can sit in a parked vehicle adjacent to the observation point. This same technique can be followed to record license plates as cars enter downtown parking garages.

2. Walking--An area such as a commercial strip that is congested with vehicles scattered throughout requires the data collector to walk. It is too difficult and often dangerous to drive under these conditions and effectively collect the data.

3. Moving vehicle--Large parking lots (e.g., regional shopping centers and commuter parking lots) that hold great numbers of vehicles can efficiently be surveyed from a moving vehicle. Traffic is usually light and the lot can be driven slowly to ensure accurate collection of data.

SURVEY METHODOLOGY

Two methods of recording data have been used effectively to date. These are forms and voice tape. Each has advantages and disadvantages that, in part, depend on the location as well as the individual collector.

1. Forms
   a. Use when small number of vehicles per site expected or small sample collected
   b. Maximum of 300 license plate numbers per hour can be recorded
   c. Requires less training and skill than voice recording
   d. Forms easily controlled and audited
   e. Writing, however, must be legible so as not to confuse 8 and B, 2 and Z, etc.
2. Voice tape
   a. Use when large sample required and vehicles concentrated
   b. 500 plate numbers per hour easily recorded
   c. More effective when collecting from moving vehicle
   d. Diction and enunciation very important (use words instead of letters--"Able" or "Apple" for A, etc., and say "Stop" after each license plate number)
   e. More equipment, greater expense, and sometimes technical problems
   f. Data preparation personnel must be trained to keypunch accurately from voice tape
   g. Greater potential for error

Cameras and hand-held keyboard entry directly to tape are also available but have not been used in VOS and thus cannot be evaluated at this time.

CONCLUSION

Experience to date has demonstrated that the VOS can provide a cost-efficient, highly useful data input to the overall transportation planning process, which can be further enhanced by follow-on surveys of motor vehicle owners. In addition, the ability to computer-map motor vehicle origin by census tract (or other small area) provides the professional and nonprofessional alike with an immediately understandable picture of the commuter and nonwork trip "marketplace" by specific destination. And when coupled with total vehicle ownership by census tract, demographics, etc., this service provides other measures such as "market penetration" relationship between public and private transportation use at the small geographic area level.

The service includes output tapes, statistical reports, and computer mapping. It is important to restate that name and address of registered owners are not available to the commercial or private sector and only available to the public sector when written approval is granted by the appropriate state motor vehicle authorities.

# Analysis of Employee Residential Locations for Transit Planning

RAI PARVATANENI AND TIMOTHY LAMBERT

**The development of a data base that describes the residential locations of employees working in the Detroit central business district (CBD) and adjoining major activity centers is described. The data base helped to conduct immediate and short-term transit service planning functions of the Southeastern Michigan Transportation Authority in the Detroit metropolitan area. This data-base development was undertaken because of the limitations of the existing sources that describe the work-related travel. Data describing the employee residential locations of selected major employers were gathered from personnel departments. The employers provided either an address list of their employees or summaries by zip-code locations. The residential locational descriptions of 33 555 employees for the CBD and 34 583 employees for the adjoining activity center represented sample rates of 31 and 52 percent of the total employment. An expansion methodology was developed and deployed to project sample data to the total employment population for 1980. Further, 1985 residential location projections were made by using the base-year data and regional population and employment-growth factors. The base-year location data at census-tract level for each employer or groups of employers and summaries for the total employment became valuable information in instituting peak-period route services; existing services were modified and route-effectiveness measures were developed. The base-year and 1985 data were also used in short-term transit service planning.**

The Southeastern Michigan Transportation Authority (SEMTA) plans, constructs, and operates public transportation facilities and services. Although the authority's area of jurisdiction covers the seven counties of southeastern Michigan, SEMTA primarily serves suburban to downtown Detroit commuter travelers and travel demands between suburban communities. Under a purchase-of-service agreement, SEMTA is also responsible for Detroit services operated by the City of Detroit Department of Transportation (DDOT).

Although SEMTA was created in 1967, the authority's operations actually began in 1971, with the first of several purchases of private carriers. Over the years, SEMTA ridership has steadily increased. Ridership since 1974 has increased at an annual rate of 13 percent from 7.1 million to more than 13.4 million annual passengers. Because of the trend toward ridership increases, SEMTA will have to carefully monitor and plan for future service improvements that will accommodate the potential ridership.

Currently, a significant portion of the travel market for SEMTA services is those traveling for work purposes. A recently conducted transit user survey indicates that approximately 90 percent of peak-period SEMTA users and 50 percent of DDOT users belong to this group. But only 25-30 percent of the total downtown-oriented commuting travelers use the transit service; therefore, potential exists to enlarge the transit market by attracting automobile users to public transportation.

To better serve commuter travel and increase transit ridership by diverting automobile users to public transportation, the current commuter travel behavior should be better understood. A review of existing information on commuter travel behavior showed many limitations for use in service planning.

Although the U.S. Census Bureau, through the decennial and annual housing surveys (1,2), provides information on the residential and employment ends of journey-to-work travel, there are several drawbacks for use in service planning. Primarily, the release of the census data, often three to four years from the survey date, makes these data less useful in route planning, which requires a more current data base.

Second, census data are gathered on a small-sample basis, which yields aggregate travel movements in the region. Although these data are at the analysis zonal level when released, they do not focus sufficiently on the trip end. That is, information on the commuter's work location is not specific to a particular establishment; rather, it is limited to respective analysis zones in the employment center. Although this allows trip patterns to be identified, marketing efforts at specific, high-potential employers are not possible. The Institute of Transportation Engineers Committee 6A-12 (3) examined the applications and limitations of the 1980 census data and recommended additional data-collection activities, including the employer surveys to supplement the census data.

