SUMMARY AND CONCLUSIONS

The methodological investigation reported in this paper discloses some disturbing facts about underreporting that typically go undetected. The effect of such information reduction can be substantial when the survey data are used uncritically for assessing mobility levels and for determining modal shares and overall travel activity levels in terms of duration and length. The paper identifies a number of methods, indicators, and relationships that permit the analyst or planner to upgrade the results of surveys by means of careful adjustments.

REFERENCES

1. W. Brög. Cooperation Between Empirical Research on Transport Behaviour and Transport Policy Making. Presented at Conference on Social Aspects of Transport: How to Use Social Research in Transport Policy Making, Chichester, England, Sept. 1980.
2. W. Brög and A.H. Meyburg. Influence of Survey Methods on the Results of Representative Travel Surveys. Presented at the 61st Annual Meeting, TRB, Jan. 1982.
3. A.H. Meyburg and W. Brög. Validity Problems in Empirical Analyses of Non-Home-Activity Patterns. TRB, Transportation Research Record 807, 1981, pp. 46-50.
4. W. Brög. Design and Application of Improved Methodologies for Large-Scale Surveys. Presented at 61st Annual Meeting, TRB, Jan. 1982.
5. W. Brög. Mobility and Lifestyle--Sociological Aspects. Presented at Eighth International ECMT Symposium on Theory and Practice in Transport Economics, Istanbul, Turkey, Sept. 1979.
6. W. Brög and A.H. Meyburg. Consideration of Nonresponse Effects on Large-Scale Mobility Surveys. TRB, Transportation Research Record 807, 1981, pp. 39-46.
7. W. Brög and K.H. Neumann. The Interviewee as a Human Being--A Critical Contribution to the Discussion of Methods of Empirical Social Research. Presented at the 44th Seminar on Ways and New Ways of Data Collection, European Society for Opinion Surveys and Market Research, Jouy-en-Josas, France, Nov. 1977.
8. M.F. Wermuth and G. Meerschalk. Zur Repräsentanz Schriftlicher Haushaltsbefragungen (Representativeness of Written Mail-Back Household Surveys). German Ministry of Transport, Munich, Germany, 1981.
9. W. Brög and A.H. Meyburg. The Non-Response Problem in Travel Surveys--An Empirical Investigation. TRB, Transportation Research Record 775, 1981, pp. 34-38.
10. M.I. Clarke, M.C. Dix, and P.M. Jones. Error and Uncertainty in Travel. Transport Studies Unit, Oxford Univ., Oxford, England, Working Paper 45, 1979.
11. S.W.M. Hendrikx. Methodische Ergebnisse im Ablauf einer Landesweiten Kontinuierlichen Erhebung zum Verkehrsverhalten in den Niederlanden (Methodological Results in the Course of a Nationwide Continuous Survey Of Travel Behavior in the Netherlands). Dutch Central Statistics Office, Heerlen, 1979.
12. Socialforschung Brög. KONTIV 75-77 (Continuous Survey of Travel Behavior), Overview Vols. 1-3. German Ministry of Transport, Munich, Germany, 1977.

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

Estimation of Cross-Cordon Origin-Destination Flows from Cordon Studies

UZI LANDAU, EZRA HAUER, AND ITZHAK GEVA

When traffic counts obtained in a cordon study are supplemented by information about the origin and destination of a small sample of trips crossing the cordon, it is possible to obtain estimates of the prevailing origin-destination (O-D) flows. The purpose of this paper is to describe a coherent method for the identification of the maximum-likelihood estimates of vehicular O-D flows. The solution procedure proves to be relatively simple. The estimates obtained are of flows between the cordon stations, flows between the stations and the area inside the cordon, flows between traffic zones, etc. The method is illustrated by a detailed numerical example. A real-life application of the estimation method to the downtown Toronto cordon is described. It appears that it is possible to obtain much-needed O-D information at relatively little extra cost.

