

# **Census Transportation Planning Products Research - Assessing the Utility of the 2006-2010 CTPP Five-Year Data**

## Summary Report

*prepared for*

**American Association of State Highway and Transportation Officials**

*prepared by*

**Cambridge Systematics, Inc.**

*with*

**Kevin Tierney**



*report*

# Census Transportation Planning Products Research - Assessing the Utility of the 2006-2010 CTPP Five- Year Data

*prepared for*

**American Association of State Highway and Transportation Officials**

*prepared by*

**Cambridge Systematics, Inc.**  
115 South LaSalle Street, Suite 2200  
Chicago, IL 60603

*with*

Kevin Tierney

*date*

**January 31, 2016**

---



## Table of Contents

<b>1.0</b>	<b>Introduction .....</b>	<b>1-1</b>
<b>2.0</b>	<b>Issue Monitoring.....</b>	<b>2-1</b>
<b>3.0</b>	<b>User and Expert Opinions .....</b>	<b>3-1</b>
3.1	User Survey .....	3-1
3.2	Peer Exchange and CTPP Oversight Board Interviews .....	3-2
	ACS Based CTPP Considerations: Data Content .....	3-3
	ACS Based CTPP Considerations: Geographic Delineation.....	3-4
	ACS Based CTPP Considerations: Multi-Year Data Accumulation.....	3-4
	ACS Based CTPP Considerations: Margins of error .....	3-4
	ACS Based CTPP Considerations: Data Perturbation .....	3-5
	CTPP Data Access Software.....	3-5
3.3	Recommendations .....	3-6
<b>4.0</b>	<b>Practical Views of the Issues .....</b>	<b>4-1</b>
4.1	Current Use of the Data and Articulating Future Needs .....	4-1
	Long-term Census ACS Improvements .....	4-1
	Potential value-added enhancements to CTPP.....	4-2
	Additional multi-way demographic/socioeconomic tables and flow tabulations .....	4-2
4.2	Multi-Year Data Accumulation .....	4-2
4.3	Working with Large Margins of Error .....	4-3
4.4	CTPP Data Access Guide .....	4-4
4.5	Comparison of 2006-2010 CTPP and LEHD Worker Flows.....	4-4
<b>5.0</b>	<b>Summary and Looking into the Future .....</b>	<b>5-1</b>



## 1.0 Introduction

This study was performed to assess the utility of the 2006-2010 Census Transportation Planning Products (CTPP) data for the transportation community under the leadership of the American Association of State Highway and Transportation Officials (AASHTO). The main goal of the project was to better understand the extent to which issues related to the shift to the American Community Survey (ACS) from the Decennial Census Long Form are acknowledged and addressed in the dissemination and use of CTPP data.

As the Census Bureau has moved to the ACS format, the survey sample size has shrunk dramatically when compared to the now discontinued Long Form.

Since 2005, ACS data summaries are published annually for national, statewide, and large geographic areas with a population of 65,000 and greater. Summaries for smaller geographic areas including individual Census Tracts, counties and places are based on five-year accumulations of survey responses.

ACS data also include a measure of margins of error (MOE) associated with each estimate. The range created by the MOE around the estimate indicates the level of precision in measuring that estimate with a 90-percent confidence level. Since the migration to ACS, the Census Bureau reports all their estimates with MOEs and strongly recommends that users incorporate this uncertainty into their analyses.

Comparing multi-year data from the ACS with one-year data from the Census Long Form became more challenging due to the need to incorporate MOEs and the difference in timeframes where data from a single point in time are compared with estimates averaged over three or five years.

The changes in the data collection methodology for Census journey-to-work data raise additional issues:

- Because of small sample sizes and low response rates, the five-year CTPP data at small geographies has exhibited some quality issues.
- Aggregating zonal data and computing new measures using the existing estimates increases the amount of uncertainty which may have been typically overlooked.
- Due to confidentiality issues and regulations from the Census Bureau's Disclosure Review Board (DRB), 180 of the 343 tables (53 percent) from 2006-2010 CTPP contain perturbed data that are disclosure proofed. These perturbed data cells would otherwise be redacted by the Census Bureau to protect respondents' anonymity.
- It is anticipated that users would like to compare CTPP to another data source to gain confidence and test consistency across data sources. However, comparison data sources may also have issues of their own reflecting important differences in how they are produced.