Similarly, the home-based work travel data developed from the application of traditional travel-demand modeling chain would not provide the detail needed in service planning. The Transportation and Land Use Study (TALUS) (4), conducted in 1965, represents travel patterns now obsolete due to significant regional urban sprawl, varied energy supplies, and demographic changes during the past 15 years.

Because of these limitations and the recognized need to supplement the 1980 U.S. Census results, a data base describing the residential locations of employees in selected activity centers was developed to enhance the understanding of work travel that affects the design of transit services. This paper describes the procedures for collecting representative residential location data and a methodology for expanding the sample data to the total employment population in selected employment centers. The collection of the sample location data base was focused on major employers in the Detroit metropolitan region because of the large number of work trips generated by these firms. The application of the expansion methodology resulted in the estimated census-tract level residential locations of all employees working in the Detroit central business district (CBD) and adjacent central functions area (CFA). This distribution represents the total potential work-travel market for the delivery of public transportation.

This paper also describes a methodology to forecast 1985 residential locations of CBD/CFA employees by using the 1980 base-year data and regional population and employment-growth factors. Further, it provides a summary of the varied applications of the data in both immediate and short-term transit planning.

## STUDY AREA

Although public transportation is provided in all seven counties, the primary service area is the City of Detroit and the adjoining three-county area (Figure 1). Currently, the transit service outside this area mainly serves the elderly and the handicapped and, to some extent, internal travel within a few satellite cities.

As in most large U.S. cities, the maximum peak-period travel is oriented toward the most densely business-populated area of the region. This area, shown in Figure 2, covers two activity centers, namely, the Detroit CBD bounded by the freeways and the CFA adjoining the CBD on the north side. The employment densities are 125 547 employees/mile$^2$ and 24 015 employees/mile$^2$, respectively. This paper describes the collection of the employee residential location data through contacts with employers in these areas and the analysis of those data to develop the potential transit demand for the travel made from various points in the tricounty area to the two activity centers.

## COLLECTION AND PROCESSING OF EMPLOYEE RESIDENTIAL LOCATION DATA

Employee residential data were collected from major employers to produce a representative sample of various industrial employment categories. The approach for data collection consisted of a planning/marketing staff team that provided the employer with the following information in an arranged meeting and requested the residence location of their employees:

1. Description of existing transit services to the employment location,
2. Planned service improvements,
3. Company's opportunity to participate in planning, and
4. Company's opportunity to market public transit to its workers.

The interest exhibited by the employers in reaction to the team's marketing approach accelerated the rate of data collection and added to SEMTA's credibility in the business community. Requests were made for employees' home address lists (names deleted) with the street address, the name of the city/township, and the zip code. Data generation usually posed no problem, since most employer personnel files were computerized. However, some employers did not provide specific address data for reasons of confidentiality but did provide aggregated summaries of the number of employee residences in each zip-code area. A record of the information gathered for each employer is maintained on standardized forms. These contain the following information:

1. Business name and address,
2. Business contact person,
3. SEMTA contact person,
4. Contact date,
5. Level of aggregation (i.e., addresses versus zip codes), and
6. Format of data (i.e., hard-copy printout and/or magnetic tape).

Figure 1. Southeastern Michigan region.

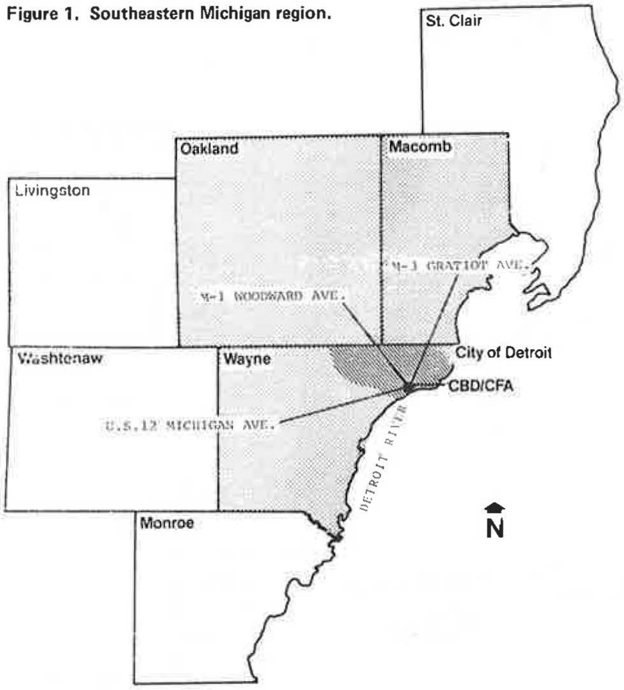
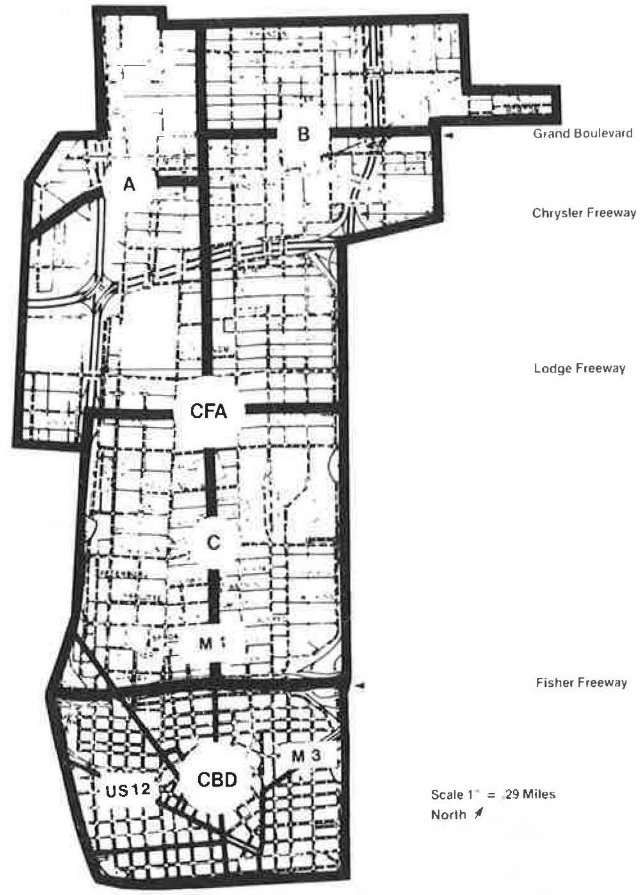



