Survey Methodology, December 1991
Vol. 17, No. 2, pp. 183–194
Statistics Canada

Sampling Flows of Mobile Human Populations

GRAHAM KALTON[1]

ABSTRACT

Surveys are often conducted of flows of persons, such as: visitors to museums, libraries and parks; voters; shoppers; hospital outpatients; tourists; international travellers; and car occupants. The sample designs for such surveys usually involve sampling in time and space. Methods for sampling flows of human populations are reviewed and illustrated.

KEY WORDS: Mobile populations; Exit polls; Traffic surveys; Time and space sampling; Systematic sampling.

1. INTRODUCTION

Most surveys of human populations are household based, typically with a sample of households selected with a multi-stage sample design, and individuals sampled within the selected households. The household survey is a powerful method for collecting data on a wide range of characteristics about the population, such as social, demographic, economic and health characteristics and the population's opinions and attitudes. The method is, however, not so effective for studying the characteristics of mobile populations. Two types of mobile populations may be distinguished: those who do not reside regularly at a fixed location, such as nomads and the homeless; and members of the general population who belong to the mobile population under study because they are in transit, such as visitors to libraries and parks, voters at polling booths, shoppers, hospital outpatients, travellers, and car occupants. This paper reviews sample design issues for this latter type of mobile population.

Although there are many surveys concerned with flows of mobile human populations, the general sampling literature contains little discussion of the sampling issues involved. The purpose of this paper is to describe the sample designs commonly adopted for surveys of flows of human populations, to discuss some of the special sampling issues faced, and to illustrate the range of applications for such surveys. The next section of the paper reviews the general time and space sample design used for sampling persons in transit and some of the issues involved in employing this design in particular situations. Section 3 then illustrates the application of the design in a range of different settings. Section 4 presents some concluding remarks.

2. SAMPLING IN TIME AND SPACE

It will be useful to consider a specific example in describing the general time and space sample design for sampling flows of human populations. Suppose that a survey of visitors to a summer sculpture exhibition in a city park is to be conducted to find out the visitors' socio-economic characteristics, how they heard about the exhibition, what means of transport they used to get to the park, and perhaps their views of the exhibition. Suppose that the exhibition is held from April 1 to September 30 in the year in question, that it is open from 10 a.m. until 6 p.m. daily, and that there are three sites where visitors enter and leave the exhibition.

[1] Graham Kalton, Survey Research Center, University of Michigan, Ann Arbor, Michigan 48106-1248, U.S.A.

The sampling frame for a survey of this type is usually taken to be a list of time interval/site primary sampling units (PSUs). This frame is constructed by dividing the time period of the survey into a set of time intervals for each site. A simple construction of PSUs for the current example would be to divide each exhibition day at each site into two time intervals, one from 10 a.m. until 2 p.m. and the other from 2 p.m. until 6 p.m. A more complex construction of PSUs could involve time intervals of different lengths on different days and/or at different sites. Once the PSUs are defined, a two-stage sample design is often employed. At the first stage a sample of PSUs is selected, and at the second stage a sample of visitors is drawn, usually by systematic sampling, in the sampled PSUs.

The actual specification of the sample design for a survey of persons in transit within the two-stage sampling framework depends on features of the mobile population under study and of the survey data collection procedures. A key feature is the nature of the flow of the mobile population. In particular, is there a predictable variability in the rate of flow across PSUs? For instance, is the flow at one site higher than that at another site, or are the flows at some time intervals (say, Saturday afternoons) higher than those at others? Also, is the flow within a PSU a smooth one throughout the time interval or is it uneven, with visitors arriving (or leaving) in sizeable groups? Both these aspects of flow affect the sample design for the survey.

If the flow is fairly uniform across the PSUs, and if the PSU time intervals are the same, then the number of visitors per PSU is approximately constant. In this case, the PSUs may be sampled with equal probabilities, and a constant subsampling fraction can be applied within the selected PSUs to generate an equal probability, or epsem, sample of visits. The PSUs can be classified in two or more dimensions (*e.g.* day of week, time of day, and site), and a carefully balanced sample across these dimensions can be obtained using lattice sampling (Yates 1981; Cochran 1977 and Jessen 1978).

