

Bus Route O-D Matrix Generation: Relationship Between Biproportional and Recursive Methods

PETER G. FURTH AND DAVID S. NAVICK

Planners must sometimes synthesize transit route origin-destination (O-D) matrices with limited data, usually on-off counts and sometimes a small or outdated O-D survey sample. When a small O-D sample is available, iterative methods such as the biproportional method that begin with the sample as a seed matrix can be used, adjusted to match on-off totals. When only on-off totals are available, the recursive method of Tsygalnitsky has been found to match O-D patterns on some routes better than others. This method is in fact a special case of the biproportional method using an implicit null seed matrix that contains information on directionality and minimum trip length. It illustrates why the recursive method is inappropriate when there is significant competition between routes, and offers a correction for when on-off data have been aggregated to the segment level. Estimation errors are then compared to help indicate how large the seed sample should be in order to produce a more accurate estimate than an estimate produced with a null seed.

A route-level origin-destination (O-D) matrix (trip table) gives the number of passengers traveling between each pair of stops or stations on a transit route in a particular direction. It can be specific to any period of interest, from the individual vehicle trip to an entire day. A route-level O-D matrix is an important descriptor of passenger demand that has been used for such analyses as systematic route evaluations (1,2), route and schedule design for short-turning (3), zonal service (4), limited-stop service (5), and complementary express and local service (6).

A route-level O-D matrix can be obtained by directly sampling passengers. The typical passenger survey, in which passengers fill in a questionnaire asking where they boarded and where they plan to leave, leaves a lot to be desired. Response rates are often low, and vary according to critical factors such as trip length—did the passenger have enough time to fill out the questionnaire?—and origin-stop—did the passenger get a seat? Is this stop in a low literacy neighborhood?—which may bias the results. A special purpose survey method, called by one author the “no questions asked” method (2), appears to overcome this nonresponse problem. Passengers are given origin-coded cards when they board and are asked to return the cards when they alight. By careful collection of the cards by alighting stop, O-D information is obtained. Practitioners report response rates of over 90 percent (2,7). However, this method is not in common use, because it requires one checker at each door and careful pre-trip preparation.

Far more common and easier to obtain than O-D data are on-off counts. In the context of O-D matrix generation, on-

off counts represent row and column totals. It is not difficult for a ride checker to obtain a 100 percent sample of on-off counts, and measurement error is generally agreed to be quite small. Therefore an O-D matrix whose row and column totals agree with the on-off counts should be preferred to one obtained by simple expansion of a small O-D sample. Of course, there are many possible O-D matrices whose row and column totals match the on-off counts. The problem of O-D matrix synthesis is to generate an O-D matrix that agrees with a given set of row and column totals and that meets some criteria of being the best or most likely O-D matrix. Ben-Akiva et al. (8) describe three methods for combining a small O-D sample with on-off counts: the biproportional method, constrained maximum likelihood, and constrained generalized least squares. All three of these methods involve iterative computations. The first two are preferable because the third sometimes generates negative matrix entries, even though all three yield very similar results. The biproportional method is computationally more attractive, is better known, and has been used in a variety of contexts (9–11). In further work, Ben-Akiva (12) shows how the maximum likelihood approach can be used to derive estimation methods that combine various imperfect sources of information. In an application to transit route O-D estimation, his assumptions about the structure of the nonresponse bias lead again to the simple biproportional method.

It is often the case, however, that a small O-D sample is not available, or that the small sample is so small or suspected of bias that an estimate based on it may not be reliable. A method for synthesizing a route-level O-D matrix from on-off counts alone was proposed by Tsygalnitsky (13). It is a very simple method involving a single pass of recursive calculations, and can be done by hand (although use of a spreadsheet or computer program is still advisable). This method has also been used by London Transport in at least one study, presumably having been developed independently (1). Tsygalnitsky found that his recursive method fit well with data from Toulouse, France. Simon and Furth (7) also tested it against O-D data from two routes in Los Angeles, and again found a good fit, although the fit on one route was better than that on another. Ben-Akiva et al. (8) tested the recursive method against O-D matrices generated using the biproportional and constrained maximum likelihood methods for two Boston area routes and found that it yielded matrix estimates that differed substantially from the estimates obtained by the iterative methods based on a small-sample O-D survey.

Although Tsygalnitsky's recursive method and the biproportional method are motivated from different assumptions,

the recursive method is actually a special case of the biproportional method. The biproportional method takes an initial matrix, called a seed matrix, and factors it to match on-off counts. The seed matrix contains information concerning the preferences for the various O-D pairs. Typically, the seed matrix is an O-D sample. If there is no O-D sample to begin with, a reasonable guess is to use a "null seed," one that assumes that every permissible O-D pair is equally preferred. It is demonstrated that Tsygalnitsky's recursive method is the same as the biproportional method using a null seed.

This insight makes it possible to better analyze which method is more appropriate under various circumstances. A small O-D sample contains valuable site-specific information about O-D pair preferences but is also subject to sampling error and nonresponse bias. A null seed has no sampling error or nonresponse bias but lacks site-specific information. In addition, two common factors—aggregation of stops into segments and competition from other routes—are shown to be in contradiction to the assumptions underlying the null seed, and consequently the recursive method should not be expected to perform well under these circumstances.