The pattern of tripmaking in an urban area is precisely and succinctly described by an origin-destination (O-D) flow matrix. This is why O-D matrix estimates should serve as basic information for traffic management and transport planning tasks. Unfortunately, methods for obtaining such estimates are time consuming and costly. This is why, even in major metropolitan areas, information about the prevailing pattern of tripmaking is often sketchy.

Many cities conduct cordon studies periodically. Traffic into and out of the cordon area is counted and some inferences about traffic flow patterns and trends are possible. However, a cordon study does not yield information about O-D flows. The idea explored in this paper is the possibility of attaching a small-sample O-D survey to the routine cordon counts and of using the combined information for the estimation of the prevailing O-matrix of vehicle flows crossing the boundary of the cordon at least once.

This idea is in line with other recent developments, which all rely on better utilization of the ubiquitous traffic-count information for O-D estimation. A detailed review of such models is available (1-3).

For some simple transport systems one can obtain good estimates about the O-D flows by using traffic counts only (4). In more complex systems, regularities of travel behavior have to be invoked in the form of similarity with other systems (5,6), patterns of the past, or models of the gravity form (7-12).

The estimation procedure developed in this paper departs from previous work in that travel behavior is brought into estimation by the specific information contained in small samples obtained by a survey. It is therefore not necessary to rely on elusive microstates (as in entropy models), to argue by induction and analogy, or to trust general interaction-at-distance or route-choice models. Rather, we will find the matrix of O-D flows, which is consistent with the observed traffic counts and which is most probable in view of the O-D samples obtained. Thus the task is to solve a constrained maximum-likelihood problem.

PROBLEM FORMULATION AND SOLUTION

Consider a cordon line surrounding a cordon area. There are n survey stations on the cordon line. In a conventional cordon study, vehicles entering and leaving the cordon area are counted at each survey station. Accordingly, let O_k be the number of vehicles entering the cordon area during a specified period of time at station k (k = 1, 2, ..., n) and let D_l be the number of vehicles leaving the cordon area during the specified period of time at station l (l = 1, 2, ..., n).

To allow estimation of O-D flows at each survey station, a random sample is selected from the inbound and outbound vehicles. From the drivers of the selected vehicles, additional information is obtained. For a vehicle entering the cordon area, the exit station (l) is ascertained. If the trip ends inside the cordon area, the index O is used. For a vehicle leaving the cordon area, the entry station (k) is determined. The index o will signify a trip starting inside the cordon area. Accordingly, t_{kl} will be the number of vehicles in the random sample obtained at station k from those entering the cordon area that report leaving the cordon area via station l. t_{kl}' is the corresponding number of vehicles in the random sample obtained at station l from those leaving the cordon area.

The practicalities of obtaining t_{kl} and t_{kl}' will be discussed later in the context of a real application to the downtown cordon area of metropolitan Toronto. In this section, the focus is on finding estimates of the vehicular flows T_{kl} from station k to station l by using the cordon-count data (O_k and D_l) and sample information (t_{kl} and t_{kl}'). When this problem has been solved, it will prove simple to obtain estimates of vehicular flows T_{ijkl} from zone i to zone j passing stations k and l.

Traffic engineers and transportation planners are accustomed to solving this problem in a heuristic manner. Data from roadside interviews or license plate surveys are factored to approximate traffic counts. The disadvantage of factoring is that there are many possible ways of going about the task and as many different solutions. In addition, it is difficult to make factored estimates to match all counts. In contrast, the solution derived below is unique. It has the merit of being "best" in the sense that it identifies the array of O-D flows (T_{kl}*), which maximizes the probability of observing the specific values of O_k, D_l, t_{kl}, and t_{kl}'.

Having established the notation and declared the approach to the solution, we now formulate the problem in mathematical terms.

Consider the random sample t_{ko}, t_{ki}, ..., t_{kn} obtained at station k. This sample is drawn from the flows T_{ko}, T_{ki}, ..., T_{kn}, which are unknown. Only their sum $\Sigma_l T_{kl} = O_k$ is known.