In many cases, the level of flow varies across the PSUs in a manner that is partly predictable. For instance, the attendance at the sculpture exhibition may be known to be generally higher in the later shift each day and at the weekends, and particularly low on Mondays. Thus the PSUs comprise different numbers of visitors, that is, they are PSUs of unequal sizes. The usual procedure for handling PSUs of unequal sizes is to sample them with probabilities proportional to their sizes (PPS), or estimated sizes (PPES). In the current context, the actual PSU sizes are not known in advance, and estimated sizes must therefore be used. Sampling the PSUs with PPES works well provided that reasonable estimates of the sizes can be made. When PSUs are selected by PPES sampling, then the application within the selected PSUs of subsampling fractions that are inversely proportional to the estimated sizes of the PSUs produces an overall epsem sample of visits. In general, an attraction of PPES sampling (with reasonable estimates of size) is that the subsample sizes in the PSUs do not vary greatly from one PSU to another. This feature is of especial value for conducting the fieldwork in surveys of persons in transit. When time/site PSUs are sampled by PPES sampling, lattice sampling cannot be applied for deep stratification. Instead, controlled selection may be employed for this purpose (Goodman and Kish 1950; Hess *et al.* 1975).

An important consideration in any two-stage sample design is the allocation of the sample between first-and second-stage units, that is, how many PSUs to select and how many elements to select per sampled PSU. In the case of surveys of persons in transit, that allocation is strongly affected by the fieldwork procedures to be used and the nature of the flow within the PSUs. The aim of the design is to make full use of the fieldworkers assigned to a sampled PSU while maintaining a probability sample of persons entering (or leaving) the site during the sampled time interval.

Many surveys of persons in transit use self-completion questionnaires, in which case the fieldwork process for the two-stage design described above consists of counting persons as they enter (or leave) the sampled site during the time interval, selecting every kth person for a systematic sample, and asking the selected persons to complete the questionnaire. If the flow is light and evenly spread throughout the time interval, one fieldworker may be able to handle all the tasks involved. When this is so, the sampling interval k can be chosen to give the fieldworker time to perform all the tasks in an unpressured way. If, however, the flow is heavy, either constantly or intermittently, two fieldworkers may be needed, one simply to count entrants (or leavers) and identify sampled persons, and the second to hand out the questionnaires and to instruct respondents on how they should be completed and returned. With this fieldwork arrangement, the sampling interval can be chosen to keep the second fieldworker as fully occupied as possible, while making sure that he or she is able to distribute questionnaires to all (or at least nearly all) of those sampled. Nonresponse can be a major concern with the self-completion mode of data collection. It is often possible to keep nonresponse to an acceptable level when sampled persons complete and return the questionnaire at the site. However, when they are handed the questionnaire with the request to complete it later and return it by mail, the level of nonresponse can be very high and, moreover, there is generally no way of following up the nonrespondents.

When face-to-face interviewing is used for data collection, the fieldwork team for a PSU usually contains one counter and a small team of interviewers. The size of the interviewer team depends on the regularity of the flow and the length of the interview. Since persons in transit are likely to be unwilling to be delayed for long, interviews are necessarily mostly short. Longer interviews may, however, be possible if the sampled persons are in a waiting mode, such as waiting in line or in an airport departure lounge. The choice of sampling interval has to be such that there is always (or nearly always) an interviewer free to interview the next sampled person, and that the interviewers do not spend too much time waiting for the next sampled person to be selected. If the flow is irregular, allowance needs to be made to accommodate the peaks (for instance, the arrival of a coachload of visitors to the sculpture exhibition).

The PPES selection of the PSUs works to equate the subsample size for each sampled PSU. For face-to-face interview surveys, the interviewer load is thus roughly the same for each selected PSU, and hence the same-sized interviewer team can be used for each PSU. A problem occurs, however, when the PPES measure used in selecting the PSU at the first stage is seriously in error. For example, a thunderstorm may substantially reduce the number of visitors to the sculpture exhibition on a particular Saturday afternoon, or an unforeseen holiday may substantially increase the number on another day. In the first case, applying in that PSU a sampling interval inversely proportional to its estimated size will leave the interviewers largely unoccupied, whereas in the second case it will result in a workload that the interviewers cannot handle. A modification that may be adopted in such cases is to change the sampling interval at the start of data collection to one that is more suitable for the flow actually encountered. Since this modification destroys the epsem property of the sample, weights are needed in the survey analysis.

