelling unit method in the case of a formal panel study.

SURVEY FORMAT

Given that an acceptable sample has been developed, one further fundamental decision remains to be made. This centers on the selection of formal panel techniques vs the use of repetitive random samples or partial overlap designs. (In a repetitive, randomized design the same single set of households is not included at each stage of the survey. Separate random samples are drawn for each survey wave from a common sampling frame.) The primary benefits to be gained from a panel study (in which the same, randomly selected households are subjected to continuing analysis throughout the study period) are an improvement in statistical inference (2, p. 475), a reduction in overall costs and a considerable increase in the total level of information obtained. Counterbalancing these advantages is a problem due to sample and response decay, sample maturation, surveillance and possible external invalidity.

The severity of these problems is significantly reduced in both of the alternative designs cited. The use of repetitive random samples, for example, removes all question of sample maturation and response decay. It also allows for a relatively high degree of external validity, especially if the randomized sample is inserted into a continuing, quasi-experimental design, such as "separate sample, pretest-post-test control" design. A partial overlap design (in which a specified sample turnover—say, 25 percent or 33 percent—occurs on each survey wave) similarly avoids all major problems of sample maintenance and surveillance. In neither case, however, are the total survey costs as low nor the total amount of longitudinal data obtained as great as for the panel study.

As in the case of the sampling frame, there is no simple, optimal choice between the three possible formats. Each has its own peculiar advantages and disadvantages,

none of which significantly outweigh those of the other two. In fact, for a continuing, monitoring operation in which a major concern is the development of a longitudinal data set defining the temporal stability of household travel, a case may be made for the simultaneous use of all three, with the panel study forming a core around which supplementary randomized or overlapping designs may be structured. This is the approach selected here and illustrated subsequently in an example panel design.

SURVEY OPERATION

Perhaps the most tempting pitfall in any continuing survey is the collection of too much information. It is extremely important, therefore, that the range and intensity of the monitoring questions, the frequency with which they are asked and particularly their pertinence to travel analysis be strictly controlled. Considerable caution should be exercised lest the monitoring operation turn into a general behavioral surveillance, with all of the obnoxious implications that such a process bears for the encroachment of privacy.

Given that the core of the proposed monitoring operation is to be formed by a continuing panel study, it is suggested that the self-administered travel diary illustrated in the Appendix and used successfully in the case study described be adopted as the primary data collection mechanism. The diary is designed for flexible use over periods varying from one day up to several weeks. The range of information requested is compact and comprehensive, and the diary format has been proven acceptable to a sample of laymen. Its use in the case study did not reveal significant response bias, either in terms of its acceptance by different types of household or of its sustained use as a travel record over a period of several weeks. Where additional information over and above that recorded in the diary is required, supplementary questionnaires could be used. It should be stressed again, however, that the temptation to ask too many supplementary questions should be vigorously avoided.

Use of the personal travel diary as the primary data collection mechanism gives rise to a potential problem of data redundancy. Although a recent empirical analysis emphasized the essential temporal instability of household travel, it also revealed a considerable degree of daily repetition, particularly from the viewpoint of the respondent who may be forced to record the details of a single, repetitive trip for each of several days for several weeks (3). Two suggestions may be made to reduce the labor of this exercise.

First, those trips which are known to be highly repetitious—i.e., the journey to work and the journey to school may be pre-coded—and, once their regular pattern has been established, full details may be requested only on the occasion of a divergence from this pattern. This technique was tested on ten of the sample families included in the case study and was found, subject to adequate surveillance, to yield acceptable results.

Second, a much more effective device in reducing the redundancy and labor of daily data recording is to vary the period and technique of the monitoring process. During the case study, data were collected on a day-to-day basis over the entire survey period. Obviously, while this is feasible and in fact desirable for a limited 4-week study, it is both undesirable and unnecessary in a continuing study.

It is suggested, therefore, that an operational format be adopted in which the sample families are requested to maintain a daily diary for an initial period of 3 to 6 weeks. During this time they would be subjected to training and surveillance at the rate of two calls per week. At the end of this period, the contact frequency would drop to an average of two days per month, with the respondents being requested to maintain a daily trip record only for those two specific days. The selection of the successive contact dates would be randomized within the selected sampling frame.

