

Alternative Data Collection Options

George Wickstrom, *Kensington, Maryland*
Elaine Murakami, *Federal Highway Administration*

The current recommendation from the Census Bureau to Congress is for Census 2000 to include in the traditional long-form survey those data items used in transportation planning, such as income, vehicle availability, and journey to work information, and to begin full implementation of the continuous measurement process in 1999. In the review of alternatives to the census during this conference, two primary directions were discussed:

1. Replacing the long-form survey with the continuous measurement process: In the event that Congress determines that the Census 2000 form will be restricted to only those items needed for apportionment and the Voting Rights Act (age and race), the Census Bureau has developed an ongoing survey program that has the potential to replace the long-form data with 3- or 5-year accumulations of data for small geographic units and with annual data for large geographic units.

2. Losing the long-form data altogether: Congress is in a budget-cutting mood, and it is still unclear whether the long-form data will be included in Census 2000 or whether continuous measurement will be implemented beginning in 1999.

IMPLICATIONS

Quality of Data for Small Geographic Units

The quality of data for small geographic units achieved through the decennial census cannot be achieved as cost-effectively through any other alternative. The credibility of the Bureau of the Census permits the highest achievable response rate for any large governmental survey process. By including the sample questions, that is, the long form, in the process of the 100 percent count for apportionment, a high response rate is also achieved for the sample survey. This high response rate, and the sample size of one-sixth of the households nationwide, permits local and state governments to use high-quality data for small geographic units, such as census tracts and block groups, for program planning and implementation, site selection, cor-

ridor analysis, and so forth. The long-form sampling ratio, which varies between urban and rural areas, permits analysis in rural areas that would be compromised with the standard sampling ratio currently proposed for continuous measurement.

Consistency and Comparability

One of the greatest benefits of the decennial census is that the data are consistent and comparable with those from previous censuses. A recent FHWA document included tables from 1960, 1970, 1980, and 1990 censuses on journey-to-work characteristics in metropolitan areas and for the nation as a whole (1). Such trend data from a reliable and credible base with national consistency cannot be easily replaced without serious thought and testing.

The transportation planning profession would face a serious loss of investment in existing data and models, and would need to invest in new data sources and new or revised processes using these new data sources. Under the U.S. Department of Transportation (DOT) Travel Model Improvement Program (TMIP), still newer procedures are being developed that are heavily reliant on census baseline data.

Shifting the Cost Burden

The marginal cost of the long form is low. The current Census Bureau cost estimate for Census 2000 is \$3.9 billion. A 1995 National Research Council report (2, p. 127) states that the estimated marginal cost of the long form with sampling for nonresponse would be \$200 million to \$400 million. The transportation community could conceive of compromising by giving up sample size and reducing the sample from one-sixth of all households as done in 1990 to one-eighth of all households nationwide, but this is not likely to result in much cost savings.

Without the census long form (and also no guarantee that continuous measurement will be implemented), the data collection burden shifts to metropolitan planning organizations (MPOs) and states. In preparing for this conference, TRB conducted a survey of MPOs on use of 1990 census data. Although one-third of the agencies responded that they would probably replace the long-form survey with smaller regionally conducted surveys, two-thirds responded that they most likely would not replace the data. As alternatives to the census, larger agencies said that they were more likely to conduct regional surveys and traditionally have had multiple sources of transportation data. Smaller agencies, many of whom report that census data are not used in their transportation plans and programs, may be the most at risk for losing local small-area data. Consultants working on behalf of MPOs for small and medium-sized areas reiterated their reliance on the census long-form data for building new travel demand models. These smaller areas have much more flexibility and responsibility under the Intermodal Surface Transportation Efficiency Act (ISTEA).

ESSENTIAL DATA

Two critical uses of census long-form data in transportation are for population and employment forecasting and travel demand forecasting, used in the regional transportation planning and air-quality analysis required under ISTEA and the Clean Air Act Amendments. With forecasting horizons of 20 to 30 years, census data provide MPOs and state departments of transportation with a stable source of household characteristics and journey-to-work flow information.