The probability of observing this sample is approximately as follows:

$$\left[\left(\sum_l t_{kl}\right)!/\prod_l (t_{kl}!)\right] \times \prod_l (T_{kl}/O_k)^{t_{kl}} \quad (1)$$

The symbols Σ_l and Π_l denote summation and product over l = o, 1, ..., n. (The multinomial probability model embodied in Equation 1 is only approximate because it assumes "sampling with replacement." As long as the sample is a small fraction of the population, the assumption seems proper.)

An expression analogous to Equation 1 can be written for every random sample obtained at each of the n survey stations for both inbound and outbound flows. Therefore the probability of observing all random samples is given by the following:

$$\left(\prod_{k=1}^{n}\left\{\left[\left(\sum_l t_{kl}\right)!/\prod_l (t_{kl}!)\right] \times \prod_l (T_{kl}/O_k)^{t_{kl}}\right\}\right)\left(\prod_{l=1}^{n}\left\{\left[\left(\sum_k t_{kl}'\right)!/\prod_k (t_{kl}'!)\right] \times \prod_k (T_{kl}/D_l)^{t_{kl}'}\right\}\right) \quad (2)$$

The probability in Equation 2 is a function of the n x n array of unknown flows T_{kl}. We wish to identify that array T_{kl}* for which this probability is maximum. However, the solution must satisfy the following traffic-count constraints:

$$\sum_l T_{kl} = O_k \quad \text{for } k = 1,2,\ldots,n$$
$$\sum_k T_{kl} = D_l \quad \text{for } l = 1,2,\ldots,n \quad (3)$$

By using the method of Lagrange multipliers, we find the following:

$$T_{ko}^* = t_{ko}/\alpha_k \quad k = 1,2,\ldots,n$$
$$T_{ol}^* = t_{ol}'/\beta_l \quad l = 1,2,\ldots,n$$
$$T_{kl}^* = (t_{kl} + t_{kl}')/(\alpha_k + \beta_l) \quad k = 1,2,\ldots,n;\ l = 1,2,\ldots,n \quad (4)$$

The 2n unknown Lagrange multipliers α_k and β_l are determined by an iterative algorithm that uses the 2n Equations 3. The algorithm is described and illustrated by a numerical example in the section on solution procedure.

So far, the station of entry (or the cordon area) has been regarded as the origin and the station of exit (or cordon area) as the destination. However, for many purposes the traffic zone in which a trip commences is regarded as the origin and the traffic zone in which the trip terminates as the destination. In such cases, one wishes to have an estimate of the zone-to-zone O-D matrix (T_{ij}) rather than the station-to-station O-D matrix (T_{kl}).

By using the results of the above analysis, it is easy to obtain an estimate of the zone-to-zone O-D matrix if the drivers sampled from the flow provide information not only about stations of exit or entry but also about trip origin and destination.

If we generalize the previous notation, let t_{ijkl} be the number of vehicles in the random sample obtained at station k from those entering the cordon area that report zone i as the trip origin and zone j as destination and leave the cordon area by station l, t_{ijkl}' be the corresponding number of vehicles in the random sample obtained at station l from those leaving the cordon area, and T_{ijkl}* be the best estimate of the vehicular flows from zone i to zone j via stations k and l.

If we formulate the problem again as constrained

Figure 1. Cordon line with survey stations.