Primary reliance would be placed throughout on the self-administered questionnaire. After the initial 3- to 6- week period each successive contact would be made by mail (i.e., a self-administered diary is mailed to the family two or three days ahead of the sample day), supplemented by a telephone call on the immediately preceding day and a subsequent personal call by a member of the survey staff to collect and vet the completed form. During this personal contact, a series of brief checks would be performed

on the completed diary by the staff member. Further detailed checks would take place in the survey office as part of the audit procedure (see Appendix). Supplementary contacts, as necessary, would be made primarily by mail or by telephone. Minimum reliance would be placed on extensive personal interviews. It is suggested that considerable use be made of telephone contacts to detect the consequences of a particular, short-lived environmental "shock" and possibly to evaluate public reaction to a particular experiment in system control. An example of the latter may be the efficacy of helicopter surveillance of freeway congestion and the broadcasting of alternate routes over commercial radio channels.

The panel study outlined makes up the core of the proposed monitoring operation. It provides the nucleus of a capability to detect and evaluate changes in travel demand on a continuing period, and at the same time measures public reaction to transportation innovation and change. Extended over a period of several years, it would provide a base of longitudinal data of considerable utility to the transportation analyst. Its format is entirely flexible. The duration of the initial, intensive survey period might vary from 3 to 6 weeks; the frequency of subsequent contact might range from one day/month to one day/week. (Extension or subtraction beyond these extremes is likely to result in significant over or under collection of information.)

Centered around this core would be a variety of less frequent, non-panel designs. The structure of these supplementary studies might also vary considerably, depending on the intensity of the panel coverage, the precise information requirements, and the size of the urban area. One might consider a three monthly, randomized set of two-day samples, or an annual, one-week study, again based on a wholly randomized sample selection procedure. Similarly, the supplementary studies might incorporate a degree of successive overlap—say, 25 percent or even 50 percent—and they might involve either the use of the standard diary format or a supplementary, simplified questionnaire. Contact might be made by mail, by telephone (preferable for the simpler forms of survey), or by conventional home interviews. The extensive use of the home-interview technique obviously suffers from the primary disadvantage of high relative cost. It is, however, likely to provide the best return on a given investment for the more complex forms of investigation, while the maintenance of a permanently employed, highly trained survey staff introduces considerable scale economies into a continuing transportation study with obligations to provide limited amounts of up-to-date survey data. The coordination of the proposed monitoring scheme within such a continuing transportation study is discussed in the next section.

Some form of respondent incentive was found in the case study to be essential to the efficient operation of a permanent monitoring scheme. It is suggested that using trading stamps redeemable at a local store for goods of the respondent's choice, and employed effectively in the case study, be utilized for this purpose.

A truly accurate estimate of the level of payment necessary to maintain a sustained, rather than a month-long response period cannot be deduced directly from the case-study findings. In the absence of more adequate information, it is suggested that a payment level of 50 trading stamps/day plus a terminal bonus of 500 stamps be utilized during the initial 3- to 6-week period, followed by regular monthly payments of 250 stamps/month. This is equivalent (for the stamps used in case study) to an annual expenditure per family, for the first year, of $13.81, followed by $7.50 for each succeeding year.

DATA MANAGEMENT CONSIDERATIONS AND INTEGRATION OF A MONITORING STUDY WITHIN A CONTINUING TRANSPORTATION STUDY

A continuous monitoring scheme generates a rather terrifying set of data management problems. The author's experience with a sample of only 104 households suggests strongly that any expansion of the diary technique to incorporate 500 to 1000 household units should be accompanied by a relatively sophisticated data-handling capability. A detailed discussion of the necessary data management format would take more space than is available here. However, a number of general comments may be made.

The data-management system may take one of a number of different forms. It should, however, consist of the following: (a) a data bank, in which a continuing set of basic travel information is retained on a permanent basis; (b) a complementary environmental record containing details of the occurrence of specific environmental shocks; (c) a sequential analysis set, in which running "averages" are maintained of a set of pertinent travel statistics (e.g., daily trip-making variances, visitation rates, etc.).