BUS ROUTES ANALYZED

Repeated reference is made to four bus routes that have been previously analyzed. Lines 16 and 93, analyzed by Simon and Furth (7) are operated by the Southern California Rapid Transit District. For Line 16, virtually complete O-D data, encompassing 266 passengers, were obtained from five inbound short-turning trips over a 5-mi radial route containing 40 stops. For Line 93, virtually complete O-D data were obtained on four a.m.-peak (383 passengers) and four p.m.-peak (273 passengers) trips. Four trips were local trips covering the entire 140-stop route from downtown Los Angeles to the San Fernando Valley, three trips were short-turned in North Hollywood (about 90 stops), and one p.m. trip ran express from downtown to the valley. Routes 77 and 350, analyzed by Ben Akiva et al. (8), are operated by the Massachusetts Bay Transportation Authority. These routes were analyzed inbound in the a.m. peak and outbound in the p.m. peak. The available data consist of a small O-D sample augmented by on-off counts. Route 77 is a heavily used radial route, 5.5 mi long, running through the suburb of Arlington into Harvard Square in Cambridge. In the a.m. peak, 2,148 passengers were counted, and O-D data were obtained from 54. In the p.m. peak, 1,617 passengers were counted, with O-D data obtained from 138. Route 350 is 15.2 mi long, with a large collection/distribution section in suburbs north of Boston, connected by express operation to selected stops in Cambridge and downtown Boston. In the a.m. peak, 485 passengers were counted, with O-D data obtained from 76. In the p.m. peak, 200 passengers were counted, with O-D data obtained from 61.

TSYGALNITSKY'S RECURSIVE METHOD

Tsygalnitsky's recursive method proceeds stop by stop, distributing alightings at each stop among origin stops in proportion to the number of people from each origin stop who are eligible to alight. To be eligible, passengers must have

traveled a minimum distance, and must not have alighted previously. Taking each stop as a node, with nodes consecutively numbered from 1 to n in the direction of travel, let

$$\begin{aligned} t_{ij} &= \text{passenger trips from } i \text{ to } j, \\ t_{i\cdot} &= \text{boardings at } i = \sum_j t_{ij}, \\ t_{\cdot j} &= \text{alightings at } j = \sum_i t_{ij}, \\ m_i &= \text{first node at which passengers who board at } i \text{ are} \\ &\quad \text{eligible to alight } (m_i \geq i), \\ E_j &= \text{set of nodes that can serve as origins for passengers} \\ &\quad \text{alighting at } j, \\ e_{ij} &= \text{number of passengers who boarded at } i \text{ who are eli-} \\ &\quad \text{gible to alight at } j, \\ e_{\cdot j} &= \text{total number of passengers eligible to alight at } j = \\ &\quad \sum_i e_{ij}, \text{ and} \\ f_j &= \text{fraction of eligible passengers who alight at } j = \\ &\quad t_{\cdot j} / e_{\cdot j} \end{aligned}$$

Initially, set $e_{ij} = 0$ for all (i, j) except when $j = m_i$, in which case set $e_{ij} = t_{i\cdot}$. Computation begins with the first node at which passengers are eligible to alight; call it Node k . After calculating $e_{\cdot k}$ and f_k , let

$$t_{ik} = e_{ik} f_k \quad \text{for all } i \in E_k \quad (1)$$

Stop if $k = n$; otherwise update:

$$e_{i, k+1} = e_{ik} - t_{ik} \quad \text{for all } i \in E_k \quad (2)$$

and advance to the next node (let $k = k + 1$) and return to Equation 1.

Simon and Furth call this method a fluid analogy, because passengers on the bus are likened to a thoroughly-mixed fluid out of which alighting passengers are drawn at each alighting stop in proportion to their representation in the fluid. Newly boarding passengers are added to the fluid after they have met the minimum travel distance criterion. (This minimum distance may be expressed in stops, distance, or time units, and may vary from stop to stop.) Ben-Akiva et al. (8) call it an intervening opportunities method, because it follows the logic of classical intervening opportunities models in giving priority to closer destinations.

BIPROPORTIONAL METHOD

Additional notation that will be used is

$$\begin{aligned} s_{ik} &= \text{seed matrix,} \\ A_i &= \text{overall adjustment factor for row } i, \\ B_k &= \text{overall adjustment factor for column } k. \end{aligned}$$

The seed matrix contains information about relative likelihoods of O-D pairs to be chosen by travelers. It may be a small-sample O-D matrix or an out-of-date O-D matrix. If no empirical seed matrix is available, a seed matrix can be created by an analyst to reflect information available on preferences between O-D pairs, as done by Furth (14) for vehicular traffic at an intersection.

