Introduction
Katherine F. Turnbull, Texas Transportation Institute

The Transportation Research Board (TRB) conference, Using Census Data for Transportation, was conducted October 25-27, 2011, at the Arnold and Mabel Beckman Center of the National Academies in Irvine, California. The conference was supported by the American Association of State Highway and Transportation Officials (AASHTO), Census Transportation Products; the Federal Highway Administration (FHWA), Office of Planning; and the Federal Transit Administration (FTA).

The conference brought together 115 individuals from transportation agencies and organization at the national, state, regional, and local levels, and from the private sector and academia. The conference provided a forum for participants to share experiences using Census data in transportation planning and decision making and to discuss opportunities, limitations, and challenges associated with using Census data, data available from the private sector, and data from global positioning systems (GPS) and other technologies. Participants were able to discuss research and training needs associated with applying Census data to transportation planning and decision making.

The conference began with general sessions highlighting activities at the Census Bureau and using data at the federal, state, and local levels. Breakout sessions included presentations on the use of Census data and data from other sources for a wide range of transportation applications. Two breakout discussion group sessions allowed participants to share their ideas and experiences, discuss challenges and opportunities, and identify research and training needs. The discussion breakout groups were organized around the five topics of content specification; integration with other sources and private sector data; data dissemination, data access, tools, and models; funding and institutional arrangements relating to the Census and alternative sources; and research and professional development. The final general sessions included an update on research activities and a data power user’s forum. The discussion group leaders summarized the key points and common themes in the closing session.

The full conference proceedings have been prepared and are in the TRB review process. This document contains the summaries of the breakout discussion groups presented in the closing session. These summaries may be of benefit to the sponsoring TRB committees, AASHTO, FHWA, FTA, and other groups in developing research problem statements and other follow up activities.
Reports from the Census Conference Breakout Groups

Breakout Group A: Data Content–Specification (Geographic Precision, Currency, Accuracy, Content)

Ed Christopher, Federal Highway Administration, Leader
Kristen Rohanna, San Diego Association of Governments, Co-Leader

The topic for Group A was data content specification – geographic precision, currency, accuracy, and content. The group was comprised mainly of representatives from MPOs. As a result, the conversation and discussion focused on the data needs of MPOs.

Participants indicated that the Census product they use the most is the 5-year ACS tract and block group data. They would like to use block data more. Participants noted they use the most recent available data and recently-released products. The 3-year ACS is used to examine county-level and large city flows. Participants noted the need for five-years of data for small areas.

Participants identified a number of different uses of Census and ACS data. Data used included population counts and household counts from the 2010 Census, the 3-year ACS flows, and the ACS household size. The 2010 Census household and population data provide the base for forecasts and estimates. Issues associated with the ACS household size data were discussed. The calculated ACS household size did not match the published household size data. The frequency of different data products was discussed. Participants noted the need for different tabulations every year, rather than full data sets and tabulations every year. A full data set every three years or every five years was identified as a viable alternative that would meet the majority of needs.

The group spent a lot of time discussing ACS household size data. The ACS provides data on the number of persons per household. If you divide the total number of people by the total number of households, however, you get a different number. While this difference is to be expected given the weighing of the data, the calculated persons per household tends to pass the reasonableness test and appears to be more appropriate than the ACS persons per household data. Examining this issue through a research project was suggested.

Participants identified household income as one of the most frequently used data elements. An issue was identified with household income in the 5-year ACS data at the block group level. The block group level data is provided uncollapsed. Aggregating the data and making collapsed categories can be problematic. Issues are encountered in aggregating the cells. It was noted that the Census Bureau has provided a footnote that more than four categories should not be collapsed together because the formula for calculating the margin of error will not compute correctly. Identifying examples of aggregating zonal ACS data for multiple block groups is needed.

Participants discussed concerns about using Census data and data from other sources, and providing data to other agencies and the public, when there are questions of accuracy. These
situations can put MPO staff in the position of having to defend data that they have questions about.