The study consisted of four consecutive steps that are designed to gather information from various users with a varying degree of expertise and interest. The four steps are as follows:

- Issue Monitoring;
- User Survey;
- Peer Exchange; and
- Utility Assessment.

“Issue Monitoring” focused on identifying potential issues for practitioners and analysts. As a first step, a literature review was conducted and was supplemented by experiences of the data user community typically featured in the CTPP newsletter and in the issues presented to Federal Highway Administration’s (FHWA) CTPP Support Staff.

The “User Survey” targeted individuals with a higher degree of institutional knowledge at state and regional agencies who have more in-depth experience with CTPP and other relevant transportation data. The survey also collected information on various uses of the data.

The “Peer Exchange” stage collected feedback and suggestions from a panel of experts by filtering the survey results through their knowledge, incorporating their own experience with the issues they face, and identifying strategies to deal with those issues. The peer exchange participants outlined a set of key practical issues, provided guidance to address them, and suggested future directions for the state of practice and research.

The CTPP Oversight Board Interviews sought to address the same objectives with a set of expert CTPP users who have knowledge of both the CTPP user experience and with the program administration.

The report titled “[Information Gathering, User Survey and Peer Exchange](#)”, includes the findings of Tasks 1 and 2. More specifically, the report outlines a series of issues identified by users in Task 1, identifies and categorizes the ongoing uses of the CTPP and ACS five-year data, and presents the insights about users’ and experts’ experience with the data, available tools, and technical support.

The “Utility Assessment” step included a series of analyses to demonstrate a sample of practical implications of multi-year data compilation. The assessment included a review of using Longitudinal Employer-Household Dynamics (LEHD) worker flows, and featured a set of recommendations to facilitate addressing specific issues on accessing data and working with large margins of errors. The Task 3 Report titled “[Utility Assessment](#)” synthesizes the findings in these analysis stages, documents the key issues related to ACS based CTPP data utilization, and recommends steps to be taken in future CTPP data releases along with future research activities.



## 2.0 Issue Monitoring

In order to frame and categorize the ongoing uses of the Census Transportation Planning Products (CTPP) and American Community Survey (ACS) data by the transportation community, a focused review of recently published media was conducted as a first step for the project. The findings of this effort were used in the following tasks to identify potential focus areas for the data user surveys, peer exchange meeting, and in-depth interviews with CTPP Oversight Board members. The review included the CTPP listserv, CTPP newsletter, Transportation Research Board (TRB) documents, agency project reports, and insights from the FHWA CTPP Support Staff. The major topics that surfaced were:

- Limitations of the disclosure-proofing data perturbation;
- Small area workplace allocation problems;
- Workplace geocoding issues; and
- Small sample size challenges.

The information gathering exercise identified potential areas to focus in assessing the utility of the five-year CTPP. Task 2 sought to collect more detailed data via surveys and in-person interviews to allow the project team to explore users' views and experiences in greater detail. The set of key issues were also analyzed in Task 3 to portray the extent of potential impacts and to demonstrate how those impacts are realized at different parts of the nation.



## 3.0 User and Expert Opinions

This section summarizes the findings of a data collection effort that was made to better understand the perspectives of Census Transportation Planning Products (CTPP) users toward the five-year American Community Survey (ACS) based product and the ongoing uses of the CTPP and five-year ACS data. The information collected by the user survey was synthesized with the feedback received during the Peer Exchange and interviews with CTPP Oversight Board members.

### 3.1 User Survey

The web survey was conducted in September 2014 to shed light on the breadth of the CTPP and ACS data usage by planners at state departments of transportation (DOTs) and metropolitan planning organizations (MPOs). Potential survey respondents were contacted by e-mail and invited to go to a website to complete the survey. In addition, bulk invitations were sent to various email lists and planning organization mailing lists, including, among others, the following:

- CTPP listserv;
- Travel Model Improvement Program (TMIP) listserv;
- Association of Metropolitan Planning Organizations (AMPO); and
- National Association of Regional Councils (NARC).