The method is to alternately balance rows and columns to match the desired row and column totals until convergence. Initially, we set $t_{ik} = s_{ik}$. Then, for iteration h , rows are balanced:

$$t_{ik}^h = t_{ik}^{h-1} a_i^h \quad \text{for all rows } i \quad (3)$$

where the balancing factor a_i^h is the ratio of the desired row i total to the current row i total. Next, columns are likewise balanced:

$$t_{ik}^h = t_{ik}^{h-1} b_k^h \quad \text{for all columns } k \quad (4)$$

where the balancing factor b_k^h is the ratio of the desired column k total to the current column k total. Since balancing columns upsets the balance of the rows, the process is repeated until convergence is reached, that is, until, after balancing the columns, all the row totals agree (to some arbitrary tolerance) with the desired row totals. Reflecting the logic of the calculations, one name that has been used for this method is "iterative proportional fit." The name "biproportional method" derives from the form of the final estimate for cell (i, k) , which is

$$t_{ik} = s_{ik} A_i B_k \quad (5)$$

where the overall balancing factor for row i is $A_i = \prod_h a_i^h$ and the overall balancing factor for column k is $B_k = \prod_h b_k^h$. It is well known that the biproportional method has a unique solution (15,16). In general, there is no closed form or single-pass recursive algorithm for determining the overall balancing factors, which must therefore be found by an iterative method such as the iterative proportional fit.

The biproportional method has been derived in several different ways. Several authors, including Ben-Akiva et al. (8) and Lamond and Stewart (16), derive it as a case of minimizing a measure of discrepancy between the estimate and the seed. Hauer et al. (10) derive it as the most likely realization of a random (either Poisson or multinomial) process in which the seed represents the known occurrence rates. Ben-Akiva (12) derives it as the maximum likelihood estimate of the population trip rates, assuming that the seed is a random sample subject to sampling bias, and the relative bias is a product of two factors, one from the origin stop and one from the destination stop.

RECURSIVE METHOD AS SPECIAL CASE OF BIPROPORTIONAL METHOD

The estimates produced by Tsygalnitsky's recursive method are actually a biproportional form. Implicitly underlying the recursive model is a null seed containing information on whether travel is permitted or not, based on directionality and minimum trip length, given by

$$s_{ik} = \begin{cases} 1 & \text{if travel from } i \text{ to } k \text{ is permitted} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The recursive method also implies the following restrictions on the seed: $s_{in} = 1$ for all i , and for all $k < n$, $s_{i,k+1} = 1$ if $s_{ik} = 1$.

Theorem

The recursive method is a special case of the biproportional method in which the seed matrix is the null seed matrix given

by Equation 6. More specifically, the recursive estimates $t_{ik} = e_{ik} f_k$ (Equation 1) are equivalent to the biproportional estimates $t_{ik} = s_{ik} A_i B_k$ (Equation 5), where

$$A_i = e_{in} \quad \text{for all } i = 1, \dots, n \quad (7)$$

$$B_n = f_n = 1 \quad (8)$$

and

$$B_k = \frac{f_k}{\prod_{j=k}^{n-1} (1 - f_j)} \quad \text{for all } k = 1, \dots, n - 1 \quad (9)$$

Proof (by construction)

Because the biproportional method has a unique solution, it is sufficient to prove that estimates produced by the recursive method have a biproportional form. Consider column n (i.e., let $k = n$). By inspection, it is clear that Equations 1 and 5 are equivalent. Now consider column $n - 1$ (i.e., let $k = n - 1$). By construction, the recursive method yields $e_{i,n} = e_{i,n-1}(1 - f_{n-1})$ if travel from i to $n - 1$ is permitted. Rearranging, we obtain

$$e_{i,n-1} = \begin{cases} \frac{e_{in}}{1 - f_{n-1}} & \text{if travel from } i \text{ to } n - 1 \text{ is permitted} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Substituting for $e_{i,n-1}$ yields

$$t_{i,n-1} = s_{i,n-1} e_{in} \left(\frac{f_{n-1}}{1 - f_{n-1}} \right) \quad (11)$$

which is a biproportional form with the balancing factors given by Equations 7 and 9.

Now consider column $n - 2$ (i.e., $k = n - 2$). By similar argument,

$$e_{i,n-2} = \begin{cases} \frac{e_{i,n-1}}{1 - f_{n-2}} & \text{if travel from } i \text{ to } n - 2 \text{ is permitted} \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Combining Equations 11 and 12 yields

$$e_{i,n-2} = \begin{cases} e_{in} \frac{1}{(1 - f_{n-2})(1 - f_{n-1})} & \text{if travel from } i \text{ to } n - 2 \text{ is permitted} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Substituting for $e_{i,n-2}$ in Equation 1 with $k = n - 2$ again yields a biproportional form, with balancing factors given by Equations 7 and 9. Similar reasoning can be applied to each successive column $k = n - 3, n - 4, \dots$. Equation 13 becomes generalized to

$$e_{ik} = \begin{cases} \frac{e_{i,k+1}}{1 - f_k} = e_{in} \frac{1}{\prod_{j=k}^{n-1} (1 - f_j)} & \text{if travel from } i \text{ to } k \text{ is permitted} \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

from which the equivalence of Equations 1 and 5, using the substitutions given by Equations 7 and 9, is obvious. *Q.E.D.*

This theorem provides a framework for determining which method of transit route O-D matrix generation, the recursive or biproportional method, is better. Because the two methods differ only in which seed is used, the question can be reframed in terms of which seed is better, a null seed or a seed derived from exogenous data such as a small sample. We have already mentioned a few empirical studies of the methods. The remainder of this paper examines theoretical deficiencies of both the null seed and the small sample seed in common situations along with some experimentation, offering further guidance as to which seed is most appropriate in various situations.