The need to incorporate the 2000 population geography of the TADs into the standard ACS product line was discussed. It is just another level of geography that is built from Census tracts. Participants could not identify any barriers to this request, which should be a no-cost item. Encouraging the transportation data community to promote this request would benefit all data user groups.

The group spent a lot of time discussing accuracy levels and margins of error. The Missouri State Data Center is color coding the margins of error included the various data tables. The margin of error categories highlighted are less than 15 percent of the estimate, 15 percent-to-35 percent of the estimate, and over 35 percent of the estimate. Participants discussed what margins of error could be accepted and what is a good margin of error. It was suggested that the accepted margin of error may vary depending on the variable and the level of geography. It was also noted that different terms are used, including margin of error, CV, and standard error. Additional research on this topic would be beneficial.

We discussed possible data sets and products that are not currently being provided by the Census Bureau. An asymmetrical work flow for multiple geographies was suggested as one desired product. The 5-year ACS asymmetrical flows with constant geography are being provided. Using different sized zones for residential data and workplace data is one option that would be worth exploring. The need for four-or-five dimensional tables in the CTPP was also suggested. Another suggestion was providing a fully perturbed data set for any type of cross tabulation at any geographic level. This data set could be used for modeling and many other applications.

A number of existing data gaps and data needs were identified by participants. Environmental justice data at small areas, including disability, environmental quality data, and public health data represent some examples of these needs. Better active transportation data, including bicycle and pedestrian data, was also identified as a need. It was noted that federal grant applications often require data that are not available. Better communication and coordination among federal agencies related to data requirements and availability was suggested.

Participants also discussed employment data, noting that LEHD data are not enough. Additional data on the universe of employees, the distribution of earnings by NAICS, the distribution of wages and earnings at the workplace, a variable identifying the relationship between individual earnings and households income, and EEO-like tabulations at smaller geographic levels to provide a better indication of transportation options by race and occupation were identified as beneficial.

Improving the American FactFinder tool was noted as important. With continuous sampling, the idea of considering the potential cost savings by not providing the same data products every year was suggested. The cost savings could be used to fund larger sample sizes.
Breakout Group B: Integration with Other Sources and Private Sector Data
Guy Rousseau, Atlanta Regional Commission, Leader
Nathan Erlbaum, New York State Department of Transportation, Co-Leader

Participants in the breakout group came from universities, a small MPO, a large MPO, a large DOT, a consulting firm, and a telematics data provider. The breakout group identified and discussed four major issues. These issues addressed alternative types of data sources, private sector data and potential risks associated with the lack of standardization and availability across all geographies, possible obstacles associated with the use of private sector data and methods to overcome these obstacles, and potential uses of LEHD data, including combining LEHD data with ACS data.

The discussion during the first session focused on using private sector data, including data from business establishments and credit collection agencies. Using marketing files with household characteristics available from private sources was discussed, along with the potential risks associated with using these data for addressing transportation issues. Many participants indicated there were good working relationships with state labor departments. Employment and socioeconomic data from these departments are used in transportation planning and modeling. Using CTPP data in rural transit planning was noted. Other participants noted that CTPP data was used to validate private sector data and that the cross-tabulations were of interest. The ability to share private sector data with others to validate results was noted as important.

During the second session participants spent more time discussing key transportation problems and how the data from different sources can best be used to identify and analyze possible solutions. An observation was made that there are many stable measures provided by the NHTS. Further, the CTPP, ACS, and LEHD can be used to provide aggregate geographic-level controls for expanding sample data resources. It was noted that the data being used today is partly a function of the legacy tools and models used in transportation planning, project selection, and operations.

Real-time telematics data, including data provided by GPS and other devices by both the private and the public sectors, bypasses the traditional trip generation, trip distribution, and trip assignment process, because it is already assigned to the network. Questions with these data include the sample size, the potential of self-selection biases, and other potential issues. Approaches for using data not intentionally collected for transportation uses to answer transportation-related questions were discussed. Potential issues and risks associated with using these data were noted. One participant noted the challenges of providing needed data to decision makers due to business cycle influences changing data collection schedules.