Figure 2. CBD and CFA activity-center boundaries.




For those employers who supplied address information and not the zip-code summaries, the data were summarized by zip code through manual tabulation. The reason for these tabulations of address data is that the data could be readily used in the service planning; a delay of about six months, which is the normal time needed for complete geoprocessing (i.e., assigning addresses to census tracts) of the data and preparation of summaries, would have made the data inadequate for use. To perform automated geoprocessing in a cost-effective way, it was better to accumulate several employer data sets.

The address information of six CBD employers was geoprocessed by using the U.S. Census ADMATCH computer program to allow analysis at a finer level of geography. The geoprocessing consists of associating the census tract number (1970 tract geography) with each of the address records by using the DIME Geographic Base File (GBF) and the ADMATCH programs. This resulted in an address-to-census tract match rate of approximately 75 percent.

Records were unmatched if the address was outside the immediate tricounty area, address input was misspelled, or the address had an inexact or new street identifier. The remainder of the unmatched records were manually geoprocessed to identify the census tracts. Then summaries of residence locations by 1970 census tracts were derived from the geoprocessed data.

The analysis reported in this paper used the employee residential location data for 6 employers at census-tract level, which accounted for 15 000 employees, and for another 10 employers at zip-code level, which accounted for another 17 000 employees. This represented a sample of more than 32 000 workers (30 percent) of 106 715 employees in the CBD. For the CFA, the data from 10 employers at zip-code level were used, which accounted for 34 583 employees (52 percent) of a total of 66 042.

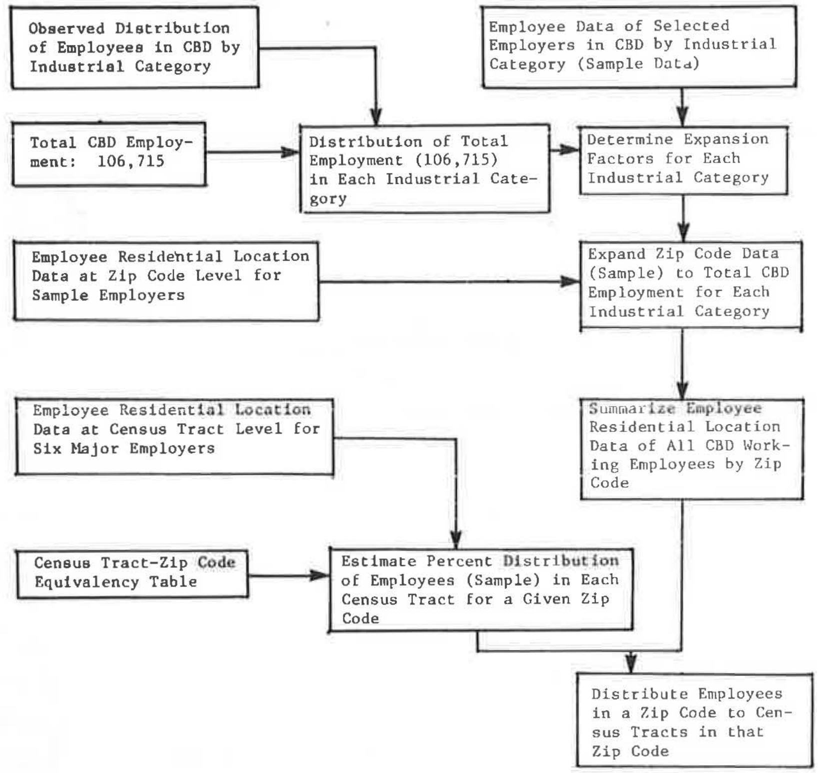
EMPLOYEE DATA EXPANSION FOR DETROIT CBD

To identify the potential travel-demand areas, the sample employee home location data were systematically expanded to the total CBD employee population. The purpose of this process was to develop total demand set at a disaggregate level (census tract) so that the data could be used directly in transit planning. Because the CBD and the CFA had distinctly different sample rates, separate methods were developed and used for expanding sample data to the total residential location distributions.