One interesting corollary of this theorem is that when the seed has the form of a null seed, the recursive method provides a single-pass algorithm for finding the biproportional solution. Another corollary is that the recursive method is reversible; that is, it will yield the same results if one works backward or forward along the route. In this sense, the recursive method is not myopic like other intervening opportunities models. It appears to be myopic since it determines demand to stops along the route without explicitly considering what opportunities lie further downstream.

FACTORS AFFECTING APPLICABILITY OF RECURSIVE METHOD

The fact that the recursive method is the same as the biproportional method with a null seed helps indicate the types of routes and situations in which the recursive model can or cannot be expected to perform well. It can be expected to perform well when there is little a priori reason to believe that anything other than the popularity of the origin and destination stops is responsible for the demand for travel between O-D pairs. In the two situations suggested by Tsygalnitsky (13), a null seed appears plausible. These situations are (a) an express route with a collection segment outside the city and a distribution segment downtown, with travel permitted only between the collection and distribution segment; and (b) a short local route free from interference (e.g., competition) with other routes.

However, there are other situations in which a null seed violates a priori knowledge of trip-making behavior, the foremost being when significant competition from other routes affects demand. For example, imagine a local route between Segment A and Segment E, with several intermediate segments. If there is another route that goes express from Segment A to Segment E, we would expect that the express route would capture most of the demand from A to E. The seed matrix for the local route should therefore have a relatively low propensity for stop pairs that are served by the express route, rather than equal propensities throughout. Likewise, if two local routes begin at a common intersection uptown and end at a common location downtown and use different paths to get there, the travel market that can use either route will be split between the routes, lowering on both routes the propensity to travel between stop pairs served by both routes. Other network effects can affect travel propensity along a route as well. For example, a large transfer volume from a

feeder route can influence propensity for travel between that transfer point and other stops on the main route because these transferring passengers may have a high propensity to go to certain portions of the main route, but not to other portions (e.g., there may be a more expedient path to some portions of the main route than via that transfer point).

Long local routes may be another example of the unsuitability of a null seed. Travel propensity is commonly agreed to decline with distance, except for very short distances, where competition with walking yields the opposite effect. On short routes, travel time differences between different O-D pairs are sufficiently minor that an equal-propensity seed is still plausible. But on long routes, even if there is no competition from other routes, propensities should be expected to be smaller for long trips than for short trips. It has yet to be shown how long a route can be before the null seed assumption becomes unrealistic.

Results reported in the literature confirm these expectations. For example, Tsygalnitsky found that his method performed very well on the two routes he tested, one an express route with separate collection and distribution areas, the other a short local route. Simon and Furth (7) found that Tsygalnitsky's recursive method worked very well on a short local route, but that on a longer route with competition from express routes, it overpredicted very long trips. It should be noted that, because the average trip length is determined by the given on and off totals on the route, any model for O-D matrix generation must yield the correct average trip length. Therefore, an overprediction of long trips must be accompanied by an overprediction of short trips. Ben-Akiva et al. found that the recursive method overpredicts very long and very short trips, particularly on Route 350, a long route with competition from express service. It is not clear, however, whether the discrepancies on these longer routes arise because of interference from competing routes, from route length, or from using segment-level data.

O-D MATRIX ESTIMATION WITH SEGMENT-LEVEL DATA

An important factor affecting the applicability of the recursive method is whether the on-off counts are by individual stop or by segment (aggregations of stops). With stop-level data, travel along the diagonal of the O-D matrix (i.e., beginning and ending at the same stop) is not permissible, but with segment-level data, travel along the diagonal is permissible. Although the recursive method recognizes only the dichotomy permissible/not permissible (1 or 0), the possibility of intrasegment travel calls for a finer level of gradation. Intra-segment travel in a segment with n stops is an aggregation of n^2 stop pairs. If propensity is 1 for stop pairs for which travel is permissible and 0 for ineligible pairs, then the average propensity for stop pairs contained in that segment is at most $(n - 1)/2n$, which is less than 0.5. Similarly, average travel propensity between nearby segments can be less than 1 if the minimum travel distance is greater than one stop, because the pair of nearby segments could contain stop pairs that are ineligible for travel. In the example given in Figure 1, a null seed containing only 0's and 1's at the stop level is shown to be equivalent at the segment level to a matrix of average

a. Stop-Level Propensity
(Minimum travel distance = 3 stops)