The survey content was developed incorporating the key findings in the “Issue Monitoring” step, and covered the following topics:

- Respondent’s degree of familiarity with CTPP data
- Reasons that non-users have for not using CTPP data
- Use of CTPP
- Characterization of the importance of CTPP variables
- Use of CTPP related data products
- User perceptions of CTPP software and documentation
- CTPP issues

202 respondents participated in the survey. Nearly 63 percent of the respondents (124 out of 202) had hands-on experience with CTPP and/or Census data, and the remaining 37 percent of the respondents were labeled as non-users.

The respondents’ high level of CTPP familiarity and knowledge is exhibited in their answers. Nearly 80 percent of respondents agreed (either strongly agreed or somewhat agreed) that they have a good understanding of the Census ACS data collection processes. While, nearly two-thirds were involved with delineating traffic analysis zones for their regions or states.

Respondents were experienced with each of the CTPP general table types, and their experiences with CTPP data extended across the many recent (and not so recent) data products.

**Users** of CTPP data tables feel strongly that these data are valuable resources; more than three quarters of the respondents felt that the CTPP data tables provide a great deal of value compared to the ACS tables.

**Non-users** of the CTPP data indicated that the most important reasons for not having used the CTPP data products were circumstantial—they have not yet used the CTPP but expect to in the future and/or others within their organization use the CTPP so they have not needed to themselves. Much smaller percentages of non-users strongly agreed with the statements suggesting that the ACS CTPP data have problems which limit their usefulness.

Based on a check-all-that-apply question, about three-fourths of the respondents' organizations use CTPP data to support **travel demand modeling**. About half develop **data profiles and summaries**. More than 40 percent of respondents' organizations use the data to support **transit planning**, and almost 30 percent to analyze **bicycle/pedestrian issues**. Nearly one-third of the organizations use CTPP to support **environmental justice** analyses, and 20 percent to perform analyses involving **race and ethnicity**.

In addition to the multiple response questions about areas where CTPP data are used by organizations, we also asked respondents to describe their most recent usage of the ACS CTPP data in an open-response format. Nearly 50 percent reported that the most common recent uses of CTPP data involved simply obtaining household, workplace, or home-to-work flow data to better understand a **transportation market**. For 30 percent, the most recent use was to support **travel demand modeling**, and for the remaining 20 percent the most recent uses were for specific planning studies.

Next, the survey indicated that CTPP data were completely unhelpful for four percent of the analysis efforts. The CTPP data were able to completely support about 40 percent of the analyses reported by the respondents. However, for the majority of the efforts (57 percent) CTPP data did not provide everything that was needed to complete the analyses. The following main themes have risen up as the reasons for concern or dissatisfaction:

- Questions about data accuracy
- Issues with small sample sizes for small geographic areas
- Concerns with using data collected over multiple years
- Need for additional cross-tabulations
- Software issues.

Furthermore, the survey focused on a set of ACS-based considerations such as content; geographic delineation; multi-year data accumulation; and margins of error; and perturbation, and included a set of questions on data dissemination and training materials.

The Task 1 and 2 report provides in-depth discussions of survey responses for each issue outlined above. The next section synthesizes the survey findings for each issue incorporating peer exchange discussions and interviews with the CTPP Oversight Board Members.

## 3.2 Peer Exchange and CTPP Oversight Board Interviews

A selected sample of web survey respondents that were using the CTPP and ACS data, along with researchers identified in the literature search and CTPP users identified by AASHTO, were asked to participate in a half-day Peer Exchange to provide additional information and insights regarding ACS-based CTPP data. The Peer

Exchange Meeting was held with 16 participants from various sectors in the transportation industry and hosted by the Atlanta Regional Commission at their offices in Atlanta, Georgia on October 20, 2014.

The Peer Exchange participants were able to provide greater depth than the web survey respondents and to go into greater detail regarding ACS-based CTPP products with a particular emphasis on future planning of CTPP data products.

To further gain understanding of the perspectives of expert users, in-depth interviews were held with eight members of the AASHTO CTPP Oversight Board. These guided conversations with project team members and individual board members were held in November and December 2015. The discussions were loosely organized around the same issue themes as outlined in Section 3.1.

The Board Member interviews provided perspectives of people who have advanced knowledge of the CTPP data uses and products, as well as the challenges and issues of the CTPP program itself. The time gap between the peer exchange meeting and the Oversight Board interviews, while largely circumstantial, ended up being of benefit since new program issues regarding future Census table limitations arose during that time period.