Research on using data from the private sector and emerging technologies in transportation planning and operations was identified as needed. Telematics and information technologies are changing rapidly. There is much more interest in visualization today and showing people the results of different alternatives graphically. The historic static approach of presenting information does not recognize available technologies.
It was suggested that we need to better understand how GPS and other telematics data can be used as a substitute for traditional transportation data, in combination with the Census and other data sources, and as a complete replacement for existing data sources. Telematics data can illustrate the real-time daily traffic flows, and adaptive behavior of motorists. Research to better understand the benefits and the limitations of these data would be beneficial.

There were two graduate students from Georgia Tech University in the breakout group. They are working on a very interesting project using employment data and marketing data from a private firm. The marketing file includes extensive data on household and individual characteristics. The project is examining how these data can be linked to other data sources, such as vehicle registration data from the Georgia Department of Motor Vehicles, and how to extract behavioral lifecycle cluster information. The desired result is a pseudo total travel survey-type record at a much lower cost than a traditional household travel survey. This approach could provide timely data and address many of the issues with traditional household surveys. How it all fits together is still a question, but it deserves a closer look.

It appears that private sector data providers and telematics data providers are willing and interested in working with public sector agencies, universities, and other groups to explore opportunities to use their data in transportation planning, project development, operations, and evaluation. We should follow up and explore how these data sources can be used and how they can be combined with data from the Census, NHTS, CTPP, and other sources. The potential to make significant advancements should be examined.

Breakout Group C: Data Dissemination, Data Access Tools, and Models

Keith Killough, Arizona Department of Transportation, Leader
Kathleen Lindquist, Washington State Department of Transportation, Co-Leader

Breakout Group C focused on models, data access tools, and data dissemination. We discussed issues and opportunities related to these topics, possible research projects to address identified concerns, and stakeholders for the research.

The first issue we discussed was potential improvements to data disclosure protection methods. A possible research project related to this issue is to identify the concerns, to examine the validity of these concerns, and to develop approaches for addressing valid concerns. The need for more research on the utility of perturbation versus suppression was discussed. The stakeholders for these issues and research include MPOs, state departments of transportation, and AASHTO.

The second issue we discussed was updating employer referencing systems. We discussed the need for improved coordination between stakeholders to avoid duplication of effort in compiling employer listings. A research synthesis examining how various regions are compiling and updating employer referencing lists would be beneficial. The possible coordination with the QCEW process could be explored as part of the research synthesis. The major stakeholders for this research are MPOs.
The third issue we discussed was estimating network travel distances. Participants noted that network travel distances or proxies are currently available from a number of sources, but that travel times by mode are equally important to many users. It was suggested that in the long term, private data sources, such as TomTom and INRIX, may provide the best option for obtaining needed data. Research on current methods and possible future directions was identified as beneficial. State departments of transportation and MPOs were identified as they key stakeholders for this research and data.

Coding workplaces when insufficient address data is provided represents the fourth issue we discussed. The suggestion to resolve this issue was cross checking multiple sources. The identified stakeholders for this information were states and MPOs.

Participants discussed the adequacy of current methods to disseminate the data. It was suggested that making the data available on the Internet was critical. It was noted that the American FactFinder is not very friendly. It was further suggested that obtaining a 3.0 version of American Fact Finder would benefit all groups. The CTPP Access Tool was noted as good, especially the Beyond 20/20 data dissemination and visualization software. Participants felt the CTPP Listserve works well and provides a community-based peer-to-peer system. It was further suggested that a Wiki could be created and used to share information and best practices. Other topics discussed included the CTPP Status Report, the need for face-to-face training, an electronic guidebook, a BTS bookstore, LEHD training, and the ACS Retrieval Tool.

Participants noted that this conference and the breakout groups provide an excellent example of the benefits of face-to-fact interaction. People shared their experiences during the discussion of various issues. Often when one person described a problem in their area, someone else was able to share information on how they had addressed the same or similar issues. The exchange of ideas and information was excellent.