The expansion process is shown in the flowchart presented in Figure 3. The flowchart depicts the zip-code level expansion process from the sample data to the total CBD employment. The flowchart then shows the process by which these zip-code level data representing the total CBD employment were distributed to the census-tract level. The census-tract distribution was based on the distribution of the address data of the previously mentioned six CBD employers. The assumptions used in this process were as follows:

1. That major CBD employers in a single industrial category show similar employee home location distributions and that, conversely, the distributions would differ between types of industries; and
2. That the distribution of employee living patterns among various census tracts within a given zip code is not in the same proportion to the number of households in the census tracts or the number of persons in the tracts. (This hypothesis was validated by comparing the distribution of employee residential locations of selected employees who had supplied their address data against the distribution of households.)

Figure 3. Process for deriving residential locations of all CBD employees.




## Derivation of CBD Expansion Factors

A direct expansion of the sample data to the total employment was not appropriate because the sample data were an inaccurate representation of the total employment in the CBD and the respective proportions for each type of industry. This meant that an expansion within similar industrial categories was more appropriate.

The industrial classifications used in developing regional small-area forecasts were considered appropriate in this analysis. Some adjustments to these classifications consisted of further groupings of regional categories when there was insignificant sample size in any category. The final industry categories used were the following:

1. Manufacturing (automobile and other);
2. Transportation, communications, utilities;
3. Wholesale and retail trade;
4. Finance, banks, insurance;
5. Public administration; and
6. Natural resources, construction, business and professional services.

All major employers who provided location data were classified in one of the categories presented in Table 1. To maintain confidentiality, the names of employers are not shown.

Table 1 also lists the total number of employees from the sample in each category. The data from one automotive employer were excluded from Table 1 because this employer recently moved from the suburbs to the Detroit CBD and its residential locations were atypical.

Table 1 also estimates the total number of employees working in the CBD by industrial category. These estimates were based on the control total of 106 715 employees, distributed in each category, based on employment data from the Michigan Employment Security Commission (MESC).

The next column in Table 1 indicates the total number of employees, excluding the one automotive company and all federal government employees. A total of 5674 federal employees were estimated to be working in the CBD. The residential location data from these employers were not available, but home locations are most likely distributed throughout the region because, unlike local government employees, no residency requirements exist for federal employees. The last column in Table 1 lists the expansion factor to project sample data to the control totals within each industry category.

## Expansion Process at Zip-Code Level

The actual expansion from the sample data to the control totals within each category was performed at zip-code level; the expansion factors are those shown in Table 1. When more than one employer was listed in a single category, the numbers of employees within each zip code were combined. The process is illustrated in Table 2 for the industry category Finance, Banks, Insurance for the three zip codes 48015, 48026, and 48043.

The actual process was completed with a standard computer package, which allowed for an automated expansion. The process was performed with all the zip codes; this resulted in the expanded data set, which was then adjusted to include the location data of the automotive company and the federal employees.

## Distribution of Zip-Code Data to Census Tracts

The expanded residential data were further disaggregated to the census tracts based on the observed location distribution of the six CBD employers. These data gave the percentage distribution of employee residences among the census tracts within each of the zip codes. In order to derive the total regional distribution, a zip-code census-tract equivalency table was used.

Table 1. Derivation of CBD expansion factors.

| Industry Category | Sample Employer | No. of Employees in Sample | Observed Distribution[a] (%) | CBD Employees by Category | CBD Employees Without Automotive Company and Federal Employees | Expansion Factor |
|---|---|---|---|---|---|---|
| 1. Manufacturing (automobile and other) | Automotive company | 504 | 5.57 | 5 944 | 4 184[b] | 8.302 |
| 2. Transportation, communications, utilities | Utility 1, utility 2, utility 3 | 7826 | 10.3 | 10 992 | 10 992 | 1.405 |
| 3. Wholesale and retail trade | Retail store | 3705 | 13.4 | 14 300 | 14 300 | 3.859 |
| 4. Finance, banks, insurance | Bank 1, bank 2, health insurance company | 8236 | 23.7 | 25 291 | 25 291 | 3.071 |
| 5. Public administration | City government, county government | 8230 | 17.5 | 18 675 | 13 001[c] | 1.580 |
| 6. Natural resources, construction, business and professional services | Accounting company 1, accounting company 2, accounting company 3, hotel, engineering consulting company | 3887 | 29.46 | 31 438 | 31 438 | 8.087 |

[a]Observed distribution was derived based on Michigan Employment Security Commission (MESC) data.
[b]Total employment adjusted to exclude recently relocated automotive company (1760 employees).
[c]Total employment adjusted to exclude federal government employees (5674 employees).

Table 2. Zip-code level expansion process (CBD) for one industrial category.

| Zip Code | No. of Employees in Sample: Bank 1 | Health Insurance Company | Bank 2 | Total No. of Employees in Sample | Expansion Factor | After Expansion |
|---|---|---|---|---|---|---|
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| 48015 | 0 | 7 | 1 | 8 | 3.071 | 25 |
| 48026 | 3 | 30 | 9 | 42 | 3.071 | 129 |
| 48043 | 13 | 181 | 29 | 223 | 3.071 | 685 |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |

The data manipulation for deriving the percentage distribution of residence locations for each zip code in the region would have been a tedious process. However, computerized techniques made this distribution possible. The result of this process was the estimated residential location data at census-tract level for all CBD employees (102 723). The final data format is shown below:

| 1970 Census-Tract Number | No. of Employees Who Live in This Census Tract and Work in Detroit CBD |
|---|---|
| 1001.00 | 50 |
| 1001.01 | 111 |
| 1001.02 | 69 |
| . | . |
| . | . |
| . | . |
| 7115.0 | 22 |

The total employment as contained in the file was less than the control total of 106 715 for two reasons. First, there are employees who work in the CBD but live in Canada and outside the region. Second, the regional DIME/GBF file includes only the tricounty area; those working outside the tricounty area were not included.