A general limitation to the systematic sampling of visitors at selected PSUs is that if the sampling interval is made long enough to enable interviewers to cope with peak flows, they spend much of their time without work. On the other hand, if the sampling interval is reduced, the interviewers are more fully occupied, but they cannot cope with peak flows. Various methods have been proposed to circumvent these problems (Heady 1985). One procedure is to take a systematic sample of times (say, every 10 minutes) and to select the next visitor to enter after each sampled time. This procedure might have fieldwork attractions, but it does

not produce a probability sample of visitors. Persons arriving in busy periods are less likely to be chosen, as are those who travel in groups, and the walking habits of persons travelling in groups may affect the chances of selection in unknown ways. The sample generated by this procedure is clearly not an epsem sample. An attempt can be made to compensate for the selection bias that operates against visitors arriving in busy times by dividing the time interval for selected PSUs into a set of much shorter intervals, and keeping a log of arrivals in each such interval. Then weighting adjustments can be employed to compensate for the variation in the flow across the shorter intervals.

Another alternative procedure to systematic sampling of visitors is to take the next person to enter (or leave) after the last interview was completed. With this procedure, the first persons to arrive after gaps in the flow, perhaps the leaders of groups, clearly have greater chances of selection. Also interviewers may deliberately speed up or slow down their current interview in order to avoid or to select a particular individual. For these reasons, variants on this procedure that select the nth person after the completed interview, where n might be set at 2, 3, 4 or 5, have been employed. These alternatives to straightforward systematic sampling of visitors make more effective use of interviewers' time, and hence enable larger samples to be obtained for a given fieldwork budget. However, they produce nonprobability samples, with the risk of selection bias that this form of sampling entails. Probability sampling provides the security of objective statistical inference without the need for assumptions about the sample selection process. With nonprobability sampling, assumptions need to be made about the way the sample was generated, a common assumption being that all the elements in the population have an equal chance of selection. Failure of the assumptions can lead to serious bias in the survey estimates.

Visitors may be sampled either as they enter or as they leave a location. If data about the visitors' activities in and opinions of the location are required, then leavers need to be sampled. In other cases, the choice between sampling entrants and leavers may depend on the nature of the flows. It may, for example, be difficult to sample and interview people leaving a theatre because they leave en masse and because they will not want to be delayed. On the other hand, they may be readily sampled and interviewed as they line up to enter the theatre.

In concluding this section, attention should be drawn to the fact that the samples described here are samples of visits not visitors. The standard two-stage design may produce an epsem sample of visits, but this is not the same as an epsem sample of visitors unless each visitor visits the place under study (the sculpture exhibition) only once (or they all visit the same number of times). For most flow surveys, the visit, rather than the visitor, is the appropriate unit of analysis. There are, however, situations where the analytic unit is problematic. Using the visit as the unit of analysis, the researcher might readily accept visits to the sculpture exhibition on two separate days as distinct visits, but might not be willing to treat two entries on the same day (one, perhaps, after leaving briefly for refreshments) as two visits. The use of the visitor as the unit of analysis presents severe problems because of the issue of multiple visits, and the fact that visitors will not be able to report their multiplicities. They may be able to recall past visits reasonably well, but they will usually be unable to forecast future visits accurately.

3. SOME EXAMPLES

This section presents some examples of surveys of flows of human populations in order to indicate the wide range of applications and to illustrate some of the special considerations that arise in particular settings.

3.1 A Survey of Library Use

A survey of the use of the 18 libraries at the University of Michigan was conducted in 1984 (Heeringa 1985). Each sampled person exiting a library was asked whether he or she had used the library's materials and services during that visit. If so, the person was asked to complete a short self-completion questionnaire of seven questions on the materials and services used. Most of the 5,184 respondents completed the questionnaires on the spot and returned them to the survey fieldworkers; others sent them back by campus mail. A response rate of 96% was obtained.

The sample design followed the two-stage time/site sample design described in Section 2. The survey covered the full 1984 calendar year. Each day the libraries were open was divided into 10 two-hour time intervals, starting at 7.30 a.m. and lasting until 3.30 a.m. the next morning, the two-hour interval being chosen on the grounds that it was a suitable shift for the fieldworkers. The PSUs were then defined to be time interval/library combinations. The PSUs were selected by PPES sampling, where the estimated size for a PSU was the estimated number of persons exiting from that library in the specified time period. Rough estimates of these numbers were derived from average daily usage based on November, 1983, turnstile counts where available, and on librarians' estimates where not, and on an assumption that library exit volume was twice as high between 9.30 a.m. and 5.30 p.m. as at other times. The libraries were stratified into four types, and within each stratum controlled selection was employed to give a proportionate distribution of the sample across libraries, days of the week, and time intervals.