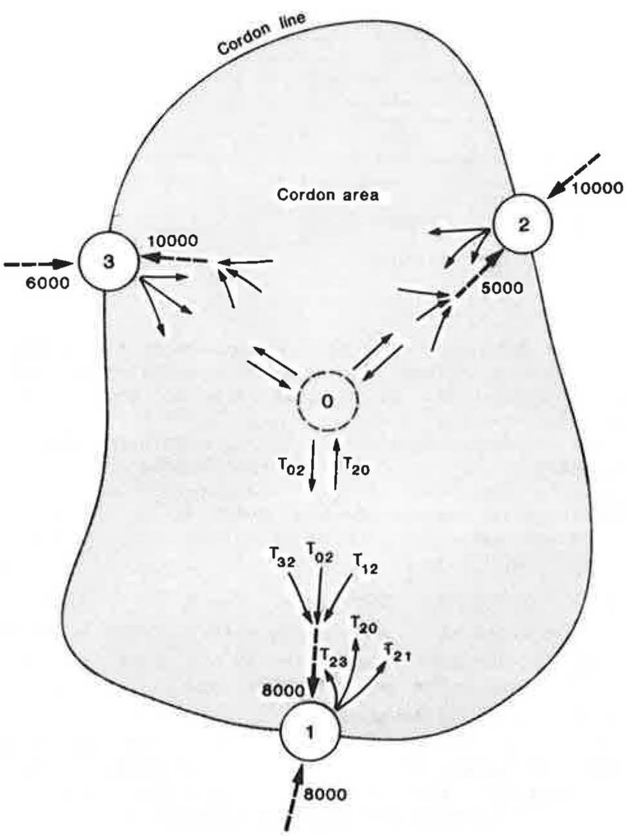


Table 1. Results of sample O-D survey.

Inbound Samples					Outbound Samples			
	To Station					To Station		
From Station	0	1	2	3	From Station	1	2	3
1	10	-	20	30	0	5	10	20
2	30	20	-	40	1	-	40	40
3	40	40	40	-	2	10	-	50
					3	40	20	-

maximization of likelihood, it can be shown that

$$T_{ijkl}^* = [(t_{ijkl} + t_{ijkl}')/(\sum_{i,j} t_{ijkl} + \sum_{i,j} t_{ijkl}')]T_{kl}^* \quad (5)$$

Thus the estimated flow between k and l (T_{kl}^*) is apportioned to zone pairs according to their proportion in the sample passing through k and l. Of course,

$$T_{ij}^* = \sum_k \sum_l T_{ijkl}^* \quad (6)$$

The problem has been formulated and solved by assuming that O-D samples are available at all cordon stations in both the inbound and outbound directions. This may not always be the case. When only inbound or only outbound samples are missing, the solution procedure remains without change. However, when the outbound sample at, say, station 3 is missing and the inbound sample at station 7 has not been obtained, there is a gap in the information for the estimation of T_{13}, T_{03}, and T_{10}. If stations 3 and 7 are on minor roads, the analyst may combine them with neighboring stations and treat two or more stations as one. This is tantamount to assuming that the trip origins and destinations in the stations combined are similar, an often justifiable assumption. Alternatively, a two-step estimation procedure can be used. In the first step the above model is applied to stations for which sample O-D information is available; in the second, the flow not allocated in the first step is distributed by using an assumption of equally likely outcomes. The two-step process is incorporated in the computer code used for the application described in the last section.

Table 2. Factoring up by using inbound samples and counts.

	To Station				
From Station	0	1	2	3	Row Sum (O)
0		(1222)	(2667)	(1 444)	
1	1667	-	3333	5 000	10 000
2	2666	1778	-	3 556	8 000
3	2000	2000	2000	-	6 000
Column sum (D)		5000	8000	10 000	

Table 3. Factoring up by using outbound samples and counts.

	To Station				
From Station	0	1	2	3	Row Sum (O)
0	-	454	1143	1818	
1	(1793)	-	4571	3636	10 000
2	(2546)	909	-	4545	8 000
3	(79)	3636	2285	-	6 000
Column sum (D)		5000	8000	1000	

SOLUTION PROCEDURE

The flow estimates are given in Equation 4 as a function of parameters α_1, α_2, ..., α_n; β_1, β_2, ..., β_n. The values of these parameters are unknown and have to be determined so that the sum of the flows at each station equals the corresponding traffic counts. The process of solving for α_i and β_j is straightforward and is best explained with reference to a simple numerical example.

Consider a cordon line with three stations (Figure 1).

The stations are numbered consecutively; a dummy station (0) represents the cordon area proper. Shown alongside the dashed arrows are the inbound and outbound flows counted at the cordon stations. Their tributary flows (T_{kl}) are schematically represented by the short solid arrows. These are the flows that are unknown and for which estimates need to be obtained.