Items seen as critical to obtain in a consistent, nationwide survey, with data reported for small geographic units, were household size, household composition, vehicles available to the household, and household income. (It is assumed that household size and composition will be available as part of the census 100 percent population and housing unit count.) There are now more cars than households and more cars than licensed drivers. It is important to know

the relationship of vehicle availability to household characteristics. Regional surveys cannot obtain accurate income data and typically ask for broad income category ranges such as less than \$20,000, \$20,000 to \$35,000, and so on. The specificity of income data and income sources in the decennial census is irreplaceable at the local level.

The next most critical data are work location and travel mode to work (including vehicle occupancy). These are used in travel demand model calibration and validation, and specifically in checking trip distribution and developing mode-split models. These uses were documented in the case studies presented at this conference, which may be found in Volume 2 of these proceedings.

Finally, those items that are currently included but could be replaced at the local level with small samples are departure time to work and total travel time.

ALTERNATIVES TO LONG FORM

The alternatives to the long-form survey that were discussed are as follows:

- Replace the long-form survey with continuous measurement as used in the American Community Survey (ACS),
- Expand the Nationwide Personal Transportation Survey (NPTS),
- Conduct regional household surveys,
- Conduct workplace surveys,
- Conduct on-board transit surveys, and
- Use secondary data sources, such as motor vehicle registrations and employment security files.

Replacement of Long Form with Continuous Measurement

The continuous measurement process as used in the ACS is currently being tested by the Bureau of the Census. This survey has been suggested as a replacement for the census long-form survey, so the decennial census would be limited to a headcount for congressional apportionment and for meeting the requirements of the Voting Rights Act. Other items currently collected on the long form would be moved to the ACS, a survey of 400,000 households each month (for 3 years, 1999 through 2001) to replace the long-form survey and subsequently to survey 250,000 households each month on an ongoing basis to provide annual estimates of population characteristics by state and congressional district. Small-area data (such as those for census tracts and traffic analysis zones) would be provided using accumulations of 3 years or 5 years of data collected during all 12 months of the year rather than on April 1 every 10 years. The Bureau of the Census sees one of their problems as having a budget that spikes once every 10 years and would prefer to present Congress with a budget that is more stable from year to year.

The current Census Bureau recommendation is to include the long-form survey in Census 2000 but to move toward use of administrative records in 2010. Although the transportation planning profession agrees that continuous measurement offers a wonderful opportunity for having annual data for large geographic units, the annual data on journey-to-work modes, vehicle ownership, and travel times are very valuable, especially at the national and large metropolitan area levels. However, conference participants were concerned that data for small geographic areas, especially the journey-to-work flows, might not be available because it would be dependent on continuous funding from Congress.

The transportation community would like a chance to evaluate the data from the continuous measurement test with a "hands-on" approach. The current continuous measurement test does not compare the long-form survey and continuous measurement directly, but is more an operational test of continuous measurement. (That is, results from a "point-in-time" larger sample are not being compared with accumulations of smaller samples over many

years.) It should be determined whether there are response rate differentials by urban and rural characteristics and other sociodemographic characteristics such as race and income and what impact these differentials may have on the final results.

In particular, the long-form survey is necessary for the year 2000 so that valid comparisons can be made with data collected by continuous measurement. It would be better to have the long-form survey and a 3-year test of continuous measurement in selected geographic areas rather than full implementation of continuous measurement simultaneously with the long-form survey. Average data for travel mode and journey-to-work flows should be examined to evaluate how annual data for large geographic areas and accumulated data (for 3 or 5 years) for small geographic areas can be used and to see where problems may occur.

Conference participants wondered why full implementation for continuous measurement is scheduled for 1999–2001 when it has not been determined whether it will work. The NRC panel, using estimates provided by the Census Bureau in 1993, reported that continuous measurement at 250,000 housing units per month would cost \$615 million over 10 years, but that this estimate was not high enough (2, p. 129). An estimate of the cost for conducting continuous measurement beginning with 400,000 housing units per month for 3 years and subsequently declining to 250,000 housing units per month has not been developed by the Census Bureau. Under the assumption that the unit costs are proportionately the same (\$5.14 million per month for 250,000 housing units per month and \$8.22 million per month for 400,000 housing units per month), the cost for the current proposal can be estimated over 10 years as \$728 million.

There were strong expressions that it seemed unrealistic to expect Congress to fund both the long-form survey in 2000 and full continuous measurement starting in 1999. The question was raised whether the Census Bureau was asking for much more than Congress would realistically fund in the hope that at least one of the alternatives would be selected.