For each selected PSU, a systematic sample of persons exiting the library was selected for the survey, with the sampling interval being determined to yield an overall epsem sample of visits. Fieldworkers were provided with a record sheet of integers from 1 up to 430, with the selected numbers marked on them. All they then needed to do was check off a number for each person exiting the library, and select the persons associated with the sample numbers. An advantage of this scheme is that fractional sampling intervals are readily handled. Where the exit volume for a sampled PSU was expected to be low, one fieldworker was assigned to perform both the counting and the contacting of sampled persons. Where the exit volume was high, two fieldworkers were assigned, one to count and one to contact sampled persons. There was also a need for more than one fieldworker for libraries with more than one exit.

3.2 A Survey of Museum Visits

A face-to-face interview survey of visitors leaving the National Air and Space Museum in Washington, D.C. was conducted from mid-July until December, 1988 (Doering and Black 1989). The interview, which took about four to six minutes to complete, collected data on the sampled person's socio-demographic background, place of residence, activities on the visit, exhibits of special interest, reason for visit, the size and type of group if part of a group visit, and mode of transport used. Children under 12 years old and persons working at the museum were excluded from the survey. Data were collected from 5,574 respondents, with a response rate of 86%.

Each day in the survey period was divided into two half-days. Interviewing was conducted on one half-day every second day, alternating between mornings and afternoons. During the summer season, three public exits from the museum were in operation, while later in the year only two of them were open. During the selected half-days, survey data collection was rotated on an hourly basis between the exits that were open.

The fieldwork team for an exit at a sampled hour comprised one or two counters and two interviewers. The lead counter used a mechanical counter and a stop watch to keep track of the number of persons exiting, and to maintain a record that gave the numbers of persons exiting in each 10-minute interval in the hour. The lead counter also identified the persons to be interviewed. The selection of sample persons was made in order to keep the interviewers fully occupied. The lead counter noted when an interviewer had completed an interview and was ready to begin another one, and then chose the fifth person exiting after that time as the next sampled person. The 10-minute flow counts were used in the analysis to develop weights to compensate for the variation in the chance of selection associated with the variable flow of persons across time.

The distinction between the "visit" and the "visitor" is particularly salient for this survey. Persons could, of course, visit the museum on several days throughout the survey period, and also could visit the museum several times on a given day. This latter possibility is particularly likely with the National Air and Space Museum because entry to the Museum is free, and hence there is no incentive to enter only once. Given this situation, it may be appropriate to define multiple entries on one day as a single visit for some types of analysis. For some purposes, this definition could be applied by restricting the analysis to those exiting for the first time on the sampled day.

3.3 Exit Polls

A number of major news organizations conduct polls of voters on election days in the United States (Levy 1983; Mitofsky 1991). Voters are sampled as they leave polling places. Those selected are asked to complete a short and simple self-completion questionnaire, and to deposit the completed questionnaire in a ballot box. A typical questionnaire contains around 25 questions asking how the respondent voted, what the respondent's position is on key issues, what opinions the respondent has on various topics, and what are the respondent's demographic characteristics. Refusal rates for the CBS exit polls have averaged 25% for recent elections (Mitofsky and Waksberg 1989).

The sampling of voters for election polls usually employs a straightforward two-stage sample design. At the first stage a stratified PPES sample of voting precincts is drawn, where the size measure is the number of voters in the precinct. At the second stage a systematic sample of voters leaving the polling place is selected, with a sampling interval chosen to produce an approximately epsem sample of voters within states. Usually only one interviewer is assigned to each selected precinct. The fieldwork is straightforward when a polling place has a single exit, and the interviewer is permitted to get close to it. When there are two or more exits, interviewers alternate between the exits, covering each one for set periods of time. When this applies, the sampling interval has to be modified accordingly. In some states interviewers are not allowed to approach within a certain distance of a polling place, and this can create problems if it results in voters departing in different directions before the interviewer can contact them.

3.4 Ambulatory Medical Care Survey

The U.S. National Ambulatory Medical Care Survey (NAMCS) employs a flow survey design to collect data on visits to physicians' offices for physicians in office practice who direct patient care (Bryant and Shimizu 1988). The NAMCS has been conducted a number of times since it was introduced in 1973. For each survey, data collection has been spread throughout the survey's calendar year in order to provide annual estimates of visit characteristics. Individual sampled physicians have, however, been asked to provide information for a sample of their visits occurring in only one week. The annual coverage is achieved by asking different sampled physicians to report on different weeks of the year.