### ACS Based CTPP Considerations: Data Content

Users identified the most and least frequently used variables and provided additional variables that they wish to see in the CTPP data.

The Peer Exchange participants reviewed the survey respondents' variable prioritization and brainstormed about opportunities to add new CTPP variables and tables. The Peer Exchange participants acknowledged that, like the survey respondents, most of their "wish list" items would require additional data or different data collection by the Census Bureau, so are probably infeasible. Participants discussed the value of nonwork travel and different delineation of workplace (multiple job holders, more relevant definition of part-time and full-time).

The Peer Exchange participants did note that it would be feasible to develop more three-way residence and workplace tabulations for CTPP without requiring changes in data collection. They also noted that there was the possibility to provide added-value tabulations such as home to work trip distances through the implementation of post processing.

Unlike the survey and Peer Exchange, the discussions with the Oversight Board interview respondents of CTPP content was colored by the Census Bureau's recent mandate to significantly reduce the number of CTPP tables for the next five-year tabulation. Because this reduction process was underway, the Oversight Board interview respondents discussed the process of selecting the tables to be dropped, rather than specific tables. In general, the Oversight Board interview respondents felt that reductions in tables could be made without greatly impacting the overall value of the CTPP, and that the essential tables for both large and small geographic delineations were being preserved, along with the most important flow information. Respondents were generally less concerned with the volume of available tables than with the data quality of the tables and with timely release of the data.

On the other hand, a few Oversight Board interview respondents raised a concern that the table cuts could have unforeseen consequences. One of the benefits of CTPP has been that it held information that planners might not have an immediately identifiable need, but that could arise in an unforeseen project or through a public or political request. The smaller CTPP will be less flexible than previous iterations of CTPP were.

## ACS Based CTPP Considerations: Geographic Delineation

Survey respondents use a range of geographic delineations in their CTPP analyses but small area data were seen to be essential by the users. Peer Exchange participants agreed with survey respondents that small area (TAZ level) data are essential for travel flow analyses and supported the idea of having demographic/ socioeconomic tabulations available at a more aggregate geography while providing travel flow data at the most detailed geographic level possible.

Unlike the general CTPP user population, the Oversight Board interview respondents generally expressed reservations about using the CTPP data at the TAZ level. Most of the Oversight Board interview respondent agencies had used TAZ level data for analyses, but these users raised concerns with Census TAZ-to-travel demand model TAZ inconsistencies and with the higher margins of error at these detailed geographic delineations.

Most of the Oversight Board interview respondents felt that tract level data were very important, because the geographic definitions were large enough to support the sampled data and detailed enough to support regional analyses and some project-level analyses. County level data were valuable in larger multi-county regions to provide useful summaries of general commuting trends, but were of little value in smaller regions.

## ACS Based CTPP Considerations: Multi-Year Data Accumulation

While not showing a strong support for using a rolling sample and compiling a multi-year data, the overall response does lean toward agreement that the continuous ACS data collection effort provides improved data quality. While they believe that multi-year accumulation of data makes analysis more difficult, more than half of the respondents stated that they treat multi-year accumulated data the same way as point estimate data for their analyses.

The Peer Exchange participants were more concerned about the effects of the multi-year accumulation than the web survey respondents. Many of the participants were concerned that five-year data collection period is problematic for regions that are rapidly changing, and proposed research to test potential problems introduced by multi-year data accumulation. Section 4.2 provides an overview of a set of analyses on this topic conducted during this study.

While recognizing that data usefulness is diminished because of the need to aggregate data across years, rather than having it for a specific point in time, the Oversight Board interview respondents were interested in having CTPP data releases be as frequent as possible. However, these expert users understood well why there needed to be accumulation of five years of data at the current ACS sampling rates to achieve reasonable estimates for small area geography and felt that a five-year frequency represented an improvement over the previous decennial based products, and that it should be continued.

Several Oversight Board interview respondents pointed out that just as important as CTPP frequency is the currency of the data. Minimizing the amount of time between the covered ACS data collection period and the data and product release was felt to be essential.

## ACS Based CTPP Considerations: Margins of error

More than 87 percent of respondents said they understood the margins of error concept. However, the users were split on the question of whether the accounting for margins-of error is making the analysis more difficult.