Transportation planners are concerned about data access and cost. If the small-area data are available in new accumulations each year (one more year added), will MPOs and state departments of transportation want a new file each year? Probably not, but since some may want a new file in 2004 and some in 2006, it seems that it would be most efficient if DOT worked to get all MPO needs processed on a regular schedule in a nationwide coverage. Thus, instead of each MPO's requesting its own special tabulations, it might be better if a standard tabulation package (similar to the Census Transportation Planning Package) was worked out for each year.

Under the continuous measurement scenario, MPOs and other local government agencies would need to have an ongoing relationship with the Census Bureau to maintain the Master Address File and the Topologically Integrated Geographic Encoding and Referencing (TIGER) files and to geocode workplace locations.

Nationwide Personal Transportation Survey

Currently, NPTS is a sample of approximately 21,000 households (21,000 in the 1995–1996 survey and 22,000 in the 1990 survey). This sample size is sufficient for national travel trend analysis but insufficient for small-area geographic flow patterns. NPTS, conducted over a 12-month period, includes all trips made in daily travel, detailed information about vehicle use and acquisition, and, unlike the census, is not restricted to the journey to work. Also, because the current survey method is by telephone retrieval, the costs are close to \$150 per household, similar to regional household travel survey costs using trip and activity diaries with a telephone recruitment and retrieval method (P. Stopher, unpublished data, NCHRP Project 20-5, Topic 26-03, Methods for Household Travel Survey: mean costs by two survey methods, \$104–\$128 per completed household).

Potentially, the NPTS could be expanded to a large sample, sufficient to provide data for updating and validating trip generation rates, and conceptually could be expanded to a very large sample to provide small-area flow data. However, DOT has not planned or budgeted for such a large survey. One benefit of expanding the NPTS compared with some other alternatives would be that nationwide consistency in survey method and implementation would be retained.

Regionally Conducted Household Travel Surveys

An alternative to expanding the NPTS would be for each state and MPO to work out its own data needs. Currently, many MPOs supplement census data with small sample surveys, predominantly ranging in size from 1,000 to 3,000. The sample rate is approximately 1 in 300 households compared with 1 in 6 households for the long-form survey. If the census data become unavailable, surveys of 5 percent of the region's households might be conducted as they were in the 1950s and early 1960s. Each metropolitan area did its own survey. The goal was to measure the origin-destination flow using a matrix of approximately 100 to 200 zones. This was done in only the very largest metropolitan areas: Boston, Chicago, New York, Philadelphia, and San Francisco. The inclusion of the journey-to-work question in the 1960 census and the development of the first Urban Transportation Planning Package (UTPP) flow data for a metropolitan area were in part to reduce overall data collection costs, so that although the census may cost more because of the inclusion of long-form questions, local agencies are saving significant amounts of money by the ability to conduct very small sample surveys (less than 1 percent of households) rather than the 5 percent sample collected in the past. Including the journey-to-work questions in the census allows regional agencies to limit the sample sizes where all trips, not just the journey to work, are surveyed.

Workplace Surveys

Another alternative would be to survey people at their place of work rather than at their residence by conducting workplace surveys. This approach could also differentiate between workers and visitors at workplace locations. Commercial business files or Bureau of Labor Statistics files could be used to identify workplaces and sample employers to conduct journey-to-work surveys of employees. Some MPOs currently conduct these surveys at a regional level. The greatest difficulty has been getting cooperation from small employers; therefore, response bias is a problem. Because there are significant differences between large and small employers in such benefits as parking provision, transit, emergency-ride-home provision, and other related benefits such as on-site daycare programs, this response bias would need to be addressed.

Conducting these kinds of surveys at a national level rather than having each state or metropolitan area design and conduct its own would ensure consistency and comparability.

Transit On-Board Surveys

Transit on-board surveys have traditionally been conducted by most major transit systems. This information can give details about current transit users, but it is difficult to evaluate transit markets or potential riders. The combination of small geographic area data with household and person characteristics and mode choice in the census provides a wealth of data that is used in the short term for transit route planning and longer-range transit system development. On-board surveys do, however, permit better data on such items as access mode to transit and transfers. Surveys can be customized for specific transit systems, which may be known locally by acronyms and nicknames that currently cannot be easily included in a nationwide study. On-board transit surveys, similar to other surveys, also have been suffering from declining response rates and rising costs.