To From		A					B		C			D	
		1	2	3	4	5	6	7	8	9	10	11	12
A	1	0	0	0	1	1	1	1	1	1	1	1	1
	2	0	0	0	0	1	1	1	1	1	1	1	1
	3	0	0	0	0	0	1	1	1	1	1	1	1
	4	0	0	0	0	0	0	1	1	1	1	1	1
	5	0	0	0	0	0	0	0	1	1	1	1	1
	6	0	0	0	0	0	0	0	0	1	1	1	1
B	7	0	0	0	0	0	0	0	0	0	1	1	1
	8	0	0	0	0	0	0	0	0	0	0	1	1
	9	0	0	0	0	0	0	0	0	0	0	0	1
C	10	0	0	0	0	0	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	0	0
D	12	0	0	0	0	0	0	0	0	0	0	0	0

b. Segment-Level Equivalent, Showing Average Propensities

To From		A					B		C			D	
		1	2	3	4	5	6	7	8	9	10	11	12
A	1												
	2												
	3												
	4	6 / 36					4 / 6		23 / 24			1	
	5												
	6												
B	7	0					0		2 / 4			1	
	8												
C	9	0					0		1 / 16			2 / 4	
	10												
	11												
D	12	0					0		0			0	

FIGURE 1 Stop-level null seed and segment-level equivalent.

propensities that include, besides 0's and 1's, fractional values ranging from 1/16 to 23/24.

Because the recursive algorithm itself does not permit fractional propensities, applying it to segment-level data will bias results, because this method forces all those fractional propensities to be 1's. For example, Ben-Akiva et al. apply the recursive method to segment-level data, setting the minimum travel distance to zero in order to make intrasegment travel permissible. As should be expected, they find that the method predicts too many intrasegment trips. In contrast, Simon and Furth (7) and Tsygalnitsky apply the recursive method at the stop level, although the results are often presented at the segment level, avoiding this bias. This effect no doubt accounts in part for the poor fit found by Ben-Akiva et al. using the recursive method.

When only segment-level data are available, a method of synthesizing O-D matrices that is consistent with the recursive method at the stop level is the biproportional method with a seed matrix consisting of segment-level average propensities. An example using data from Line 93 demonstrates how using this "equivalent null seed" avoids the large bias of a naive segment-level application of the recursive method. Table 1 shows four sets of results (presented at the segment level even if the analysis was done at the stop level): (a) the actual O-D matrix; (b) the stop-level estimate using a stop-level null seed (minimum trip length = 2 stops), which is the same as a recursive estimate; (c) the segment-level estimate made using the segment-level equivalent null seed; and (d) the segment-level estimate made using a naive null seed (minimum trip length = 0 segments), which is the same as a recursive estimate made at the segment level. Three different error measures are used: relative root-mean-square error (RRMSE), root-mean-weighted fractional error (RMWFE),

and χ^2 . The RMWFE can be used to judge whether the actual data obtained agree with the model. Formulas for these measures applied at the segment level are

$$RRMSE = \left\{ \frac{K}{t_{..}} \sum_{i=1}^n \sum_{j=i}^n [(\hat{t}_{ij} - t_{ij})^2] \right\}^{1/2} \tag{15}$$

$$RMWFE = \left\{ \frac{1}{t_{..}} \sum_{i=1}^n \sum_{j=i}^n \left[\frac{(\hat{t}_{ij} - t_{ij})^2}{t_{ij}} \right] \right\}^{1/2} \tag{16}$$

$$\chi^2 = \sum_{i=1}^n \sum_{j=i}^n \left[\frac{(\hat{t}_{ij} - t_{ij})^2}{\hat{t}_{ij}} \right] \tag{17}$$

where

- t_{ij} = passenger trips from i to j ,
- \hat{t}_{ij} = synthesized passenger trips from i to j ,
- $t_{..}$ = total passenger trips, and
- K = number of matrix cells containing permissible trips.

The segment-level estimate made using the equivalent null seed is almost as good as the stop-level estimate. The equivalent null seed minimizes aggregation bias, and aggregation error, as shown by the increase in the error measures, is small. The results are markedly worse when the recursive algorithm is applied naively at the segment level. The tendency in this case to predict too many very short and very long trips is clearly seen. The same analysis was performed for Line 16 and similar results were obtained.

An attempt was also made to assess the effect of aggregation bias on the tests performed by Ben-Akiva et al. (8). Three segment-level estimates for Route 77 outbound are shown in Table 2: the "best" estimate, a biproportional estimate generated using a small-sample O-D survey as a seed; the estimate using an equivalent null seed based on a minimum trip length of three stops; and the naive estimate using a segment-level null seed (minimum trip length = 0 segments). Because stop-level data were not available, it was impossible to generate a stop-level estimate and compare the results with the true distribution. Measures of error are in comparison with the best estimate. The equivalent null seed estimate approximates what would be obtained from a proper stop-level application of the recursive method. The comparison of these cases clearly shows how the naive segment-level application of the recursive method increases the estimated number of very short and very long trips.