The nucleus of the system would be the travel data bank. This would contain details of each person- and group-trip made by each panel household, summarized in terms of the trip's origin and destination point, mode, start and end time, and trip purpose (the latter information being coded according to a two-digit classification). Permanent base data on the characteristics of the sample families would also be retained in the same source, with the environmental and sequential-analysis records aligned in concert with the structure of the main data bank.

The previous remarks are intended to be no more than the most cursory outline of an extremely complex operation. They serve merely to underline the essential need for an adequate data-management capability. This requirement in turn leads directly into a consideration of the integration of the proposed monitoring scheme within the framework of a continuing transportation study.

It may be validly assumed that any major transportation planning agency will already possess, possibly in reduced form, an ability to resolve large quantities of basic travel data. Integration of the monitoring scheme within the framework of such a study, therefore, is likely to reduce considerably the initial problems of data processing. More importantly, such an integration is both logical and also of obvious utility to the concept of a continuing planning operation. In the case of an established transportation study, the monitoring operation may be viewed as a device to measure the temporal transformation of travel demand away from that observed for the base year. Equally, the development of a longitudinal data set may provide a means for verifying, and if necessary modifying, the basic set of predictive relationships.

In the case of a newly initiated study, yet to collect its basic data set, the concept of an extended monitoring scheme is even more intriguing. Here there is a chance to modify and experiment with the established dogma of data collection. One may conceive of a variety of possible sampling formats: a spatially diffuse, randomized, cross-section sample based on a 2-, 5- or 7-day trip-record and a sampling ratio of only 3 percent to 10 percent rather than the usual 5 percent to 20 percent; a conventional one-day home-interview sample, supplemented by spatial clusters of 25 to 30 panel households; a combination of an original 7 percent one-day, base study; a regular 6-monthly or annual random 3-day sample (at a rate of, say, $\frac{1}{4}$ percent for a city of 400,000); and a continuing 500 household panel analysis. The possible permutations are virtually endless.

In each case, the initial extended base study would provide both a more accurate estimate of an individual household's (or zone's) average daily trip-production and also of the necessary spatial distribution of trip linkages required as input into the initial planning process (3). The extended monitoring and supplemental surveys provide a measure of the changing pattern of travel demand and of the stability of the predictive estimates developed from the base data. Together, the two data sets represent a considerably more complete and meaningful information package than that obtained from a conventional cross-sectional survey. The additional costs of the longitudinal surveys, though high, would be defrayed in part by the lower expenditures necessary for the reduced-sample base studies and, hopefully, by an increased forecasting accuracy.

The preceding discussion has been couched deliberately in very general terms. No mention has been made of specific sampling rates, sample sizes or design schedules. Obviously, a useful discussion in these terms may be presented only in the context of a specific design situation—a city of 400,000 population, a metropolitan area of 3,000,000 —and a given set of data requirements. General statements regarding desirable sample sizes and sample combinations, separated from a specific design context, are meaningless. The final section of this paper, therefore, is devoted to an example of a possible monitoring scheme applied to the limited case of the Village of Skokie, Illinois.

TABLE 1

ESTIMATED ANNUAL COSTS FOR PROPOSED SKOKIE STUDY (FIRST YEAR)

Initial Contact Costs (assumed acceptance rate = 50%)	
Drawing and validating sample	$ 500.00
Interviewer time, at $2.00/hr (including travel cost)	1,130.00
Preparation and printing of survey forms	700.00
Mail, telephone and miscellaneous	50.00
Secretarial time	800.00
Total cost	$ 3,180.00
Avg. cost/accepting household	$ 14.42
Sustained Surveillance Costs (assumed "medium level surveillance" for first 4 weeks, 2 calls/month thereafter)	
Interviewer time, at $2.00/hr (including travel cost)	$10,740.00
Office audit, coding and punching of data	28,760.00
Total cost	$39,500.00
Avg. cost/household/day	$ 3.60
Reimbursement Costs (assumed repayment rate 50 stamps/day for 4 weeks + 500 bonus and 250/month thereafter)	
Total reimbursement cost	$ 2,552.00
Avg. cost/household	$ 11.60
Repetitive Random Sample (assumed acceptance rate = 75%)	
Drawing and validating sample	$ 500.00
Interviewer time, at $2.00/hr (including travel cost)	2,950.00
Preparation and printing of survey forms	100.00
Mail, telephone and secretarial time	500.00
Office audit, coding and punching of data	1,200.00
Reimbursement cost (at 50 stamps/day)	166.00
Total cost	$ 5,416.00
Avg. cost/household/day	$ 4.10
Total annual costs	$50,648.00
Annual cost/household/day	$ 4.11