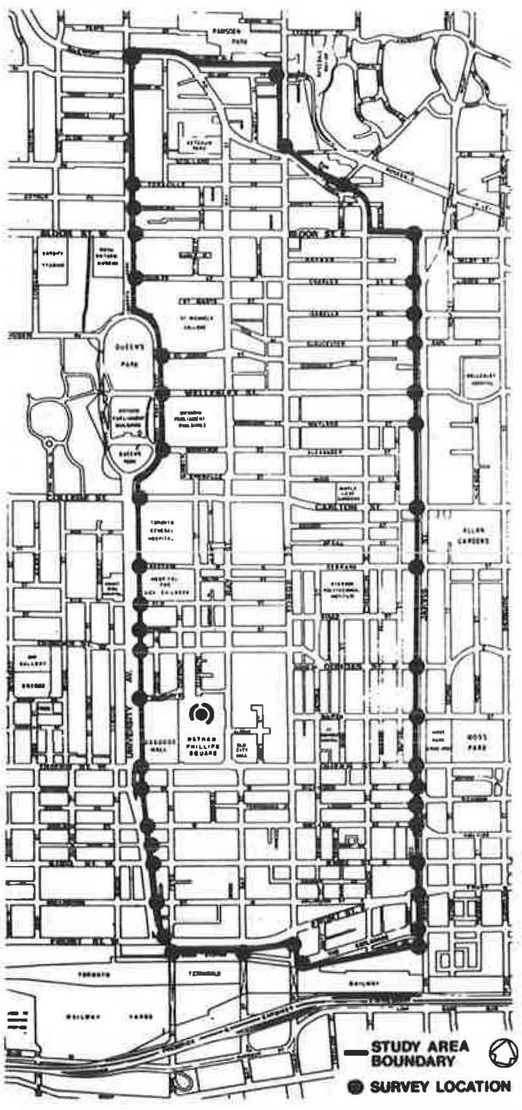
In addition to the traffic counts, a random sample of drivers is polled at each station. In this illustration we assume that entering drivers are asked only about their exit station and vice versa. The resulting information is presented in Table 1. Thus, at station 1, for example, of the 60 inbound drivers asked, 10 ended their trip inside the cordon area (station 0), 20 exited via station 2, and 30 by station 3.

Were one to use the inbound samples, in order to factor them up to the inbound traffic counts the estimates in Table 2 would be obtained. The entries in parentheses are calculated to match the column sums. Alternatively, were one to use the outbound samples and factor those up to the outbound flow, the estimates in Table 3 would be obtained.

Table 4. Estimates of station-to-station flows.

From Station	To Station 0	1	2	3	Row Sum (O)	α_k
0	–	835	1597	1 964	4 396	
1	1422	–	4513	4 065	10 000	0.007 03
2	2404	1624	–	3 971	7 999	0.012 48
3	1569	2541	1890	–	6 000	0.025 49
Column sum (D)	5395	5000	8000	10 000	28 395	
β_1		0.005 99	0.006 26	0.010 18		

Figure 2. Downtown cordon in Toronto.



Comparison of the estimates in Tables 2 and 3 highlights the difficulty inherent in heuristic factoring: Each flow can be estimated in two ways by using two different sets of field data, and the two estimates are usually different. The reconciliation of these differences by some balancing gives rise to ambiguity.

However, redundancy in data should be viewed as an opportunity to improve the quality of the estimation rather than as an embarassing nuisance. The solution obtained earlier resolves this difficulty by making good use of available information and yields unique estimates that are in some sense optimal.

The solution algorithm begins by obtaining initial estimates of $\alpha_1, \alpha_2, \ldots, \alpha_n$. In Equations 4, $\alpha_k = t_{ko}/T_{ko}^*$. The value of T_{ko}^* is at present not known. However, a reasonable starting guess may be the average of corresponding values in column 0 of Tables 2 and 3. Thus, for example, $^0\alpha_1 = 10/[1/2(1667 + 1793)] = 0.005\ 78$. Similarly, $^0\alpha_2 = 30/[1/2(2667 + 2546)] = 0.011\ 51$ and $^0\alpha_3 = 40/1/2(2000 + 79) = 0.038\ 48$. The left superscript is a counter of iterations. The value of α_k after, say, the seventh iteration will be denoted $^7\alpha_k$.