We discussed possible process enhancements that could be made at the Census Bureau or elsewhere to obtain higher-quality data. It was suggested that implementing TAD geography would be beneficial. Better responsiveness to inquiries was also suggested, including enabling one-on-one contacts with Census Bureau staff. MPOs were identified as the key stakeholder for this support, but state departments of transportation and other Census user groups would also benefit.

Another issue discussed in the breakout group concerned whether or not the disclosure process was producing usable data. As noted by other speakers, concerns about accuracy are sometimes raised. Users may never know until there is a question about the data from outside groups. Clearly and visibility noting any data concerns or data limitations was suggested as an important improvement. The user community was identified as the stakeholders for these improvements.

Participants discussed the potential for further improvements to the place of work allocation. The need for additional validation was suggested. The state of Minnesota, which requires firms to accurately report multi-site branches, was noted as a good model for other
states. Stakeholders for this topic were state departments of transportation, MPOs, and other groups.

Similar to other breakout groups, we discussed the roles technology can play to promote dissemination, analysis, and presentation of census data for transportation planning, operation, and research. One suggestion was to fund the University Transportation Centers (UTCs) to act as data clearinghouses, providing one-stop locations for available data. It was noted that stewardship of the data is needed to ensure providing long-term, ongoing availability. The need for a permanent cloud storage or a data repository was suggested. The need for a Census Bureau equivalent of the BTS’ long-term, archival library and management system was also discussed, along with considering a web-based data language and organizing system, such as the PHP. The U.S. DOT was suggested as the main stakeholder for this concept.

The final topic we discussed related to making micro level census data more easily and inexpensively available to the transportation research community. Online training and NTI/NHI-type training was suggested. The stakeholders identified for this training included TRB, AMPO, NARC, APTA, APA, and AASHTO. A final long-term theme discussed by the participants was using technology to support data collection and data sharing, including interaction among different user groups.

Breakout Group D: Funding and Institutional Arrangements Relating to Census and Alternative Sources

Thomas J. Kane, Thomas J. Kane Consulting, Leader
Alan E. Pisarski, Consultant, Co-Leader

Breakout Group D discussed funding, institutional arrangements, and cooperation. Thinking of the agencies involved in the collection and dissemination of Census data for transportation purposes as a triangle, with the Census Bureau, the U.S. DOT, and AASHTO at the points of the triangle, was suggested. There are multiple issues from a communication and a coordination point of view. These issues are both internal to the agencies and organizations and external. The external communication and coordination issues can be considered both bilateral and trilateral.

We discussed internal issues within all three groups. There was agreement that fostering improved coordination within the Census Bureau would be beneficial to overcome internal stove pipes. A key example is the need to enhance communication and cooperation between the journey-to-work program and the LEHD program. Earlier speakers from the Census Bureau indicated that improvements in coordination among these two programs are being made. We look forward to seeing the results of this enhanced communication and coordination.

Another internal issue is better integration the data to respond to the Congressional mandates. Many participants were concerned about, what are in effect, unfunded mandates. We discussed performance measurement, performance monitoring, and the different aspects of the CTPP that are not now being considered programmatic responsibilities. One could suggest that Congress needs to consider these elements in the reauthorization.
We discussed the need for funding the transportation data programs on a more substantive and ongoing basis. Most of the transportation data programs are within the research portion of the U.S. DOT reauthorization. Moving the transportation programs into ongoing programs, rather than the research portion of the reauthorization might provide a more substantial and stable funding environment.

Participants discussed an inter-federal agency concern that the U.S. DOT needs to better represent the transportation data community in dealing with other federal agencies. Many of these agencies have an interest in transportation data and use transportation data, but have not necessarily funded or supported any data collection elements. We discussed the need for the U.S. DOT to take a more proactive role in gaining cooperation and support, including financial support where appropriate. A key point is that the transportation data community is not alone in the need for good data. Many other federal agencies also rely on transportation and Census data.