TABLE 2

ESTIMATED ANNUAL COSTS FOR PROPOSED SKOKIE STUDY (SECOND YEAR)

Sustained Surveillance Costs (assumed 2 calls/month)	
Interviewer time	$ 8,540.00
Office audit, coding and punching data	10,560.00
Total cost	$19,100.00
Avg. cost/household/day	$ 3.62
Reimbursement Costs (assumed repayment rate of 250 stamps/month)	
Total reimbursement cost	$ 1,650.00
Avg. cost/household	$ 7.50
Repetitive Random Sample (as for first year)	
Total cost	$ 5,416.00
Avg. cost/household/day	$ 4.10
Sample Renewal Costs (estimated sample decay 25 percent per annum)	
Maintaining and updating sample frame	$ 250.00
Interviewer time	300.00
Mail, telephone and miscellaneous	25.00
Total cost	$ 75.00
Avg. cost/accepting household	$ 10.42
Miscellaneous	
Preparation and printing of survey forms	$ 200.00
Secretarial time	200.00
Total cost	$ 400.00
Total annual cost	$27,141.00
Annual cost/household/day	$ 4.11

A POSSIBLE MONITORING STUDY FOR SKOKIE, ILLINOIS

Skokie lies 14 miles northwest of the Chicago Loop. It has a current population of approximately 68,000, divided into some 22,000 households. An 11 percent cross-sectional home-interview sample, drawn by the Chicago Area Transportation Study in 1964, yielded 2,101 separate dwelling units. The cost of the data collection and coding operations associated with this conventional, one-day home-interview study was roughly $25,000.00 (3).

The following longitudinal format is proposed as a supplement to the 1964 CATS survey:

1. Continuing Panel Study

Sampling ratio	1 percent
Sample size	220 households
Operational format:	
Initial survey period	4 weeks
Continuing survey rate	2 days/month
Surveillance level:	
Medium	(biweekly housecalls, see Appendix)
Repayment level:	
Initial period	50 stamps/day + 500 bonus
Continuing period	250 stamps/month
Assumed sample decay rate	25 percent per year

2. Repetitive Random Sample

Sampling cycle	3 months
Sampling ratio	½ percent
Annual sample size	440 households
Survey period	3 consecutive weekdays
Repayment level	50 stamps/day

In each case, the survey mechanism is the self-administered travel diary, supplemented by semiweekly monitoring calls in the initial phase of the panel study, and in-

TABLE 3

ESTIMATED ANNUAL COSTS FOR A 2000 ELEMENT MONITORING SCHEME (FIRST YEAR)

Initial Contact Costs (assumed acceptance rate = 50%)	
Drawing and validating sample	$ 1,500.00
Interviewer time, at $2.00/hr (including travel cost)	8,500.00
Preparation and printing of survey forms	3,500.00
Mail, telephone and miscellaneous	250.00
Secretarial time	3,000.00
Total cost	$ 16,750.00
Avg. cost/accepting household	$ 8.38
Sustained Surveillance Costs (assumed "medium level surveillance" for first 6 weeks, 2 calls/month thereafter)	
Interviewer time, at $2.00/hr (including travel cost)	$172,000.00
Office audit, coding and punching of data	250,000.00
Total cost	$422,000.00
Avg. cost/household/day	$ 3.29
Reimbursement Costs (assumed repayment rate 50 stamps/day for 6 weeks + 500 bonus and 250/month thereafter)	
Total reimbursement cost	$ 27,620.00
Avg. cost/household	13.81
Total annual survey costs	$466,370.00
Avg. annual cost/household/day	$ 3.64

TABLE 4

ESTIMATED ANNUAL COSTS FOR A 2000 ELEMENT MONITORING SCHEME (SECOND YEAR)