SAMPLING ERROR AND BIAS WITH SMALL-SAMPLE SEED

It may seem that any empirical seed, whether from a small-sample O-D survey or an old O-D survey, would be superior to pleading ignorance and using a null seed. However, a null seed is not such a bad guess for many situations, being consistent with our understanding of travel behavior and having been confirmed on a few test routes. Before an empirical seed is used with the biproportional method or another iterative method, the value of its information content should be considered. Although information content can in many contexts be difficult to judge, in the case of a small-sample O-D matrix

TABLE 1 COMPARISON OF STOP-LEVEL AND SEGMENT-LEVEL ESTIMATES, LINE 93, a.m.

a. Actual O-D Matrix

From \ To	1	2	3	4	5	6	7	8	On
1	0	14	9	5	8	7	5	3	51
2		12	10	12	22	25	7	1	89
3			3	18	16	2	18	0	57
4				3	20	12	3	2	40
5					6	25	11	3	45
6						8	28	2	38
7							32	15	47
8								16	16
Off	0	26	22	38	72	79	104	42	383

b. Stop-Level Estimate

From \ To	1	2	3	4	5	6	7	8	On
1	0.0	11.6	6.4	7.6	10.3	8.3	5.5	1.4	51
2		14.4	12.1	13.8	20.0	15.6	10.7	2.4	89
3			3.5	11.8	17.2	12.4	9.4	2.7	57
4				4.8	14.0	12.9	6.9	1.4	40
5					10.5	19.5	12.2	2.7	45
6						10.3	25.6	2.2	38
7							33.7	13.3	47
8								16.0	16
Off	0	26	22	38	72	79	104	42	383

RRMSE = 0.352
 RMWFE = 0.479
 Chi Squared = 47.2

c. Segment-Level Estimate with Equivalent Null Seed

From \ To	1	2	3	4	5	6	7	8	On
1	0.0	14.1	5.8	7.0	9.9	7.6	5.5	0.9	51
2		11.9	12.2	14.8	20.7	15.9	11.6	2.0	89
3			4.0	12.0	16.9	13.0	9.5	1.6	57
4				4.2	14.7	11.4	8.3	1.4	40
5					9.8	18.9	13.9	2.3	45
6						12.1	22.1	3.8	38
7							33.0	14.0	47
8								16.0	16
Off	0	26	22	38	72	79	104	42	383

RRMSE = 0.368
 RMWFE = 0.513
 Chi Squared = 52.2

d. Segment-Level Estimate with Naive Null Seed

From \ To	1	2	3	4	5	6	7	8	On
1	0.0	9.5	5.3	7.3	10.6	8.9	7.5	1.9	51
2		16.5	9.3	12.7	18.5	15.6	13.1	3.3	89
3			7.3	10.0	14.6	12.2	10.6	2.6	57
4				8.0	11.7	9.9	8.3	2.1	40
5					16.5	13.9	11.7	2.9	45
6						18.5	15.6	3.9	38
7							37.6	9.4	47
8								16.0	16
Off	0	26	22	38	72	79	104	42	383

RRMSE = 0.534
 RMWFE = 0.641
 Chi Squared = 94.8

TABLE 2 COMPARISON OF SEGMENT-LEVEL ESTIMATES, ROUTE 77 OUTBOUND

a. Estimate using Small O-D Sample Seed

From \ To	1	2	3	4	5	6	7	On
1	0.0	14.6	16.2	3.2	1.6	1.6	92.4	130
2		6.5	56.7	6.5	13.0	84.3	298.1	465
3			9.7	3.2	1.6	38.9	228.5	282
4				6.5	3.2	4.9	197.7	212
5					0.0	9.7	110.2	120
6						40.5	367.8	408
7							0.0	0
Off	0	21	83	19	19	180	1295	1617

b. Segment-Level Estimate with Equivalent Null Seed

From \ To	1	2	3	4	5	6	7	On
1	0.0	11.8	16.2	2.4	2.1	14.6	82.9	130
2		9.2	57.9	9.4	8.0	57.0	323.4	465
3			8.9	6.1	5.5	39.2	222.3	282
4				1.1	3.4	31.1	176.4	212
5					0.0	15.4	104.6	120
6						22.7	385.3	408
7							0.0	0
Off	0	21	83	19	19	180	1295	1617

RRMSE = 0.197
 RMWFE = 0.427
 Chi Squared = 106.3

c. Segment-Level Estimate with Naive Null Seed

From \ To	1	2	3	4	5	6	7	On
1	0.0	4.6	12.2	2.2	1.9	13.3	95.8	130
2		16.4	43.5	7.8	7.0	47.6	342.7	465
3			27.3	4.9	4.4	29.9	215.4	282
4				4.1	3.6	24.9	179.3	212
5					2.1	14.4	103.5	120
6						49.8	358.2	408
7							0.0	0
Off	0	21	83	19	19	180	1295	1617

RRMSE = 0.242
 RMWFE = 0.407
 Chi Squared = 126.1

seed, the information content can be evaluated in terms of bias and sample size.

The main bias in O-D surveys is nonresponse bias, which is present if the response rate is substantially below 100 percent, a condition endemic to surveys on busy bus routes, and the nonresponding population is different in its O-D patterns from the responding population. The differences most often cited are as follows: nonresponders (a) are more likely to come from segments of the route in neighborhoods that have lower literacy or are less cooperative, or both; (b) are more likely to board where the route is crowded and they can't get a seat; and (c) are more likely to be making short trips, leaving them too little time to complete a survey. Fortunately, the first two biases are proportional to the response rates at each origin and each destination stop, and since the biproportional method correctly expands origin and destination totals, these biases disappear, as confirmed by Ben-Akiva (12). The third bias, however, remains, and can be significant, though its extent is hard to judge.