## EMPLOYEE DATA EXPANSION FOR CFA

The CFA expansion process differed from that used for the CBD for two reasons. First, sample employee residential location data collected from employers accounted for 34 583 of the 66 042 total employees in the CFA. Since the sample consisted of more than 50 percent of the total, an assumption was made that the sample was representative of the total employment population and that errors due to simple expansion would be minimal.

Second, the CFA sample data gathered were not geoprocessed and were only available at the zip-code level; hence, within zip-code areas, the distribution of CBD employee residential locations by census tract was assumed to hold for the CFA employees also. During the follow-up analysis, this assumption will be tested after CFA address data have been geoprocessed. The expansion process performed on the data is described below.

### Derivation of Expansion Factors and Expansion for Zip Codes

Because the CFA included a large geographic area, it was divided into three analysis districts to classify data down to a level suitable for service planning. The expansion process was performed separately on each of the CFA districts designated A, B, and C (see Figure 2). The sample details for each of these districts are presented in Table 3. Also shown in the table are the expansion factors for Districts A and C, where

Expansion factor = (control employment totals)/(sample employment totals).

As shown in Table 3, the expansion factor for District A is 1.388 and for District C, 2.295. For these two districts the expansion was performed at the zip-code level from the sample to the totals in the same manner as in the CBD.

Since the sample employment total for District B was very small compared with the total, a direct expansion as above was not considered appropriate. The distribution of the entire CFA sample by zip codes was derived and used for District B. This process is illustrated in Table 4.

Thus, for each of the three CFA districts, the residential location of all employees by zip codes in the region was derived. All the data manipulations were performed by using a standard computer package program.

### Distribution of Expanded CFA Data from Zip Codes to Census Tracts

To further distribute the zip-code data to corresponding census tracts, data from the Detroit CBD

Table 3. Derivation of CFA expansion factors.

| District | Sample Employer | Sample Employment | Total Sample Employment by District | Adjusted Employment Totals | Expansion Factor |
|---|---|---|---|---|---|
| A | University staff | 16 024 | 23 827 | 33 068 | 1.388 |
| | Motor company | 5 069 | | | |
| | Computer company | 2 005 | | | |
| | Hospital 1 | 729 | | | |
| B | Hospital 2 | 1 660 | 1 795 | 12 412 | -[a] |
| | Art institute | 135 | | | |
| C | Hospital 3 | 3 723 | 8 961 | 20 562 | 2.295 |
| | Hospital 4 | 1 496 | | | |
| | Hospital 5 | 1 279 | | | |
| | University (medical staff) | 2 463 | | | |
| Total | | | 34 583 | 66 042 | |

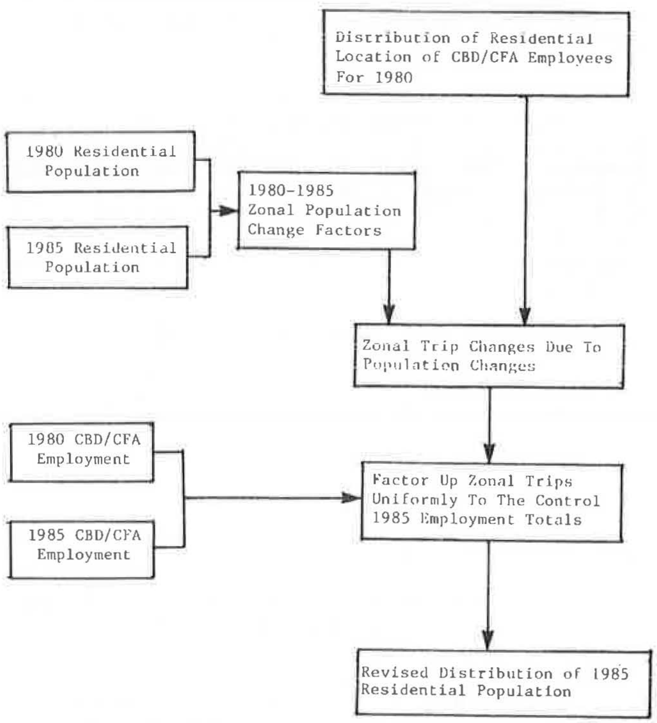
[a]Because the expansion factor was greater than 6, a different method was used for District B as explained in the text.

Table 4. Zip-code level expansion for CFA (District B).

| Zip Code | Total Sample Employment for All 10 Employers[a] | Distribution of Total[b] (%) | District B Employment Distributed[c] |
|---|---|---|---|
| . | . | . | . |
| . | . | . | . |
| 48015 | 607 | 0.017 | 221 |
| 48026 | 449 | 0.0130 | 162 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |

[a]Total sample employment = 34 583.
[b]Column 3 = column 2 divided by 34 583.
[c]Column 4 = column 3 times 12 412 (total District B employment).

Figure 4. Process for forecasting 1985 residential locations.




were used. This process consisted of determining the percent distribution of employment from the six major employer data sets for each of the census tracts in any given zip code and then distributing the total employment of that zip code to the census tracts based on the derived distribution. This process then yielded the employee residential location data at census-tract level for the CFA.

EMPLOYEE RESIDENTIAL LOCATION DATA FORECASTS FOR 1985

This section describes the methodology and its application for developing projections of residential locations for employees working in the CBD and the CFA for 1985 based on 1980 data.