Sustained Surveillance Costs (assumed 2 calls/month)	
Interviewer time	$ 46,000.00
Office audit, coding and punching data	128,000.00
Total cost	$174,000.00
Avg. cost/household/day	$ 3.60
Reimbursement Costs (assumed repayment rate of 250 stamps/month)	
Total reimbursement cost	$ 15,000.00
Avg. cost/household	$ 7.50
Sample Renewal Costs (estimated sample decay 25 percent per annum)	
Maintaining and updating sample frame	$ 750.00
Interviewer time	2,060.00
Mail, telephone and miscellaneous	75.00
Total cost	$ 2,885.00
Avg. cost/accepting household	$ 5.75
Miscellaneous	
Preparation and printing of survey forms	$ 2,500.00
Secretarial time	3,000.00
Total cost	$ 5,500.00
Total annual cost	$197,385.00
Avg. annual cost/household/day	$ 4.11

dividual house calls for each continuing panel survey contact and also for each random sample selection (for the latter, personal visits are made both at the start and end of each 3-day period). The sampling frame for both surveys is the 1964 CATS home-interview study. An annual decay rate of 25 percent is assumed within the panel sample.

The approximate annual costs of the two surveys, exclusive of data analysis and the salaries of professional personnel, are given in Tables 1 and 2. Table 1 refers to the first year of operation and Table 2 to all subsequent years. The unit survey costs in each case are based on those observed for the case study described in the Appendix.

Note the relatively high annual expenditure involved even for a small-scale study such as that illustrated here—$50,648.00 in the first year and $27,141.00 in each subsequent year. These figures are exclusive of any data analysis costs or professional salaries. Estimation of a single, lump sum of $40,000.00 to cover these two items (probably a little conservative in terms of computer costs) yields estimated total annual costs for the first and second years of slightly more than $90,000.00 and slightly more than $67,000.00 respectively.

These figures do, of course, represent a somewhat artificial situation. It is unlikely that a full-scale monitoring scheme would ever be seriously considered for an area the size of Skokie. Much more probable is the application of the method to an entire metropolitan area such as Chicago.

In this case, one may argue that a much larger sample, say, 2000 households would be required. Tables 3 and 4 summarize the estimated costs of an expanded panel study (no repetitive random sampling) applied to such a sample. In this case, allowing for an estimate of $100,000.00 for the necessary computer capability (again a rather conservative figure), the total survey expenditure during the initial 12-month period comes to $566,370.00 or an average figure of $283.14 per household. For subsequent years, allowing for an average annual sample decay rate of 25 percent and a similar computer charge, the equivalent figures are $299,450.00 and $149.73.

Appendix

THE HOUSEHOLD PANEL: A CASE STUDY OF TRAVEL BEHAVIOR

From November 1965 to March 1966, a case study was performed of the daily travel of a sample of families residing in Skokie, Illinois. Skokie is a suburban dormitory community of approximately 68,000 people (Fig. 1). The population is predominately middle class, more than two-thirds of the households living in single family dwelling units and slightly less than one-third in flats, apartments or townhouses. The study was designed both to test a set of hypotheses concerning the temporal patterns of household travel and also to evaluate the suitability of the personal travel diary as a survey mechanism.

A total of 104 families were requested to maintain a record of their daily trip-making for a period of four weeks. Information was requested on the number of people involved, the points of origin and destination, start and end times, travel mode and trip purpose for all trips made by the household during the duration of the survey.

The survey mechanism in each case was the self-administered travel diary. Three different diary formats were tested, two designed for completion by the individual and one by the entire household. Figure 2 illustrates format No. 3 (household diary), which yielded the most successful results. Note the open-ended question on trip purpose.