By using these tentative values for α_k, the first estimates of β_1 can be obtained. Thus, for example, the sum $T_{01} + T_{21} + T_{31}$ must be 5000. Substituting for T_{0k} from Equation 4,

$$(5/^1\beta_1) + [(20 + 10)/(^1\beta_1 + 0.011\ 51)] + [(40 + 40)/(^1\beta_1 + 0.038\ 48)] = 5000 \qquad (7)$$

In this equation, $^1\beta_1$ is the only unknown. Although it is not possible to write $^1\beta_1$ as a function of the constants in this equation, its value is easily determined by iterative methods. In this case, $^1\beta_1 = 0.004\ 14$ is the solution.

After a few iterations, the solution in Table 4 is reached.

The underlined part of Table 4 contains the flow estimates T_{ij}^*. Apart from minor discrepancies due to rounding off and termination of iterations, the estimates comply with observed traffic counts and are consistent with the information from the sample survey. The underlined cells represent trips that originate in the cordon area and have destinations outside it (4396), trips that originate outside with destinations inside the cordon area (5395), and total number of trips (28 395).

APPLICATION TO DOWNTOWN TORONTO

In most metropolitan areas, cordon counts are performed periodically in order to keep tabs on the trends in traffic and to provide a data base for transport planning and management. The Metropolitan Toronto Cordon Count Program was established in 1975. A count is conducted every second year on a web of cordons and screen lines. The downtown cordon (Figure 2) is a small yet central part of this web. During the summer of 1981, the normal count program for this cordon was supplemented by a sample O-D survey. For the Metropolitan Toronto Planning Department this was an opportunity to obtain estimates of O-D flows that are notoriously difficult to come by. For us, this was a chance to examine the estimation procedure in practice.

To obtain a sample of O-D information, a random set of license plates was recorded at each cordon station. The address of the registered owner of the vehicle was then traced and a map of the cordon area with a few questions was mailed to him. The person driving the vehicle at the time of the sighting was asked about the origin and destination of the trip as well as its trace on the map of the cordon area.

Of the 25 453 license plates recorded, 20 319 addresses were obtained and questionnaires were mailed to those. Some 6740 responses were obtained and of those, 5835 had usable answers and could be coded.

There are 22 major-flow cordon stations and 24 stations on minor streets. For a major road, an average sample of 180 was obtained; for minor streets, the average sample size is 80. These are close to the planned target. However, the planned sample size was selected without the benefit of a statistical survey design. Results of this study will serve to examine the effect of sample size on estimation accuracy. This examination is under way and it is hoped that its results will facilitate better survey design for future studies.

Random sampling is easier said than done. In the field this requires care for at least two reasons. First, the destination of a trip is not independent of the lane in which the vehicle travels. Thus, the sampling rule must ensure that there is no predisposition to record the plates of vehicles in, say, the curb lane. Otherwise the sample would contain an uncharacteristically large proportion of right-turning vehicles. Second, the origin of a trip is not independent of the place of the vehicle in the platoon. Platoon leaders might be straight-through vehicles with turning traffic forming the tail of the platoon. Thus, the sampling rule must be such that vehicles are selected at a uniform rate from all parts of the platoon. These problems can be obviated by, for example, registering the plates of all vehicles that end with some prespecified digits. It is more difficult to ensure randomness in the second stage of the sampling progress. One has no control over the correctness of the address to which the letter is sent, whether the registered owner was the driver, whether the driver recalls the trip, who decides to fill out the questionnaire and return it, etc. It is clear that this particular method of obtaining a sample of O-D trips will yield results that are not quite representative of trips by fleet vehicles, commercial vehicles, taxis, etc. Nor is this method particularly cheap. To obtain one response (field work + address search + postage + coding + keypunching), we had to invest approximately $1.60 and 10 min of work. It may be possible to simplify the process considerably (for example, by handing out questionnaires to stopped traffic). Experimentation with other workable sampling methods that do ensure randomness will be the subject of future work.