The basic factors considered in the 1985 projection process were the following:

1. 1980 base employee residential location data files for the CBD and the CFA,
2. Employment growth rates for the CBD and the CFA from 1980 to 1985, and
3. Changes in population from 1980 to 1985.

The process used for projecting the 1985 data is illustrated in the flowchart (Figure 4). Based on the population shifts from 1980 to 1985, it was assumed that the residential locations of CBD and CFA working employees will exhibit shifts similar to that of the entire population. The adopted regional 1980 and 1985 small-area forecasts of population (number of people) were used to adjust the 1980 employee residential location data to the population shifts. The 1985 employment forecasts for the CBD and the CFA were determined to be 123 789 and 70 731, respectively. The adjusted 1980 employee residential location data were projected to reflect 1985 employment totals.

Preparation of Base Data

A computer file was created with employment and residential information for 1980. The file holds data for 1446 regional analysis units, and since regional population forecasts are based on these analysis units, the 1980 census-tract residential location data were converted to analysis-zone geography. A census-tract and analysis-zone equivalency file was used to do this.

The input data file variables in this process are as follows:

1. Analysis unit number,
2. 1980 CBD working employees who live in this zone,
3. 1980 CFA employees from District A who live in this zone,
4. 1980 CFA employees from District B who live in this zone,
5. 1980 CFA employees from District C who live in this zone,
6. 1980 population for this zone, and
7. Projected 1985 population for this zone.

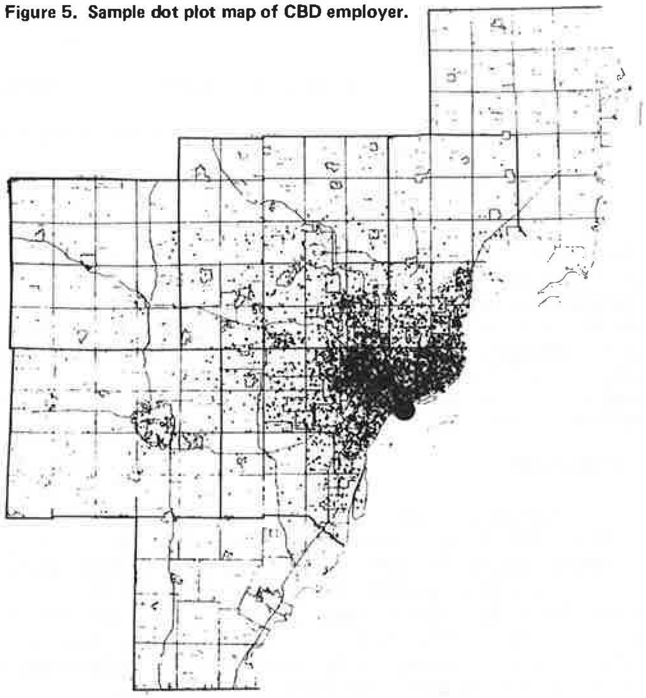


Figure 5. Sample dot plot map of CBD employer.

### Adjustments of Residential Data to Population Shifts

The basic assumption is that the employee residential locations would move in the same direction as general population shifts. If a zone experiences a reduction in population from 1980 to 1985, the employees who live in that zone and work in the CBD or the CFA will likely be reduced proportionately. Similarly, areas with increased population in 1985 will reflect an increase in employee residential locations. The adjustment factor was derived by dividing projected 1985 population by the 1980 population for each of the zones (total of 1446) in the file. Thus,

Adjustment factor (i) = [1985 population for analysis unit (i)]/[1980 population for analysis unit (i)],

where i ranges from 1 to 1446.

The adjustment factor was then applied to the 1980 residential location data. Multiplication of the 1980 data by the adjustment factor for a given zone reveals the adjusted number of employees in that zone. There is only one adjustment factor for each zone.

### Adjustment of Location Data to 1985 Control Employment Totals

The residential location data due to adjustments for the population shifts resulted in a data set with fewer employees in recognition of the decreasing regional population forecasts. The total employment observed in the adjusted data files was 98 175 for the CBD and 30 848 for District A, 11 636 for District P, and 19 337 for District C of the CFA. In spite of the decreases in the overall population for the study area, the employment for CBD and CFA is expected to increase from 1980 to 1985.

Based on the predicted regional employment-growth trends (5), the 1985 control employment was determined to be 123 789 for the CBD and 35 416 for District A, 13 293 for District B, and 22 022 for District C of the CFA. However, as observed in the expansion process for 1980, it was assumed that 3.7 and 2.2 percent of the CBD and the CFA employees will live outside the study area. Discounting for this factor, the CBD and the CFA control employment totals were 119 209 for the CBD and 34 634 for District A, 13 000 for District B, and 21 538 for District C of the CFA.

Expansion factors (EFs) were derived from the revised control totals and the adjusted employee data:

EF for CBD = 119 209/98 175 = 1.214.

EF for District A = 34 634/30 848 = 1.123.

EF for District B = 13 000/11 636 = 1.117.

EF for District C = 21 538/19 337 = 1.114.

The adjusted 1980 file for 1980-1985 differential population was factored up uniformly, based on the above expansion factors, which resulted in 1985 employee residential location data projections for each of the analysis units.

## APPLICATIONS OF EMPLOYEE RESIDENTIAL LOCATION DATA

SEMTA has used residential location data in many service-planning and corridor-analysis projects. Perhaps the most important result is that the previous planning data base and the transit demand estimation methodology, which employed population density of potential service areas as an indicator of certain types of travel habits, have been replaced. This new data base is much more efficient because it represents population densities of actual travelers with a known destination.