The sample for the NAMCS is based on a complex three-stage design, which has varied over time. A broad overview of the design will serve for present purposes; for more details, the reader is referred to Bryant and Shimizu (1988). The first stage of the NAMCS sample design is the selection of a stratified PPES sample of areal PSUs, selected with probability proportional to population size. At the second stage, physicians are sampled from lists within the selected PSUs with different sampling intervals from PSU to PSU to take account of the unequal selection probabilities for the PSUs (in the more recent surveys, different specialty classes are sampled at different rates). Sampled physicians are then assigned at random in a balanced way to one of the 52 reporting weeks of the year. Each physician is asked to record information for a systematic sample of his or her patient visits occurring during the sampled week, with the sampling interval being chosen to yield about 30 sampled visits in the week. A sampling interval of 1, 2, 3 or 5 is chosen for a particular physician on the basis of the number of office visits the physician expects during the week, and the number of days he or she expects to see patients. The fieldwork procedures consist of keeping a log of patient arrivals for sampling purposes, and then completing a short 16-item record for each sampled visit.

The NAMCS is a survey of patient visits not patients. As such, it provides useful information about the nature of physicians' work on a visit basis - the frequency of use of diagnostic tests, the therapies provided, and the demographic characteristics of the patients seen. It does not, however, provide estimates on a patient basis, such as treatments and outcomes for patients' episodes of illness.

3.5 Surveys of International Passengers

A number of countries conduct surveys of their international travellers, both those entering and those leaving the country by land, sea or air. This subsection will briefly describe the sample designs for a survey of international air passengers conducted by the United States, for surveys of international air and land travellers conducted by Canada, and for a survey of international air and sea passengers conducted by the U.K.

The United States Travel and Tourism Administration conducts an In-flight Survey of International Air Travelers to survey both foreign travellers to the U.S. and U.S. residents travelling abroad (see, for instance, United States Travel and Tourism Administration 1989). The survey is conducted through the voluntary cooperation of some thirty airlines. A stratified sample of scheduled flights is selected for the third week of each month and all passengers on those flights are included in the sample. Participating airlines are provided with a survey kit of instructions and questionnaires in appropriate languages for each sampled flight. The airline cabin personnel distribute the self-completion questionnaires in boarding areas or in flight to all adult passengers and collect them prior to debarkation. Nonresponse is a serious problem with these surveys. For the 1988 survey of visitors to the United States, one half of the flight kits issued resulted in no returned questionnaires. For flights for which questionnaires were returned, the estimated response rate for non-U.S. residents was 44% and for U.S. residents it was only 20%.

The International Travel Section of Statistics Canada conducts international travel surveys at both airports and landports in Canada. The surveys are undertaken in cooperation with Canada Customs, with customs officers being responsible for distributing the self-completion mail-back questionnaires. The account here is based on the report by the International Travel Section, Statistics Canada (1979). It reflects the survey designs that applied prior to some changes that have recently been made. The sample designs for the landports and airports have been similar, and therefore only the design for the landports will be outlined here.

At one time the sampling scheme at landports for returning Canadian residents who had spent at least one night abroad was to distribute survey questionnaires to every travel party

on every fourth day throughout the year, the days being chosen by systematic sampling. This scheme proved to be unworkable because the customs officers too often failed to apply it correctly. It was therefore replaced by a stint scheme in which a landport was assigned two periods, or stints, for each quarter of the year during which the questionnaires were to be distributed. The stints were expected to last from 6 to 10 days, with successive stints starting about 6½ weeks apart (Gough and Ghangurde 1977). The number of questionnaires sent to a landport for a particular stint was determined from the expected traffic at that port. The customs officers were then instructed to start the distribution of the questionnaires on a given day, and to continue to distribute them until none were left. This sample design is geared to operational limitations resulting from the use of customs officers, for whom the survey is of only secondary concern, as survey fieldworkers. The design has some major drawbacks, but perhaps a more serious concern is a response rate of 20% or less.

The U.S. and Canadian surveys of international travellers both rely on cooperation from other agencies in conducting the fieldwork. This cooperation has notable benefits in costs, but a price is paid in terms of a lack of ability to apply rigorous controls to the fieldwork procedures. The U.K. surveys of air and sea travellers employ more costly face-to-face interviewing procedures.

The 1984 U.K. International Passenger Survey included the three Heathrow terminals, Gatwick and Manchester airports as strata (Griffiths and Elliot 1987). Within each airport, days were divided into mornings and afternoons, and these periods constituted the PSUs. A stratified sample of PSUs was selected, and systematic samples of passengers were chosen in selected PSUs. A sample of PSUs for other airports was also included. Two alternative data collection procedures were used at seaports. At some seaports, interviewers sampled and interviewed passengers at the quayside. At others, the interviewers travelled on the ship, interviewing passengers during the voyage. In the former case, they worked shifts that covered several sailings, and the shift became the PSU. In the latter case, the crossings were the PSUs.