A computer code has been written (in FORTRAN) and fed the cordon-count data and the results of the sample O-D survey. Several O-D flow tables have been produced:

1. Flows between cordon stations with the cordon area as station 0,
2. Same as in 1 but the cordon area disaggregated into a few zones, and
3. Flows between traffic zones for trips crossing the cordon once or more.

Many other aggregations and disaggregations are possible. A copy of the computer code listing is available on request from the Department of Civil Engineering at the University of Toronto.

SUMMARY AND DISCUSSION

When cordon counts are supplemented by information about the origin and destination of a small sample of trips crossing the cordon, it is possible to obtain estimates of the prevailing O-D flows. The problem is formulated and solved as the task of identifying the most likely flow estimates. The solution procedure is relatively simple. Estimates obtained are of flows between cordon stations, flows between the stations and the cordon area, flows between traffic zones, etc. An application of the method to the downtown cordon in Toronto is described. Thus, a method is suggested that allows estimation of much-needed O-D information at relatively little extra cost.

As formulated, the model has several flaws. The vehicular flows that are the subject of estimation here are the flows that prevailed at the time of the survey. Ordinarily, one is not interested in the flows that prevailed at some specific time. Rather, one wishes to know the underlying expected value. This distinction is usually disregarded in practice and therefore may not be important. However, personnel limitation and cost preclude the conduct of cordon counts and sample surveys at all stations of the cordon simultaneously. Therefore, one cannot pretend to estimate flows that prevail on some specific day. Yet the flow estimates are not expected values either. An added weakness in the formulation is the insistance that the sum of estimates match the traffic counts exactly. Most practitioners will find this pleasing and in line with other models in use. Nevertheless, the traffic counts are random variables just as the results of O-D samples are, and a coherent model must treat them as such [see, for example, the report by Kirby and Murchland (13)].

The formulation, solution, and illustration of the problem are all presented against the background of a cordon study. One does not ordinarily think of a bus line with its stations or a freeway with its ramps as a cordon. In an earlier paper (4), Hauer, Pagitsas, and Shin discussed the task of O-D estimation on such systems when only traffic counts are available. When in addition to traffic counts one also has small-sample O-D data, the aforementioned systems can be regarded as a cordon and the method developed in this paper applied.

ACKNOWLEDGMENT

This research was supported by the National Science and Engineering Research Council (grant A8158) and the Toronto-York Joint Program in Transportation. The application to downtown Toronto was coordinated by the Metropolitan Toronto Planning Department with the assistance of the Ministry of Transportation and Communications and Metro Roads and Traffic. Their help is gratefully acknowledged.

REFERENCES

1. V.K. Chan, R.G. Dowling, and A.E. Willis. Deriving Origin-Destination Information from Routinely Collected Traffic Counts: Phase I--Review of Current Research and Application. Institute of Transportation Studies, Univ. of California, Berkeley, Jan. 1980.
2. G.M. Glen and P. Hague. Cross Cordon Traffic Flow Matrices from Roadside Interviews. Traffic Engineering and Control, Vol. 16, No. 7/8, July/Aug. 1975.
3. Y. Guz and others. Determining an Origin-Destination Trip Table Based on Observed Volumes. Presented at Annual Meeting of the Operations Research Society of America and