Secondary Data

Similar to the stated goal that the Census 2010 would use administrative records, some possible secondary data sources were explored. First, state motor vehicle registration files could be used. Some of the problems with these files are vehicles operated out of state, leased cars

whose owner is not the operator and who may be many states away, and identification of company fleets. Other possible secondary data sources are commercial employment data bases or state employment security department files. These data bases may contain total employment and may give employment by physical workplace locations. Experience by MPOs has varied; some states are more diligent about disaggregating large employers to work locations rather than one personnel office address, and some states do not permit other governmental agencies access to the files.

The biggest problem with using these secondary sources is the lack of combining characteristics; for example, motor vehicle files do not include household characteristics such as size, life-cycle stage, or income. Similarly, employment files include the industry code of the employer but do not include occupations of the workers or any information on their vehicle ownership or travel characteristics.

WILL ALTERNATIVES MEET THE NEEDS?

The consensus was that the transportation community was not yet ready to abandon the long-form survey. Eliminating the long-form survey was seen as shifting the burden of cost of data collection from the federal level to states and local government, not as an end savings to the general public. The alternatives identified left a loss of consistency and comparability.

The use and value of census products should be documented for the transportation community, and the case studies prepared for this conference are a good start. How the data are used in major investment studies, in developing long-range transportation plans, and in selecting projects for metropolitan and statewide TIPs should be shown. The support of the American Association of State Highway and Transportation Officials, the National Association of Regional Councils, the American Public Transit Association, and other organizations should be enlisted to support the data needs.

ACCESS TO DATA

Availability of data and flexibility in tabulation are the two key goals for access to all alternatives to decennial census data used in transportation planning. For either the decennial census or the ACS, the Bureau of the Census has recently been discussing the Data Access and Dissemination System (DADS), to replace most of the traditional standard census reports and products except those few profile reports necessary to show appreciation to the public for cooperation. DADS is envisioned to provide direct access to a limited number of data summaries, Public Use Microdata Samples, and a process for specifying special tabulations from confidential files. The goal is to provide tabulations rapidly and on demand.

DADS could potentially reduce the cost of producing standard and custom tabulations and provide tabulations more quickly to data users. Because users could define their own tabulations, there would not have to be predetermined and standardized tables. The transportation community needs assurances that DADS will accommodate requests not only for residence geography, but also for workplace geography and for flow (place of residence and place of work) tabulations. The ability to customize the tabulations would be beneficial. For example, income tabulations such as CTPP Part 1, Table 14 [Number of Workers (6) by Household Income (26)] or Part 1, Table 33 [Earnings of Workers (12) by Means of Transportation (11)] could use income groups appropriate to the area based on local median income values rather than using income ranges that cover a nationwide spectrum.

However, to be most expedient for the transportation community, it might be fastest for DOT to request standard tabulations for all states and all MPOs that might be a combination of a limited number of STF3 and CTPP tables so each state or MPO would not have to establish a DADS account or need to make a special request. Also, some priority might be given to a request by DOT compared with local governmental units should there be any problems in the ability of DADS to respond to tabulation requests in a timely manner. These stan-

Standard tabulations might be used to further determine custom tabulations that the local agency would need at some later point. Similarly, consultants working on behalf of local governments would be assured of some baseline data in which comparability with other areas is guaranteed.

Workshop participants also expressed the view that DADS could have negative impacts such as loss of control, reduction in data quality, and a higher impact on small MPOs because of greater limitations on their ability to access the data and user fees.

All alternatives to the census long-form survey or the ACS should have the same goals of data availability and flexibility in tabulation. BTS should examine other data dissemination programs to recommend a data access program for DOT.

REFERENCES

1. Rossetti, M., and B. Eversole. *Journey to Work Trends in the United States and Its Major Metropolitan Areas, 1960-1990*. Report FHWA-PL-94-012. FHWA, U.S. Department of Transportation, 1993.
2. Edmonston, B., and C. Schultze (eds.). *Modernizing the U.S. Census*. Panel on Census Requirements in the Year 2000 and Beyond, National Research Council, National Academy Press, Washington, D.C., 1995.