### Graphic Displays

Graphic displays effectively illustrate the residential location patterns and thus the travel patterns with CBD and CFA orientation. They simply and quickly lend a sense of the overall distribution pattern of the CBD and CFA employee populations. Two types of graphic displays have been extensively used in the transit planning. The first is a rough, inexpensive display of the number of employees per zip-code area. The number of employees in a zip-code area is represented by color coding; different colors indicate the varying density of employee residences per acre. The advantage of this method is that a single business or a group of businesses located in the same place or block can be studied for trends in residential patterns. However, this display is limited in that it does not effectively illustrate trip origin densities when the size of the zip-code areas varies.

One way to overcome this limitation is to develop a computerized dot-plotting program. This program uses the digitized census-tract or zip-code boundary coordinate files and randomly places a dot in the appropriate area for each employee (see Figure 5 for an example--a reduced version of the actual map, which is usually at a scale of 1:250 000). The dot-plotting and the color-coding techniques are two examples of graphic tools being employed in various industries. The variety of industries using these graphic capabilities has been well documented (6).

Although the dot-plotting method does give a better representation of the actual number of employees in each unit of analysis, it is also limited. In some areas, absolute numbers cannot be gleaned from the display when the computer overplots several dots in one place. This problem can be solved by using larger scale maps and smaller dots;

even with these limitations the display serves to illustrate general travel density trends. These displays are also useful in visually analyzing new and existing routes. The residential location data and the graphic displays have contributed to the development of several route-planning activities scheduled for execution within six months. The data have been used frequently to modify specific routes to increase ridership. Two examples of such efforts are presented below.

### Route Planning

A potential market on a major employment site was identified and served by establishing an express route. A major utility located on the western side of the CBD had no express route to the eastern suburbs. The residential location data analysis indicated that this company had a concentration of employees in the express service area. An extension of routing in the work-destination end of the trip resulted in a significant ridership increase on the route. Boarding counts at the utility site confirmed that this increase came primarily from that workforce.

SEMTA also rerouted a portion of a CBD local service by employing residential location data and displays. Originally, the service had a single-route configuration in the home (residential) end of this trip. Once the expanded data were derived, SEMTA staff observed that by branching (i.e., deviating route segments from the main routing), substantial numbers of potential riders would gain access to transit. The service area was widened by the route deviations, which put transit service closer to CBD and CFA employees' neighborhoods.

### Corridor Analysis

Employee residential location data are also adaptable to corridor service-planning projects, which extend to midrange (i.e., five-year) time periods. SEMTA reviewed the level of transit service provided to different areas of the region, projected the potential demand from the data set, and determined that the western suburbs were underserved, particularly by the park-and-ride commuter routes. SEMTA employed the Interactive Graphic Transit Design System (IGTDS) to analyze service to a single corridor. IGTDS is a set of computer programs developed by General Motors Corporation to assist planners in designing and evaluating transit alternatives by using computer graphics and analysis. With the location data demand set, alternative park-and-ride routes were tested, which yielded potential routes and park-and-ride lot locations.

Another example of an automated corridor analysis that uses the data is a feasibility analysis of a commuter rail system along another corridor in the region's northeast area. The modal-split models used the data to estimate ridership by various transportation modes. The modeling process consisted of validating the primary and submodal-split models by applying them to another corridor currently served by the commuter rail service for base 1980 conditions.

Model results were matched against the actual ridership on various transit modes, including commuter rail, to validate the model coefficients. After this step, they were applied to the study corridor and produced ridership estimates close to the actual ridership. These models were then applied to simulate the projected 1985 travel demand on each mode. Based on the analysis results, it seems that sufficient demand will be present in 1985 to support the proposed rail line.

Additional benefits derived from the team approach of data collection include the following:

1. Placement of a sales-ticket booth at a major employment center,
2. Establishment of sales agents at major employment sites, and
3. Enhancement of SEMTA's credibility in the regional business community.

The third point is especially important, since SEMTA pursues joint development projects and seeks to expand the employer base to employment centers not in the CBD or the CFA. Finally, nontransit benefits also resulted from the residential location data. Other government and private agencies have used these data to examine the regional demographic and economic trends.

## CONCLUSIONS

An up-to-date data base describing home-to-work travel demand is necessary and useful to conduct transit service-planning activities efficiently. Major employers helped to gather residential location data for their employees and to develop a travel-demand data base. The team data-collection efforts proved to be quite effective, as illustrated by the positive response from the employers.

The data gathered are the most current information at the level of geography needed in service planning. The data summaries and display techniques developed in this study provide insight and understanding of actual travel demand on a route and corridor basis. These summaries and displays have already been used for route modifications in the downtown area to effectively serve the employment center. A unique expansion methodology was developed to estimate total travel demand from sample data. Expanded employee residential location data at census-tract level and densities assisted in planning on the home end of the work travel. These data have also been used in service improvements and for the design of new express routes. Further, the base-year data and 1985 projections were used for short-term corridor planning.

The methodologies described in this paper are unique and do not involve hypothetical modeling theories. Rather, they provide a reliable, accurate, and up-to-date data base. To supplement the dynamic planning activities, the data base can be continuously monitored with little effort. In fact, because the data base has been widely and successfully used, the authority plans to expand its services to include other employment centers in the region. Efforts are also under way to provide a nonwork travel data base. The end product will be a comprehensive demand set that reflects current travel habits and will result in more sensitive and efficient transit planning.