Institute of Management Sciences, New York, May 1978.
4. E. Hauer, E. Pagitsas, and T.B. Shin. Estimation of Turning Flows from Automatic Counts. TRB, Transportation Research Record 795, 1981, pp. 1-7.
5. E. Hauer and T.B. Shin. Origin-Destination Matrices from Traffic Counts: Application and Validation on Simple Systems. Traffic Engineering and Control, Vol. 22, No. 3, March 1981, pp. 118-121.
6. L.G. Willumsen and D. Van Vliet. Validation of the ME2 Model for Estimating Trip Matrices from Traffic Counts. Presented at Eighth International Symposium on Transportation and Traffic Theory, Toronto, Canada, June 1981.
7. P. Robillard. Estimating an O-D Matrix from Observed Link Volumes. Transportation Research, Vol. 9, 1975, pp. 123-128.
8. H.J. Van Zuylen. Some Improvements in the Estimation of an O-D Matrix from Traffic Counts. Presented at Eighth International Symposium on Transportation and Traffic Theory, Toronto, Canada, June 1981.
9. H.J. Van Zuylen and L.G. Willumsen. The Most Likely Trip Matrix Estimated from Traffic Counts. Transportation Research, Vol. 14B, 1980, pp. 281-293.
10. H.J. Van Zuylen. The Estimation of Turning Flows on a Junction. Traffic Engineering and Control, Vol. 20, No. 11, Nov. 1979.
11. L.G. Willumsen. Estimation of an O-D Matrix from Traffic Counts--A Review. Institute for Transport Studies, Univ. of Leeds, England, Working Paper 99, 1978.
12. L.G. Willumsen. O-D Matrices from Network Data: A Comparison of Alternative Methods for Their Estimation. Proc., Summer Annual Meeting, Planning and Transport Research and Computation Co., London, England, July 1978.
13. H.R. Kirby and J.D. Murchland. Gravity Model Fitting with Both Origin-Destination Data and Modelled Trip Engineering Estimates. Proc., Eighth International Symposium on Transportation and Traffic Theory, Toronto, Canada, June 1981.

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

Toward Improved Collection of 24-H Travel Records

PETER R. STOPHER AND IRA M. SHESKIN

A major concern of many transportation-planning surveys is to collect data on a 24-h weekday period of travel for all members of a household who are five years of age and older. Traditionally, this has been done by asking household members to recollect their travel for the immediately preceding 24-h weekday period. A travel diary that has been developed to be used by each household member to record travel as it is undertaken is described. Although the concept of a travel diary is not new, several aspects of this diary are new and appear to be very effective in obtaining a response. The diary has been used in some recent surveys and the results of these applications are described briefly. In general, response to the instrument was found to be good when it was administered in an effective supporting survey context. It is concluded that this travel diary represents a good procedure for measuring travel and should be tested in comparable studies with conventional procedures.

Travel-behavior surveys are designed to obtain information about where, when, how, and with how many others the respondent and members of his or her household over the age of five have traveled during a 24-h period. Strictly speaking, when asked as a historical record, the information obtained is the respondent's perceptions of his or her behavior. In a travel-behavior survey, such perceptions are likely to be flawed significantly because a respondent is being asked to remember a sequence of events (and details about these events) that, to the average person, may have seemed unimportant when they occurred. The probability of the omission of trips or of the reporting of inaccurate trip details is heightened even more when (as is often necessary) one household member is asked about the travel of another. Another problem is created when lengthy travel records are collected following a lengthy home interview survey and both respondent and interviewer are tired. The possibility also exists that trips made on another day will be remembered incorrectly as having been made on the subject day. From the early metropolitan area transportation studies, such as the Detroit Metropolitan Area Transportation Study (1), to the present, most travel-behavior surveys have employed this method as part of the urban transportation planning process (2,3). A typical example of such a survey instrument may be found in books by Stopher and Meyburg (4) and by Domencich and McFadden (5).

A second method of collecting travel-behavior information is to intercept people in the process of making a trip. The roadside interview and the on-board transit survey (6) are the most common. The significant advantage of such a technique is that respondents are surveyed at the time when they are least likely to forget trip details. On the other hand, such surveys, for logistical reasons, must be kept relatively short and it is impossible to construct a 24-h trip record for the respondent and his or her household by using this method.

The third method is to use a travel diary in which respondents are asked to report their own future behavior. Such a technique is a cross between observing behavior (such as counting riders on a bus) and a participatory survey (7). The major problem such a technique is designed to circumvent is that of memory. Evidence that memory can be a significant problem in recalling behavior has been provided by Cantril (8). In his survey, only 87 percent of persons interviewed twice at a three-week interval gave the same answer both times about the person for whom they voted in the 1940 presidential election. By presenting a respondent with a document that needs to be filled out about travel for the next day, which can be filled out partly or fully while traveling or partly while traveling and partly at the end of the day, less information should be lost to memory problems. One problem not solved, of course, is that certain trips (such as