More collaboration within the AASHTO programs was also noted as important. The AASHTO programs are expanding in response to federal mandates and the needs of the various states. Realizing and coordinating the different data demands of these programs is important.

We discussed the opportunities for a peer exchange or a synthesis to better understand and document how state departments of transportation and MPOs are using, applying, and coordinating purchases of private sector data. This topic was also discussed in other breakout groups. Organizing the peer exchange or synthesis is an important institutional issue. It was suggested that AASHTO is the logical group to promote a better understanding of the full complement of data patterns and data users. These partners and users include state departments of transportation, MPOs, cities, counties, regional agencies, transit agencies, and other organizations. All of these groups need to be involved in a possible peer exchange.

There are other issues associated with the bilateral relationships – the relationships between the U.S. DOT and AASHTO, between the U.S. DOT and the Census Bureau, and between the Census Bureau and AASHTO. We discussed the need to strengthen the working relationship between the U.S. DOT and the Census Bureau. While there is an excellent working relationship between these two agencies, it needs to be broadened, strengthened, and expanded to higher levels. The same holds true for enhancing the working relationship between AASHTO and the Census Bureau. We discussed the need for regular communication at the top levels of all three organizations.

Funding issues, including obtaining needed resources for the transportation data programs and possible flexible funding, were discussed. Moving toward the standardizing of Census products represents one approach to addressing some of the funding issues. The weakness and limitations in the ability to transfer funds among programs within an agency, as well as between agencies was identified as a potential issue. A variety of possible funding sources and combinations were identified and discussed, including promoting funding flexibility and pooled fund projects.

A key theme through all of the items we discussed was the importance of collaboration and the importance of relationships. That message – collaboration and relationships – is critical.
for moving forward. The relationship between the U.S. DOT and AASHTO needs strengthening. The U.S. DOT needs to address collaboration and partnerships within its internal data programs in the RITA, BTS, FHWA, FTA, and the Office of the Secretary. Many of these items go beyond just the CTPP.

Breakout Group E: Research and Professional Development

Elaine Murakami, Federal Highway Administration, Leader
Alison K. Fields, U.S. Census Bureau, Co-Leader

Breakout Group E was tasked with examining research and professional development. The breakout group included representatives from state departments of transportation, MPOs, consultants, universities, the Census Bureau, and the FHWA. Even though there was only one person who used Census data for transit planning, there was a lot of discussion about mode choice, including transit, walk, and bike modes.

The charge to this breakout group was to consider avenues of research which could be implemented in the next three-to-five years to inform the transportation community. Participants discussed both research priorities and professional development needs. Many of the topics discussed were similar to those noted by the previous speakers.

Participants noted that Census data is often taken for granted. People assume that Census data will always be available, that it will always be reliable, and that it will include small areas. The increasing use of Census data for social justice, environmental justice, and LEP analyses was discussed, as well as the increasing use for population synthesis, microsimulation, and ABM at large MPOs. The CTPP three-way cross-tabulations for households and the ACS PUMS are of increasing importance in these efforts.

For most small MPOs, the Census data is the only source of population, socio-economic, and travel data. The home-to-work travel flows from the ACS and the CTPP are critical components for the traditional four-step travel demand models. Many of these areas have not conducted household travel surveys for over 25 years. As a result, there is an increasing reliance on Census data, the NHTS, and transferrable parameters from other areas for travel demand models.

The flow data from the CTPP are in high demand and are valued by the transportation community and other groups. There may be opportunities for the transportation community to work with other groups using the flow data to build partnerships to support the ACS and to leverage other data sources. There may also be opportunities to coordinate with the growing interest in freight data in many areas. Another opportunity discussed by participants is the use of private sector data and other alternative sources. Private sector data from INRIX, TomTom, GM On-Star, smartphone tracking, credit card tracking, and social media, such as Facebook and Twitter, may provide viable options that should be explored. Possible issues related to self-selection bias and coverage bias need to be examined with these data sources.