## ACKNOWLEDGMENT

The opinions, findings, and conclusions expressed in this paper are ours and not necessarily those of the the Urban Mass Transportation Administration of the U.S. Department of Transportation.

## REFERENCES

1. Journey To Work. Bureau of the Census, U.S. Department of Commerce, Rept. PC 2-6D, 1973.
2. Selected Characteristics of Travel to Work in 20 Metropolitan Areas: 1977. Bureau of the Census, U.S. Department of Commerce, Rept. P-23, No. 105, Jan. 1981.

3. ITE Committee 6A-12. Preparation for the 1980 Census. ITE Journal, March 1979, pp. 41-47.
4. Profile of Southeastern Michigan TALUS Data. Detroit Regional Transportation and Land Use Study Committee, Detroit, MI, Aug. 1968.
5. Review of 1974 Version of Small-Area Forecast for Use in Alternative Analysis. Southeast Michigan Council of Governments, Detroit, MI, Oct. 1976.
6. H. Takeuchi and A.H. Schmidt. New Promise of Computer Graphics. Harvard Business Review, Jan.-Feb. 1980, pp. 122-131.

# Method for Determining and Reducing Nonresponse Bias

PETER R. STOPHER AND IRA M. SHESKIN

**Nonresponse bias is of continuing concern in participatory surveys of human subjects. It has led frequently to the adoption of expensive interview surveys in place of cheaper self-administered surveys because of relative response rates. Nonresponse bias has been estimated from comparison of early and late returns in self-administered surveys, from comparison of socioeconomic and demographic variables between the survey and census data, from special efforts to contact a sample of nonrespondents, and by assuming extreme values for nonrespondents. None of these methods is totally effective, whereas the relative economy of self-administered surveys has grown and suggests a reexamination of the value of such surveys. A method is outlined by using two survey mechanisms, including a conventional self-administered procedure, where the joint mechanism retains most of the economies but adds information on nonresponse and provides a means to increase response levels of the self-administered segment. Results from two transportation surveys are described and nonresponse biases and response levels are discussed.**

One of the first decisions in any survey design is to select the mechanism by which the survey will be performed. Input to this decision includes specification of the purpose of the survey, definition of the sampling frame, determination of desired confidence levels (and thus sample size), labor availability, time and budget constraints, types of questions that need to be asked, likelihood of obtaining accurate answers, length of the survey, and expected response rate (1). Each survey effort is to some extent unique and thus the choice among the face-to-face interview, the mail questionnaire, the telephone interview, and a number of other alternatives must be made for each survey by using a careful balancing procedure that considers the various advantages and disadvantages of each method.

One of the most important of these factors is the expected response rate because of the effects both on costs and on the unknown bias that a low response rate may introduce. More often than not, if respondents are placed in direct contact with an interviewer, the response rate is assumed to be high, generally on the quite strong grounds that refusal is less acceptable to a personal request than it might be to any impersonal approach such as a mail survey. In contrast, significantly lower response rates are assumed to occur when no personal request is involved or when the request is only to accept a survey form and not to answer specific questions. However, as Dillman (1) points out, this supposed significant advantage in response rate may be due, to some extent, to the manner in which response rates are calculated for the mail survey versus the face-to-face interview survey.

Irrespective of the survey mechanism, nonresponse occurs. It can be classified into two forms: genuine and nongenuine nonresponse (2). Genuine nonresponse is not the concern of this paper. This is defined as the nonresponse occasioned by selecting sampling units that are subsequently found to no longer be a part of the survey population (e.g., vacant or demolished houses, addresses that do not exist). In contrast, nongenuine nonresponse is defined as that nonresponse which occurs by the voluntary action of a sampled respondent not to participate in the survey. Genuine nonresponse is not of serious concern because it can be assumed generally to be a random or quasi-random occurrence that adds no significant bias to the survey data and that can be corrected largely by expanding the sample appropriately to cover its expected or encountered level (3). Nongenuine nonresponse is a documented source of bias for a number of reasons (4). It has been shown in a number of instances that those who do not respond to a survey possess generally a characteristic of direct relevance to survey measurements. For example, in surveys of travel habits and needs (an area well known to us), nonrespondents are most likely to be drawn from two segments of the population: those who travel very extensively and who therefore would be subject to much longer questioning on travel habits for a period such as 24 h and those who travel very little or not at all and who doubt the relevance of the survey to them or of themselves to the survey (5,6). This facet alone is a major cause of nonresponse bias. Others, which do not need elaboration here, include educational and income bias to written questionnaires and life-style biases associated with the state of being at home for the survey (1).

As a general rule, it can be assumed that the potential existence of and the extent of nonresponse bias caused by nongenuine nonresponse is correlated with the size of the nonresponse rate. Although it appears that little scientific evidence exists to support this hypothesis (particularly given the paucity of studies of nonresponse itself, let alone the biases and their relationship to rate), this assumption carries a fairly substantial weight of circumstantial common sense. For the purposes of this paper, it will be accepted as a reasonable postulate and not subject to further question.

Given, then, the parallel factors of an expected relationship between nonresponse bias and the common assertion that personally conducted surveys have higher response rates than impersonally conducted ones, it is not surprising that the majority of human surveys have tended to be carried out by means of direct interviewing in preference to most other methods of survey.

This paper raises three parallel concerns that derive from this state of affairs. First, some problems concerned with the calculation of response rates on face-to-face interview surveys versus mail surveys are discussed. Second, given the tremendous differences in unit costs of personal interviews