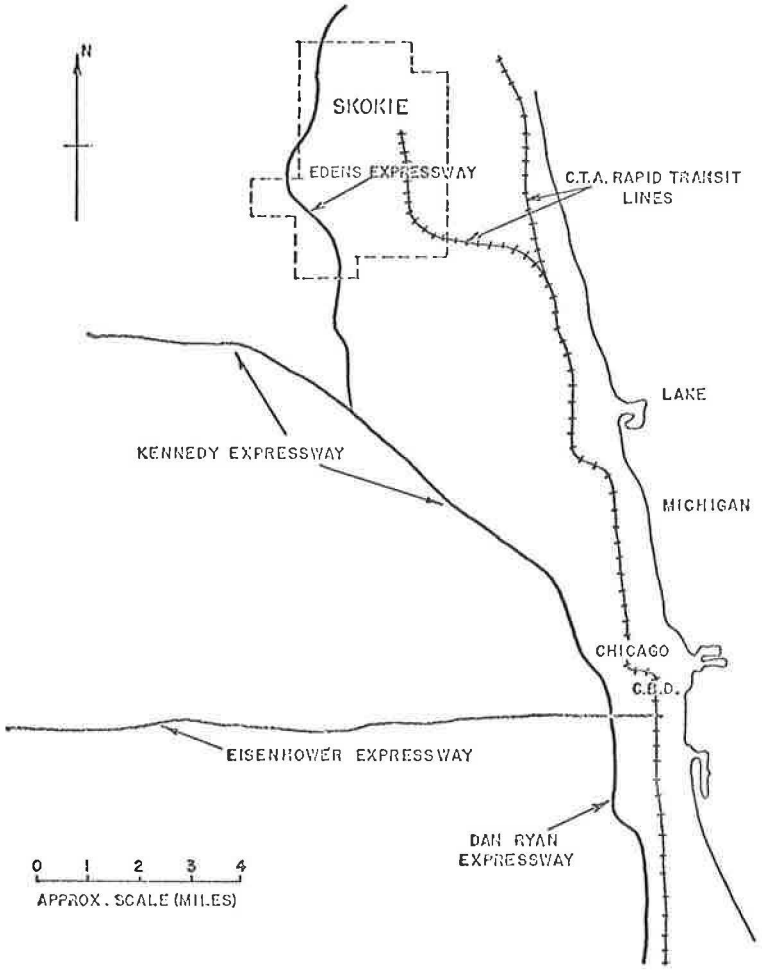

Figure 1. Location of Village of Skokie, Illinois.

HOUSEHOLD TRAVEL SURVEY

Date and Day of Week _ _ _ _ _ _ _ _ _ _ _ _ _ _

Trip No.	PERSONS TRAVELING		ORIGIN & DESTINATION OF TRIP		TRIP TIMES		REASON FOR MAKING TRIP
	Name	Method of Travel	From Name & Address of Origin	To Name & Address of Destination	Start	End	(Please give full details of why each trip was made
1	Richard	Car Driver	Home, 1570 Oak Avenue, Evanston, Illinois	Work, Tech. Inst., Sheridan Road, Evanston. (Parked in Tech lot)	8:30 a.m.	8:45 a.m.	Go to Work
2	Sandra Mandy	Walk, Bus Walk, Bus	Home, 1570 Oak Avenue, Evanston, Illinois	Old Orchard Shopping Center, Skokie (M. Fields)	9:45 a.m.	10:15 a.m.	Buy Mandy a new dress. Marshall Fields and Sachs
3	Sandra Mandy	Bus, Walk Bus, Walk	Old Orchard	The Huddle Resteurant, Orrington Hotel, Evanston	11:50 a.m.	12:30 p.m.	Meet Richard for lunch
4	Richard	Car Driver	Work	The Huddle Restaurant, Orrington Hotel, Evanston (Parked on the street)	12:20 p.m.	12:30 p.m.	Meet family for lunch
5	Richard Sandra Mandy	Car Driver Car Passenger Car Passenger	The Huddle	Home, 1570 Oak Avenue Parked in Apartment Lot	1:00 p.m.	1:10 p.m.	Take family home
6	Richard	Walk, El-subway	Home	State Highway Dept. Marina City, N. State Street Chicago (Davis St. & Randolph	1:15 p.m.	2:00 p.m.	Business Call
7	Richard	Walk, El-Subway,	State Highway Dept.	Jewell Store, Davis St., Evanston	5:00 p.m.	5:35 p.m.	Buy groceries on the way home
8	Richard	Walk	Jewell Store, Davis St., Evanston	Home	5:40 p.m.	5:45 p.m.	Return Home

Figure 2. Typical sheet from household travel diary (diary format No. 3).