Participants identified and discussed a number of research needs. One research topic focused on synthetic microdata with a system for disclosure proofed user-defined tables.
2006-2010 CTPP will include more than 200 tables, of which, some will require the use of partially synthetic records. While some tables are used by many agencies and planners, other tables are rarely used. Participants noted it would be beneficial to develop a new system that would protect microdata behind a firewall, and establish a tabulation interface with disclosure checking. A GIS component for data visualization would also be beneficial.

One of the key issues to be researched is whether or not an ACS-based microdata summary system can include all the ACS variables, or whether a microdata summary system with a more limited set of variables for transportation would be more cost effective and easier to implement. Some existing research projects were discussed at the conference, including research at the Census Bureau using the Current Population Survey that Laura Zayatz is doing and work at Westat conducted for the National Center for Health Statistics. A test using the ACS data is not yet being conducted, but the Census Bureau test system and disclosure limitations rules would be applicable to ACS and other Census datasets.

Participants would also like to see contextual variables added to the microdata records, including neighborhood characteristics related to transit accessibility and land-use. Integrating the microdata access program with GIS was also viewed as important. Because individual confidentiality must be protected, the PUMA geography linked to the PUMS are large. For many transportation planning applications, the PUMA geography limits the ability to understand land-use, such as job density and shopping accessibility, and transportation accessibility impacts on travel choice, including travel mode choice (transit, walk, and bike).

Participants discussed the potential to attach small areas summary information that would be valuable for transportation planning applications to individual microdata records without compromising confidentiality. Research is needed to identify and prioritize contextual variables, and then to test how to modify the variables (e.g., ranks or scores, not discrete values) to maintain confidentiality while enhancing the utility of the microdata for transportation data users.

To implement the microdata access program, it would be beneficial to reach out to other users of the ACS flow data at the Census Bureau. The feasibility of an ACS microdata set that includes all the variables, as well as maintaining a microdata set with a limited number of variables for smaller geographies, should be explored.

The second research topic focuses on combining administrative records with the ACS. Conference speakers discussed some of the work the Census Bureau has initiated on combining the ACS with administrative records, including but not limited to, the LEHD, StARS, and potentially relying on the use of the Census Bureau Business Register. Ron Jarmin noted that the RDC program could be used to involve transportation planners working as research partners. It was noted that because the research project is already underway, the approval process for additional research should not take as long as previous attempts at conducting transportation-related research at RDCs. Participants suggested that the Census Bureau team working on this topic also consider combining other sources, including private data, the NHTS, and other public data. Participants felt that this project is very important and the transportation data community would like to be involved and assist if possible.
Another research topic is to examine all travel, not just commute trips. Additional data, including private data sources, will be needed in this effort. Participants were very interested in the role new technologies can play in the collection, analysis, and display of data on all types of travel. The upcoming International Travel Survey Methods Conference in Chile may provide additional insight into new technologies for conducting travel surveys. There was interest in the use of smart phones with GPS and in-vehicle navigation systems for capturing data on daily travel and examining travel over time. Traditional surveys have obtained travel information for individuals for one or two days. Multi-day data is needed for modeling and other analyses.

Participants also discussed the need for freight data. Transportation planners rely on the Census and the ACS for personal travel data and commute trips. Trucking firms and telematics companies represent a source of travel data. Purchasing data from these providers and combining it with Census data and data from other sources represents an option that should be explored more.

The need for high quality, detailed geography employment data was discussed as another research area. The results from the CTPP, private sources, such as Dun & Bradstreet, InfoUSA, and LODES can vary considerably at the detailed geographic scale. Participants identified a number of questions related to these data. Why is the correlation between LODES and private employment lists at the block level so low? Why are the trip length distributions between CTPP and NHTS so different from LODES? What is the imputation and allocation method used for LODES? What is the best way to identify, address, and correct specific types of known problems, such as headquarter concentrations where multiple worksites are not included, such as school district headquarters? Research addressing these questions would be beneficial.