The effect of sample size on quality of information in an O-D matrix is also well known. A common rule of thumb is that an observation of fewer than five travelers in a cell is unreliable, since a difference of one or two people can effect an enormous relative change in the value. In the extreme case, a cell with no observations poses a special challenge, since a biproportional estimate for a cell must be zero if its seed value is zero. If a small-sample O-D matrix, aggregated to the segment level, where the segment is the level of the detail one is finally interested in, has a substantial number of cells with fewer than five observations, the information content of the seed may be so compromised by sampling error that it is worse than the information content of a null seed.

For example, the small-sample O-D surveys used by Ben-Akiva et al. (8) are all quite small, containing 61, 76, 138, and 54 responses for the four route/direction combinations studied. In the case with the greatest sample size, Route 77 outbound, only 8 of 25 segment-to-segment cells contain five or more observations, and six of these all lie in the same column of the matrix alighting at the last stop. Ten of the 25 cells contain no observations at all. An estimate based on such a seed seems risky.

Ben-Akiva et al. respond to the problem posed by cells with zero observations by offering a correction to deal with these "non-structural zeros." Even with this correction, estimates based on the empirical seed are heavily influenced by patterns that appear in the seed. Their estimate for Route 77 outbound made using this empirical seed (Table 2a, equivalent to their Table 3) contains the peculiar pattern in which, although there is substantial demand from Segments 1 to 7 (92 passengers) and from Segments 2 to 6 (84 passengers), there is virtually no demand from Segments 1 to 6 (1.6 passengers), because in the small-sample O-D survey, no one went from 1 to 6. In contrast, the estimate resulting from the equivalent null seed (Table 2b) has a much more typical pattern, assigning a far larger volume (14.6 passengers) to O-D pair 1-6. Because Route 77 is a short route and, at the time of data collection, had no significant competition from other routes, a null seed seems quite plausible. The question is whether the peculiar pattern found using the small-sample seed is a reflection of true patterns in the population, or just the spurious outcome of a random sampling process.

The effect of sample size can be addressed more rigorously. Ben-Akiva et al. provide equations for determining the approximate standard error of a biproportional estimate based on the number of observations in a cell, and also report approximate standard errors of their estimates. However, because many of their results are reported normalized to a standard grand total, the level of accuracy attained is not immediately apparent. Reversing the normalization, it was found that for the case of Route 77 outbound, the relative standard error of their estimates (standard error divided by estimate) is quite small (below 13 percent) for all six eligible cells in which the destination is Segment 7. These were the cells with many observations in the empirical seed. In the remaining 19 eligible cells of the matrix, the seed contained only 26 observations. Consequently, the relative standard error is greater than 100 percent in a majority of those cells. For the entire matrix, the average passenger volume per eligible cell is 17, and the average approximate standard error is 8.4. With a smaller sample size, as in the other three cases examined by Ben-Akiva et al. (8), errors can be substantially larger.

How large should a small-sample survey be for it to be a more reliable seed than an equivalent null seed for O-D matrix estimation? To explore this issue, a Monte Carlo simulation was conducted that repeatedly drew samples at random from the population of passengers surveyed on Lines 16 and 93, the two routes for which complete stop-level data were available. Simulated sampling was done without replacement, and there was no bias in the sampling process. Biproportional estimates were generated for various sample sizes using the simulated sample as a seed and compared with the true distribution. For each sample size, 100 to 200 repetitions were made. The final measure of fit reported is the RMSE as averaged across all cells and all repetitions for a given sample size. The results, shown in Figure 2, show how estimation error decreases with the sample size.

Also shown in Figure 2 is the RMSE resulting from a stop-level estimate using a null seed. While the three cases examined are too few to draw any firm conclusions, the results consistently show that a null seed is better than the ideal small-sample seed with fewer than 100 observations. When real-world sampling biases, response errors, and coding errors are accounted for, the sample size at which a null seed is as reliable as a small-sample survey will be still higher. On the other hand, if the route under study has significant competition from other routes, the null seed model is theoretically flawed, and so a small-sample seed will be preferred even with a sample size under 100.

CONCLUSIONS

Planners who need route level O-D matrices have had two primary approaches to use for generating O-D matrices from on-off counts. One is the biproportional and similar iterative methods that require a seed matrix containing information about relative preferences for O-D pairs. The seed matrix is usually a small-sample O-D matrix. The other technique, Tsygalnitsky's recursive method, is a computationally simple technique that requires no seed matrix. The recursive method is actually a special case of the biproportional model using a "null seed," a seed matrix in which entries are either zero or