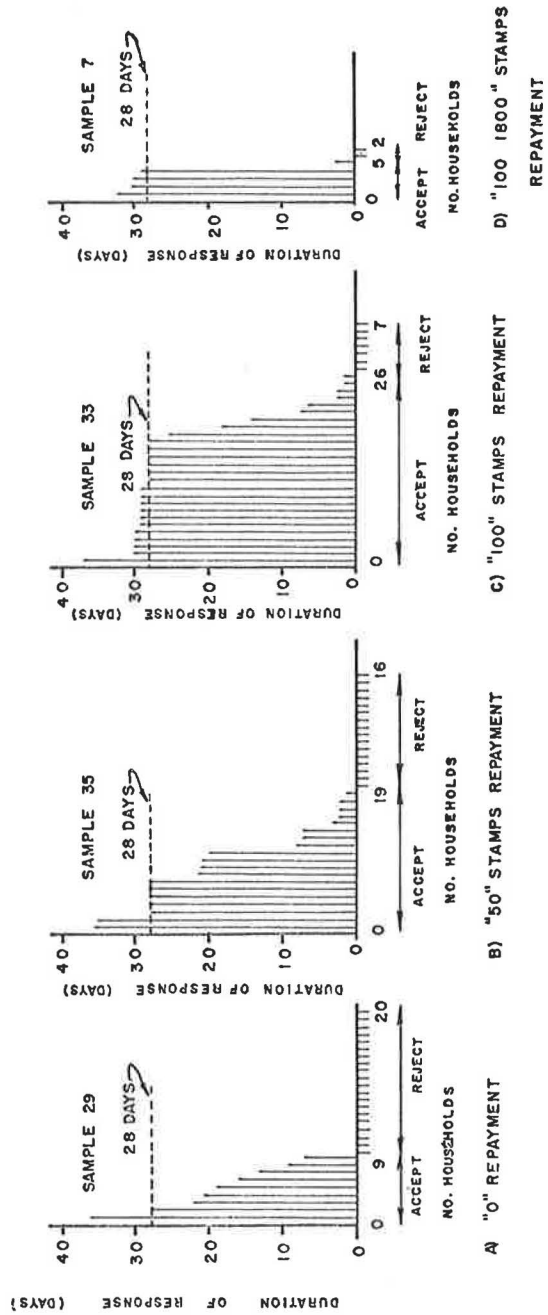

Figure 3. Distribution of sustained response rates by repayment level.

To evaluate respondent sensitivity to reimbursement and supervision, the total sample was subdivided into four repayment categories and three different surveillance groupings. Reimbursement in each case took the form of trading stamps redeemable for items of the respondent's choice at a local store. Surveillance techniques varied from regular house calls by the survey staff, to telephone contacts, logical consistency checks and the application of a simple recall questionnaire. Supplementary data on household structure, personal travel preferences and regular activity patterns were obtained through separate questionnaires.

RESULTS

Figure 3 shows the distribution of initial response to the request to maintain a travel diary. Out of the 104 households contacted, 45 rejected the request outright. Of those accepting, 30 maintained a comprehensive diary for the desired period of 28 days; the remainder kept a record for periods varying from 2 to 26 days.

The main factor influencing initial response was the prospective level of reimbursement. Once a family had agreed to cooperate, the level of information obtained was dependent primarily on the degree of surveillance provided by the survey staff. No significant bias was apparent in the socioeconomic structure of the families accepting and rejecting the initial survey contact. Simiarly, socioeconomic characteristics appeared to have little or no bearing on the maintenance of a sustained response.

It was found that as a minimum each household had to be visited by a member of the survey staff once every 3 to 4 days. At each visit the interviewer collected and checked the diaries for the previous 3 to 4 days, and also provided the family with a fresh set of blank forms. At the same time an extra series of predesigned questions were directed at the respondents to determine whether any trips had been inadvertently omitted from the daily records.

On receipt of the completed diaries in the survey office, a further, more comprehensive consistency check was performed, including both a repeat of the checks performed by the interviewer, and also an audit to determine possible errors of omission (i.e., a total failure to report certain kinds of trips), and, as the survey progressed, a check on the variance of the family's behavior from one week to the next. A detailed specification of the office monitoring procedure is given in Table 5. If any major errors or ambiguities were identified as a result of these office checks, the interviewer was instructed to clarify them during his next house call. As a supplement to the house calls, contact was also maintained daily with each family by telephone.