While almost 90 percent of the users claimed they understand the concept of margins of error and its computation, roughly half of survey respondents agreed that they generally use the CTPP estimates without accounting for margins of error. Only 20 percent of respondents feel strongly that they do address impacts of margin-of-error in their analyses.

Peer Exchange participants do pay attention to and are concerned by the generally large margins of error that accompany small area estimates, but like the web survey respondents, the estimates in the CTPP are often considered as the best estimate available regardless of the size of the margins of error. The participants expressed the need for guidance on the presentation of data with margins of error, specifically in the areas of graphical presentation, and obtaining policymaker acceptance of high and low estimates, particularly in the context of performance based planning.

Generally, the margins of error have been used by Oversight Board interview respondents qualitatively to evaluate the reasonableness of the CTPP estimates, to identify the need to use more aggregate geographic delineations, and in some cases to determine whether the estimates should not be used at all. Prior to ACS, users generally did not worry about the margins of error, but they now use them as measures of the quality of the estimates. Oversight Board interview respondents were generally familiar with several tools to help include margins of error in calculations and mapping, but the use of these tools was light. The respondents felt training on these tools and their techniques could be helpful for CTPP users.

### ACS Based CTPP Considerations: Data Perturbation

The survey respondents seemed to have some understanding of the general methods of disclosure proofing, but not a strong one. While only 13 percent of respondents strongly agree that they understand the disclosure proofing, most CTPP users somewhat agree or are neutral about the statement. Despite the challenges of dealing with the disclosure proofing, almost none of the survey respondents prefer CTPP data with suppressed values to CTPP data with disclosure proofed tabulations.

Peer Exchange participants also share similar uncertainties on how much perturbation changes the underlying data and would like to see simple results of comparisons between raw and disclosure proofed data.

Almost all of the Oversight Board interview respondents were of the mind that having perturbed data for smaller geographic delineations was preferable to not having the data at all. Provided the perturbation methods are tested and those that are doing the perturbation are familiar with how to do it, there appears to be a level of comfort with it. For the most part, Oversight Board interview respondents have been using the perturbed “B” tables without hesitation. They have been applying the same level of scrutiny of the general reasonableness of estimates to the “A” tables and the “B” tables.

### CTPP Data Access Software

The primary means of dissemination of CTPP data is the CTPP Data Access Software available online along with detailed documentation and supporting training materials. Only about one quarter of the survey respondents were regular users of the CTPP software (at least monthly). More than half the respondents used the CTPP software a few times over the previous few months. Only 8 percent viewed themselves as experts, and more than half of the respondents described themselves as skilled users. One third of the respondents viewed themselves as novice users. Most CTPP software users still find great value in the software. A sizeable portion of the respondents believe that the software has all the features that they need but the magnitude of the shares for “Neutral” and “Somewhat Disagree” levels points to room for improvement. Respondents were

also mixed in their assessments of the software's ease-of-use and documentation quality indicating a need for improvement as well.

The CTPP Peer Exchange participants, AASHTO staff, and the representative of the CTPP Software development team held a productive discussion about potential CTPP software improvements. Some of the desired features included:

- Automated Programming Interface (API) capability.
- Keyword search capability to allow users to identify table ids.
- Improved geographic aggregation, particularly map-based capabilities.
- Improved flow data visualization.
- Specific ease-of-use/intuitiveness issues with software.
- Coordination with the Census Bureau to get CTPP tables on American Fact Finder

The Oversight Board interview respondents were quite familiar with the software and with the CTPP training media through their obligations on the advisory board, rather than having a hands-on experience. The consensus opinion regarding the software is that it is quite powerful and flexible for users who have developed a level of comfort with it, but that there is a steep learning curve and this tends to be difficult for infrequent users.

One improvement suggestion in this regard is to recognize that the software could be designed to have different user experiences for frequent, more advanced users than for those who simply need a few pieces of data on an infrequent basis.

Oversight Board interview respondents found the great diversity of training opportunities to be very appealing. The respondents who had attended the in-person classes generally found them to be the most effective means of training, and some of the respondents pointed out the “marketing” value of these classes. Some Oversight Board interview respondents suggested that training be offered on what the CTPP data could be used for and how, so instead of only providing instructions on how to access the data that might be needed for a specific purpose, training should be offered on how the data would then be used.