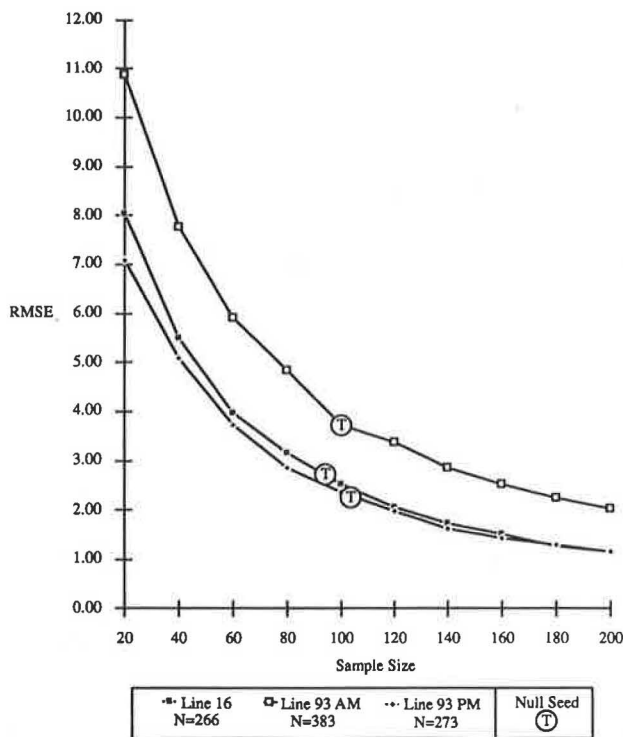


FIGURE 2 Estimation error versus sample size.

one, corresponding to whether travel between stop pairs is permissible based on direction of travel and minimum travel distance. A null seed is theoretically plausible on certain types of routes, such as relatively short routes with little interference (e.g., competition) from other routes. Empirical tests on different bus routes confirm this hypothesis.

The structure of the null seed underlying the recursive method implies that it is unsuitable for application to segment-level data. Instead, the biproportional method should be applied using an "equivalent null seed," a seed whose values are the average stop-level null seed propensity averaged over the stop pairs comprehended in a segment-level pair. This method yields results that closely approximate estimates made using the recursive method with stop-level data. It is probably the best method available for generating a transit route O-D matrix from segment-level data when there is no reliable small-sample survey or old O-D matrix to serve as a seed.

Finally, a comparison of estimation error using an equivalent null seed versus using a small O-D sample seed indicates, at least for the routes tested, that an ideal small-sample survey is preferable to a null seed when the sample size is over 100, and that a null seed is preferable when the sample size is smaller. In real-world applications, modifications to this

threshold should be made to account for imperfections in the sampling process and competition between routes.

REFERENCES

1. London Transport International Consultants, Inc. *Route, Schedule and Ridership Analysis*. Final Report/Technical Report to Dallas Area Rapid Transit. May 1987.
2. P. R. Stopher, L. Shillito, D. T. Grober, and H. M. Stopher. On-Board Bus Surveys: No Questions Asked. In *Transportation Research Record 1085*, TRB, National Research Council, Washington, D.C., 1985, pp. 50-57.
3. P. G. Furth. Short-Turning on Transit Routes. In *Transportation Research Record 1108*, TRB, National Research Council, Washington, D.C., 1988, pp. 42-52.
4. P. G. Furth. Zonal Route Design for Transit Corridors. *Transportation Science*, Vol. 20, No. 1, 1986, pp. 1-12.
5. P. G. Furth and F. B. Day. Transit Routing and Scheduling Strategies for Heavy Demand Corridors. In *Transportation Research Record 1011*, TRB, National Research Council, Washington, D.C., 1985, pp. 23-26.
6. P. G. Furth, F. B. Day, and J. P. Attanucci. *Operating Strategies for Major Radial Bus Routes*. Report DOT-I-84-27. UMTA, U.S. Department of Transportation, 1984.
7. J. Simon and P. G. Furth. Generating a Bus Route O-D Matrix from On-Off Data. *ASCE Journal of Transportation Engineering*, Vol. 111, No. 6, 1985, pp. 583-593.
8. M. Ben-Akiva, P. Macke, and P. S. Hsu. Alternative Methods to Estimate Route Level Trip Tables and Expand On-Board Surveys. In *Transportation Research Record 1037*, TRB, National Research Council, Washington, D.C., 1985, pp. 1-11.
9. P. G. Furth. Updating Ride Checks with Multiple Point Checks. In *Transportation Research Record 1209*, TRB, National Research Council, Washington, D.C., 1989, pp. 49-57.
10. E. Hauer, E. Pagitsas, and B. T. Shin. Estimation of Turning Flows from Automatic Counts. In *Transportation Research Record 795*, TRB, National Research Council, Washington, D.C., 1981, pp. 1-7.
11. M. G. H. Bell. The Estimation of an Origin-Destination Matrix from Traffic Counts. *Transportation Science*, Vol. 17, No. 2, 1983, pp. 198-217.
12. M. Ben-Akiva. Methods to Combine Different Data Sources and Estimate Origin-Destination Matrices. In *Transportation and Traffic Theory* (N. H. Gartner and N. H. M. Wilson, eds.), Elsevier, 1987.
13. S. Tsygalnitsky. *Simplified Methods for Transportation Planning*. M.S. thesis, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1977.
14. P. G. Furth. Model of Turning Movement Propensity. In *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990, pp. 195-204.
15. M. Bacharach. *Biproportional Matrices and Input-Output Change*. Cambridge University Press, London, 1970.
16. B. Lamond and N. F. Stewart. Bregman's Balancing Method. *Transportation Research*, Vol. 15B, No. 4, 1981, pp. 239-248.