Almost all the cooperating families, after a period of approximately one hour's coaching and 2 or 3 days experience with the use of the diary format, showed themselves to be capable of providing an accurate and concise record of their day-to-day travel, provided that they were constantly reminded of the need to do so.

The commonest recording deficiencies (based on the result of the audit procedures) were definition of destination points (this occurred in approximately 6 percent of all ultimately recorded trips), specification of trip purpose (10 percent of all cases), and a failure to record approximately 8 percent of trips back to the home. Errors of total omission (i.e., the initial failure to record a trip in any form) varied considerably from family to family. Overall, the average rate was slightly less than 5 percent of all ultimately repeated trips.

The use of a self-administered questionnaire raises serious problems of possible reactivity (i.e., the influencing of behavior due to the fact that it is under study). This suggests that the data obtained from such a study may not represent a valid statement of "free" or unconstrained behavior.

Although accepting the validity of this comment, the writer feels that its importance should not be overplayed. People are essentially creatures of habit in their daily trip-making. Further, they exhibit a strong predilection for the pursuit of apparent irrationalities—they do not always choose the shortest or quickest route, they do not minimize their total travel effort nor do they usually maximize their expected net benefits. Rather, they tend to pursue a pattern of behavior which satisfies their demands and does not create too high a level of inconvenience.

TABLE 5

OUTLINE OF OFFICE MONITORING PROCEDURE

General
Check for complete diary set by day and by person
Individual Trip Records—Inclusion of Data
Check for inclusion of traveler names
Check for complete home-home round trips for each person
Check for ambiguity of person trip participation
Check for method of travel, by person and by trip
Check origin and destination specification
Check for successive origin and destination links
Check trip-start times, trip-end times
Check trip-time ambiguities
Check for trip purpose details
Note any ambiguities and illegible statements
Individual Trip Records—Completeness
Construct daily travel diagram, point to point, by person and by group; check for consistency
Construct trip-timing diagram, by purpose and by person; check for consistency
Cross-check personal and household travel records (where applicable), for group-trip participation
Compute running-average, daily trip-making rates, for 9 major trip-purposes and for subset of minor purposes; check for consistency and deviations from average trip-making frequency; check for work trips (one or more/worker/day), school trips (one or more/child/day), shopping trips and social-recreational trips

The maintenance of a travel diary, if it in fact has any reactive effect, probably tends to simply underline the more blatant irrationalities of behavior. The diarist, for example, may note that consolidation of several independent trips into a single "linked" trip might reduce his total travel effort. Alternatively, he may note that increased efficiencies would result from a change in the timing of his trips or perhaps from the use of another travel mode. In all events, any significant reactive consequence of diary maintenance should be reflected in a significant change in the pattern of daily travel over the period of diary keeping.

In no case was such a change observed—in terms either of an increase or decrease in total trip-making, a displacement of trip timing or a transference between travel modes—for any one of the sample households included in this study. This is not to say categorically, of course, that no reactivity occurred. The diary mechanism may simply have been too crude to detect any such effects, or else the period of the survey may have been too short for them to become apparent.

REFERENCES

1. Davis, J. A. Panel Analysis: Techniques and Concepts in the Interpretation of Repeated Measurements. Univ. of Chicago, 1963.
2. Kish, L. Survey Sampling. John Wiley and Sons, New York, 1965, pp. 148-150.
3. Garrison, W. L., and Worrall, R. D. Monitoring Urban Travel. Report submitted to NCHRP, Sept. 1966.
4. Yates, R. Sampling Methods for Censuses and Surveys, 3rd Ed. Charles Griffin and Co., London, 1960.
5. Hansen, M. H., Hurwitz, W. N., and Madow, W. G. Sample Survey Methods and Theory. Vol. 1 and 2, John Wiley and Sons, New York, 1953.
6. Sobol, M. G. Panel Mortality and Panel Bias. Jour. Amer. Statist. Assoc., Vol. 54, pp. 52-68, 1959.
7. Sharp, H., and Feldt, A. Some Factors in a Probability Sample Survey of a Metropolitan Community. Amer. Soc. Rev., Vol. 24, pp. 650-661, 1959.