3.6 Surveys at Shopping Centers

Surveys conducted at shopping centers are of two types. One type aims to describe the shoppers' socio-economic characteristics, their areas of residence, and their shopping activities in the center. The other type uses the shopping center as a convenient location to obtain samples of people from the general population of the area.

An example of a survey of the first type is a study that was conducted to examine the impact of the opening of a hypermarket on the outskirts of the city of Southampton, England (Wood 1978). Surveys of shoppers were conducted in four neighboring shopping centers both before and after the hypermarket opened (and also at the hypermarket). At each center, the first step in the survey process was the enumeration of all the retail outlets and their hours of opening. The second step was a counting of departures of groups of shoppers from sampled shops at sampled hours, with counting being conducted for 15 minutes within the hour. The counting operation was carried out over a period of one month. Based on the counts obtained, interviews were allocated between shop types and days of the week, and to specific shops and hours. Interviewers were then instructed to interview the given number of people leaving the shop, interviewing the next person to leave after they had completed the previous interview. The sample is one of shop visits, and shoppers could visit several shops on a particular trip to the shopping center. Respondents were asked about previous visits to shops in the center on this particular trip, and also about the number of extra shops they planned to visit. These data were used to develop weights for analyses of trips.

The second type of shopping center survey uses the selected persons at shopping centers as a convenience sample of the general population. Mall intercept surveys of this type are widely used in market research (Bush and Hair 1985; Gates and Solomon 1982). The procedures are often haphazard, and the samples are potentially biased. The issues involved are reviewed by Sudman (1980), who discusses procedures for sampling shopping centers, locations at selected centers, and time periods to improve the sample designs, and by Blair (1983), Dupont (1987), and Murry *et al.* (1989).

3.7 Road Traffic Surveys

One form of road traffic survey relates to traffic passing through one or more locations. Time and space sample designs can be applied for these surveys in a relatively straightforward manner. Kish *et al.* (1961), for example, describe the sample design for an origin-destination survey of vehicles using the Port of New York Authority's bridges over and tunnels under the Hudson river during 1959. A four-stage stratified PPES sample design was used for this survey. The PSUs were combinations of eight-hour shifts and particular bridges or tunnels. A sample of these PSUs was selected at the first stage, a sample of contiguous toll lanes (locations) was selected at the second stage within selected PSUs, a sample of specific lanes was selected at the third stage within selected locations, and finally a systematic sample of vehicles was selected at selected lanes. Interviewers stayed at one sampled location for four hours, and moved each hour from one traffic lane to another according to a prescribed pattern.

Another type of road traffic survey relates to general traffic on the road. Surveys of occupants of passenger vehicles to study seat belt usage and drivers' blood alcohol concentrations are of this type. A full discussion of the complex design issues involved in such surveys is outside the present scope; instead only a few general observations will be made.

The method of data collection to be employed exerts a strong influence on the sampling procedures for a general traffic survey. Seat belt usage is mostly studied by observational methods, whereas the measurement of blood alcohol concentrations usually involves breathtesting. Shoulder belt usage of front-seat occupants can be observed in moving traffic, but lap belt usage and the seat belt usage of other occupants can be observed only when the vehicle has stopped briefly, for instance at traffic lights. Lack of street lights can preclude observation of seat belt usage at night at some sites. Breathtesting requires the vehicle to be stopped, and this can be done safely only in locations where the stopped vehicle does not hinder the other traffic. Unlike observational surveys, interview surveys that stop vehicles face a significant nonresponse problem.

An ingenious method of studying seat belt usage on interstate highways is described by Wells *et al.* (1990). For this study, an observer sat behind the driver in a passenger van that travelled at a slower speed than the prevailing traffic in the right hand lane of the highway. From that vantage point, the observer noted the shoulder belt usage of front-seat occupants of cars, light trucks, and vans that passed the observer's van in the adjacent lane.