### 3.3 Recommendations

The stakeholders were generally optimistic that CTPP could continue to provide users with extremely valuable data into the future. The first immediate challenge for CTPP identified by Oversight Board interview respondents is to maintain the CTPP's utility for users as the table reductions are made. Some respondents actually saw the reduced number of tables as possibly leading to a more streamlined and effective tool for planners, but it will be imperative that CTPP remain a “stable, reasonably comparable source from a large random sample of population” for it to maintain its core utility.

Both Peer Exchange participants and Oversight Board interview respondents cited the increasing importance of integrating CTPP and other common transportation planning datasets. The CTPP program can help this happen in several possible ways, including the provision of training on how to integrate CTPP and other datasets, the development of tools to enable users to perform these integration steps more easily, and/or the inclusion of other datasets in the CTPP program itself.



One unfortunate outcome of this has been that the different data sources have been put into competition with one another. For instance, planners are choosing to use LEHD data, and not CTPP, data because of the LEHD's more detailed geographic granularity, but are not accounting for the weaknesses in LEHD's government employment data or headquarters/branch geocoding issues. Likewise, there is growing use of analyzing cellphone probe data as a means to estimate journey-to-work flows without much consideration to the algorithms used to identify home and work locations. The Oversight Board interview respondents endorse the idea of using the different available data sources in conjunction with each other to develop improved estimates, rather than selecting just one data source.

In addition, Oversight Board interview respondents felt that there was an opportunity, and a need, to provide future CTPP users with additional guidance on how to use and interpret the CTPP data. For instance, a respondent cited the need for improved understanding of how to use multi-year estimates in travel demand model validation, which has traditionally relied on specific point-in-time data. Another respondent suggested that users could use guidance on the best ways to apply the CTPP data to develop descriptive analyses to support transit studies, environmental justice analyses, and bicycle/pedestrian planning efforts.

The Peer Exchange participants produced the following recommendations based on the expected future needs of users of CTPP:

- Streamlining by reducing the number of variables,
- Software improvement suggestions
- Greater availability of documentation of the program's table development activities,
- Long term expansion of CTPP to include access to other data sources and guidance on how to relate or combine CTPP data with other sources, and
- Research project ideas that are, described below.

Several specific research tracks were identified by the Peer Exchange participants based both on the outcomes of the Peer Exchange and the web survey:

- Compare CTPP data and alternative sources for model validation.
- Develop guidance and/or summary of practice on how to combine CTPP and other data sources, possibly including:
  - LODES,
  - National Household Travel Survey (NHTS),
  - regional and statewide travel surveys,
  - private source employment data,
  - private source cellphone data,
  - local agency employment and school data, and
- Extend ongoing research in combining CTPP and NHTS data, and
- Evaluate effects of household-based and person-based perturbation for future.

A few Oversight Board interview respondents suggested that the CTPP program (and the Census Bureau) would benefit from research into likely future trends in commuting. Since the Census Bureau will need many years to make changes to capture innovations and commuting market advances, the transportation community

should be seeking to prepare the Bureau for future trends that may be important (perhaps, such as automated vehicles and ride hailing services). In addition, research into the changing nature of the home-to-work commute could lead to the need for further changes far down the road.

## 4.0 Practical Views of the Issues

The prior tasks identified the key issues related to ACS-based CTPP data utilization. This section summarizes the outcomes of the “Utility Assessment” effort and provides further practical information and guidance to CTPP users by addressing some of the key challenges users may experience when accessing and using the data for various purposes. These activities include, but are not limited to the following:

- Revisiting the current use of the data and articulating future needs,
- Demonstrating how 1-year ACS data products vary by time and geography,
- Summarizing methods to work with large margins of error (MOE),
- Compiling useful tools and approaches that facilitate the use of CTPP, and
- Comparing CTPP with other similar data sources.

The findings from each of these steps are summarized below.