Developing better relationships with the private sector to enable the fusion of public and private data sources was also identified as a need. Private sector data that have the potential to provide detailed origin/destination matrices may lack a link to personal and household characteristics that is typically obtained from the Census-ACS data. This link between the socio-demographic and specific travel behavior is critical to travel demand forecasts and microsimulation. Therefore, methods to use these private and public sector datasets are needed. This fusing requires innovative approaches to partnerships and cost-sharing, not just technical innovations. Research on this topic would be beneficial.

Another research topic focuses on evaluating the effectiveness and utility of the synthesized CTPP, using the method outlined in NCHRP project 08-70, which was discussed in one of the general sessions. The CTPP data product for 2006-2010 ACS data will provide synthesized data that will enable the Census Bureau to provide perturbed data at detailed geographic delineations that conform to Census Bureau disclosure limitation requirements. The perturbation procedures were developed and will be applied by Westat. After the CTPP is released and data users have had the ability to use the data, it will be instructive to evaluate how much the data perturbation has affected common transportation planning applications.

Participants identified a number of training and professional development activities. A key theme of many of these efforts was on knowledge transfer. Staff turnover continues to be a concern at many transportation agencies, especially smaller MPOs. With the ageing of the Baby
Boom generation, attracting new professionals to the transportation planning and modeling field is a priority at many agencies. Participants noted that there are more professionals with GIS expertise in transportation planning, but attracting professional with expertise in computer science was also suggested as important.

Much of this discussion focused on improving communication among professionals in different agencies and recognizing generational differences. The discussion of improving communication focused on getting personnel from state departments of transportation and MPOs to work together to ensure the right people have the right information about data, training and conferences, and research findings. This improved communication would increase consistency in planning and the utilization of data for transportation planning. The discussion of recognizing generational differences focused on the fact that newer employers may be more GIS savvy, but lack experience to judge the appropriateness or reasonableness of the data. Spatial data are increasingly easier to access, but are not always correct. The ability to use the tools does not imply the skills needed in evaluating the data.

One of the training needs data identified by the group was developing a way to describe data synthesis in a way that is understandable and that addresses privacy protected data. Protecting individual confidentiality is important. Identifying a method that is palatable and understandable would be beneficial.

Developing a Transportation Data 101 manual and delivering it through a Wiki was identified as a need. The Wikis used for the TRB Travel Survey Manual could serve as an example for the Transportation Data 101 manual. The manual would provide an overview of the basic data used in transportation planning, including the Census, the ACS, CTPP, the NHTS, LEHD, household travel surveys, traffic counts, and other sources. Finally, we discussed the need to better check market and brand synthetic data, explain the issues of privacy and disclosure, and describe its use. Adding a CTPP e-learning module would be a good way to address this need in the short-term. An NHI training course on this topic might also be appropriate.

Training on privacy protected/disclosure proofed (synthetic) data was noted as needed. The 2006-2010 CTPP will have disclosure proofing applied so that data at the Census tract and TAZ level can be released. The implementation of the NCHRP 08-79 project needs to be explained simply, with a discussion of data accessibility versus data suppression, privacy protection, and usability and reliability. Expanding the CTPP e-learning to include a specific module on this topic would be beneficial. A way to describe data synthesis needs to use language that is positive and will not induce questions about believability, and focus on the privacy protection would also be beneficial.

The need for training, including exercises, for judging data reasonability and utility was also suggested. Topics which could be included are how to evaluate if data should be used, how to evaluate if data sources do not agree with each other, and what to use if the sources do not agree. Another suggested training included highlighting best practices in data use by developing one or two case studies to document the data activities at an MPO or a state department of
transportation for planning applications. Updating a document from the U.S. DOT’s TMIP completed in the late 1990’s that showcased the Portland Metro data program was suggested.

Developing training in best practices in human capacity management was also suggested. This document could include methods such as using teams and cross-training to ensure that institutional knowledge is not lost. Another suggested idea was to use outside experts to provide training.

Finally, a suggested outreach activity was to develop an “elevator speech” or short list of discussion points that enumerate the highlights of critical data elements. This product could be in the form of a brochure to ensure that executives and upper management know the value of reliable socio-demographic travel data for completing planning activities.