A more usual approach to studying seat belt usage is to take observations at road intersections and freeway exits controlled by traffic lights, and sometimes at shopping centers and parking lots (Ziegler 1983; Bowman and Rounds 1989). O'Day and Wolfe (1984) describe an observational survey of seat belt use in Michigan applying this approach. They sampled a number of areal units, sampled a number of intersections with traffic signals within these areas, sampled days for observations to be taken at these intersections, and sampled five periods of one hour each between 8 a.m. and 8 p.m. for observation on each selected day. Each hour of observation was conducted at a different intersection. The hours were selected by a scheme that alternated one hour working and one hour free, with the observers moving between

intersections in the free hours. Observations of seat belt usage were taken at the selected intersections at the specified times for vehicles that stopped at the traffic lights. When more than one vehicle was stopped, observation began with the second vehicle, because of the bias associated with the first vehicle to stop at a light. In order to obtain more detailed information on the usage of child-restraints, observations were also made on vehicles entering shopping centers and rest areas.

The usual approach to analyzing observational data on seat belt usage is to calculate the proportion wearing seat belts among those observed. Brick and Lago (1988) propose an alternative measure, the proportion of estimated time front-seat occupants are belted in eligible vehicles to the total time in eligible vehicles. For their survey a probability sample of all roadway intersections, whether they had traffic signals or not, was selected. To avoid selection bias, observers were told the site they were to use for observation and the direction of the traffic to be observed in the specified 40-minute interval of observation. The time occupants were on the road was estimated as the length of the road segment leading to the intersection divided by the estimated average speed of the traffic on that segment. This estimated time was used as a weighting factor in the analysis.

The sampling considerations for roadside breathtesting surveys are broadly similar to those for seat belt usage surveys, except that the locations for data collection need to be places where vehicles can be stopped safely. In the 1986 U.S. National Roadside Breathtesting Survey, local police officers cooperated in the survey by flagging down selected drivers and directing them to the survey interviewers (Wolfe 1986). The interviews lasted about 5-6 minutes. When an interviewer finished an interview and the respondent had taken the breath test, the interviewer would signal to the police officer to stop the next passing vehicle. Interviewing was conducted for a period of two hours at each sampled location. A count was made of all the vehicles passing the location in the sampled direction during the period, and the ratio of this count to the number of interviews conducted was used as a weighting factor in the analysis.

4. CONCLUDING REMARKS

As the examples in the previous section illustrate, fieldwork considerations and the economics of data collection play major roles in the choice of sample design for surveys of persons in transit. The length of the time interval used in defining the PSUs may, for instance, be dictated by the length of a suitable workshift for the fieldworkers, and this may result in PSUs with substantial internal variation in the rate of flow. For example, in a survey of passengers arriving at a railway station, a morning interviewer workshift may include a peak flow of early morning commuters and a low rate of flow later on. If it were not for the need to make the PSU time interval conform to the fieldworkers' workshift, it would be preferable to avoid such variation in flow within PSUs since it leads to problems in how to subsample in the selected PSUs.

When the flow of persons within a PSU is uneven, the use of systematic sampling, or any epsem sampling scheme, for selecting persons creates a variable workload over time. If this variability in workload is substantial, there are difficulties in deciding how to staff the PSU for the survey fieldwork, particularly for a face-to-face interview survey. The assignment of sufficient staff to cope with peak flows is uneconomic since interviewers will then often be inactive at off-peak times. Sometimes staffing for somewhat below peak flow may be preferable. This will introduce some nonresponse at times of peak flow because no interviewer is available to conduct an interview with some sampled persons, but it will more fully use the interviewers' time.

The most effective use of the interviewers' time is to assign them to interview the first person to arrive (or leave) after they have completed their current interview. Schemes of this type suffer the disadvantage of not producing probability samples, and hence there is a risk of bias in the survey estimates. Where cost effective probability sampling designs can be devised, they are to be preferred. However, the choice of a sampling scheme in which the first (or second, or third) person is selected after an interviewer becomes free is understandably attractive for face-to-face interview surveys when the flow is very variable and unpredictable. When this kind of scheme is employed, it is useful to take counts of the flow over short intervals of time. These counts may then be used to make weighting adjustments to compensate for the unequal selection probabilities caused by the uneven flow.

ACKNOWLEDGEMENT

I would like to record my thanks to many researchers who generously provided me with information about the flow surveys with which they have been associated.

REFERENCES

BLAIR, E. (1983). Sampling issues in trade area maps drawn from shopper surveys. *Journal of Marketing*, 47, 98-106.

BOWMAN, B.L., and ROUNDS, D.A. (1989). Restraint System Usage in the Traffic Population, 1988 Annual Report. Washington, D.C.: U.S. Department of Transportation.

BRICK, M., and LAGO, J. (1988). The design and implementation of an observational safety belt use survey. *Journal of Safety Research*, 19, 87-98.

BRYANT, E., and SHIMIZU, I. (1988). *Sample Design, Sampling Variance, and Estimation Procedures for the National Ambulatory Medical Care Survey*. Vital and Health Statistics, Series 2, No. 108, Washington D.C.: U.S. Government Printing Office.

BUSH, A.J., and HAIR, J.F. (1985). An assessment of the mall intercept as a data collection method. *Journal of Marketing Research*, 22, 158-67.

COCHRAN, W.G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley.

DOERING, Z.D., and BLACK, K.J. (1989). Visits to the National Air and Space Museum (NASM): Demographic Characteristics. Working Paper 89-1, Institutional Studies, Smithsonian Institution.

DUPONT, T.D. (1987). Do frequent mall shoppers distort mall-intercept survey results? *Journal of Advertising Research*, 45-51.

GATES, R., and SOLOMON, P.J. (1982). Research using the mall intercept: State of the art. *Journal of Advertising Research*, 22, 4, 43-49.

GOODMAN, R., and KISH, L. (1950). Controlled selection – a technique in probability sampling. *Journal of the American Statistical Association*, 45, 350-372.

GOUGH, J.H., and GHANGURDE, P.D. (1977). Survey of Canadian residents returning by land. *Survey Methodology*, 3, 215-231.

GRIFFITHS, D., and ELLIOT, D. (1987). Sampling errors on the International Passenger Survey. Unpublished paper, Social Survey Division, U.K. Office of Population Censuses and Surveys, London.

HEADY, P. (1985). Note on some sampling methods for visitor surveys. *Survey Methodology Bulletin*. U.K. Office of Population Censuses and Surveys, 17, 10-17.

HEERINGA, S.G. (1985). The University of Michigan 1984 Library Cost Study: Final Report. Institute for Social Research, University of Michigan.

HESS, I., RIEDEL, D.C., and FITZPATRICK, T.B. (1975). *Probability Sampling of Hospitals and Patients*. Ann Arbor, Michigan: Health Administration Press.

JESSEN, R.J. (1978). *Statistical Survey Techniques*. New York: John Wiley.

KISH, L., LOVEJOY, W., and RACKOW, P. (1961). A multi-stage probability sample for traffic surveys. *Proceedings of the Social Statistics Section, American Statistical Association*, 227-230.

LEVY, M.R. (1983). The methodology and performance of election day polls. *Public Opinion Quarterly*, 47, 54-67.

MITOFSKY, W.J. (1991). A short history of exit polls. In *Polling in Presidential Election Coverage*. (Eds. P. Lavrakas and J. Holley). Newbury Park, California: Sage.

MITOFSKY, W.J., and WAKSBERG, J. (1989). CBS models for election night estimation. Paper presented at American Statistical Association San Diego Winter Conference.

MURRY, J.P., LASTOVICKA, J.L., and BHALLA, G. (1989). Demographic and lifestyle selection error in mall-intercept data. *Journal of Advertising Research*, 46-52.

O'DAY, J., and WOLFE, A.C. (1984). Seat Belt Observations in Michigan - August/September 1983. Ann Arbor, Michigan: University of Michigan Transportation Research Institute.

SUDMAN, S. (1980). Improving the quality of shopping center sampling. *Journal of Marketing Research*, 17, 423-31.

STATISTICS CANADA (1979). Data Collection and Dissemination Methods for International Travel Statistics in Canada. International Travel Section, Statistics Canada.

UNITED STATES TRAVEL AND TOURISM ADMINISTRATION (1989). In-flight Survey: Overseas and Mexican Visitors to the United States. Survey Period: January-December 1988. Washington, D.C.: United States Travel and Tourism Administration.

WELLS, J.K., WILLIAMS, A.F., and LUND, A.K. (1990). Seat belt use on interstate highways. *American Journal of Public Health*, 80, 741-742.

WOLFE, A.C. (1986). 1986 U.S. National Roadside Breathtesting Survey: Procedures and Results. Ann Arbor, Michigan: Mid-America Research Institute.

WOOD, D. (1978). The Eastleigh Carrefour: a hypermarket and its effects. London: U.K. Department of the Environment.

YATES, F. (1981). *Sampling Methods for Censuses and Surveys* (4th ed.). London: Charles Griffin.

ZIEGLER, P.N. (1983). Guidelines for Conducting a Survey of the Use of Safety Belts and Child Safety Seats. Washington, D.C.: U.S. Department of Transportation.