### 4.1 Current Use of the Data and Articulating Future Needs

Typical uses of the ACS data included characterization of households, workplaces and worker flows, intercity analyses and travel demand modeling activities such as calibration and validation of trip generation and trip distribution models, and population synthesis. Other unique uses included the following:

- Smaller and mid-size MPOs that do not have travel surveys produced trip distribution data,
- Unique three-way data tabulations provided by CTPP were used for household travel survey sampling and expansion,
- Data from successive CTPP releases were compared to track changes over time,
- Travel to regional special trip attractions,
- ACS data to support corridor studies and environmental justice analysis, and
- Small travel markets such bike to work flows and travel by self employed respondents were analyzed.

The future needs were categorized into three broad categories:

- Long-term Census ACS Improvements,
- Potential value-added enhancements to CTPP, and
- Additional multi-way demographic/socioeconomic tables and flow tabulations.

#### Long-term Census ACS Improvements

Many users viewed the CTPP development process as the primary means for the transportation community to collaborate with the Census Bureau and to influence future ACS data collection. They helped identify the following potential areas for exploration:

- Second Jobs,

- Better Information on telephone availability (e.g., Cell-phone only households),
- Sub modes such as access/egress modes,
- New modes (e.g., shared ride arrangements, services offered by transportation network companies (TNCs), autonomous vehicles), and
- International commuting trips for border communities.

### Potential value-added enhancements to CTPP

The transportation data community also identified a set of potential value added improvements particularly on data dissemination to help users access multiple data sets and facilitate data fusion with other sources, such as LEHD/LODES and NHTS data. In addition, supplementing the data with travel distance information during the data compilation stage was suggested. This may require additional effort and coordination with national and local agencies, however, online travel time estimates can be used as a proof of concept.

### Additional multi-way demographic/socioeconomic tables and flow tabulations

CTPP data users identified a few potential ACS-based cross-tabulations and travel flow tabulations that would be helpful for specific analyses they recently performed or would like to perform. Driven by advanced travel demand modeling needs, these users aim to understand travel behavior needs at a person level via additional tabulations such as person level characteristics, age, gender, employment, occupation, earnings; vehicle parity with workers, adults and or household size; school enrollment and part-time arrangements to work; and internet access/use.

Some users suggested including additional tabulations using poverty and disability status. However, complexities associated with data compilation and interpretation of poverty status and the small size of the population with certain disabilities make incorporation of these items to the CTPP infeasible in the short term.

## 4.2 Multi-Year Data Accumulation

One of the concerns raised during the peer exchange relates to the five-year data accumulation that reflects the ACS data collection method. Panelists discussed how well the five-year estimates represent the conditions through the five-year period, since the 2006-2010 period included the 2008 -2009 economic crisis.

County level comparisons were conducted for six counties which had unique patterns of growth during the 2006-2010 period. These counties are:

- Louisiana, Orleans Parish County,
- New York, Bronx County,
- Pennsylvania, Lackawanna County,
- Arizona, Pinal County,
- California, San Diego County, and
- Michigan, Wayne County.

Six tables from ACS one-year data were extracted for each year in the 2006-2010 periods and contrasted with the corresponding five-year data:

- Household Size by Vehicles Available,
- Household Income in the past 12 months,
- Household Size by Number of Workers In the Household,
- Employment Status,
- Travel Time to Work, and
- Travel Mode to Work.

Based on the analysis of five-year CTPP and five-year ACS data compared to the individual one-year data sets for the same period, the following general observations can be made:

- Estimates from the five-year data in most cases provide the expected “average effect” level for that time period.
- Most five-year estimates fall near the one-year estimates for the years 2008 and 2009.
- For regions where there are major differences over time or a clear trend between 2006 and 2010, the five-year estimate would not provide a precise estimate.
- For counties that experienced significant growth such as the New Orleans Parish, five-year estimates would generally underestimate the measures being analyzed.
- For areas with substantial loss of population and employment, such as Wayne County, MI, the five-year estimates were influenced by the pre-recession data collected earlier in the five-year period (2006 and 2007). The use of the five-year average data for individual year estimates between 2008 and 2010 would result in large errors.

In general, comparing five-year estimates with five-year trends provided useful insights especially for areas that experienced unique changes in socioeconomics. CTPP users may benefit from conducting similar analysis and supplementing their findings with local knowledge to better understand how the five-year data stand within the duration they have been collected.

### 4.3 Working with Large Margins of Error

Typical uses of the CTPP data generally encompass multi-county geographies covering multiple communities that vary in size and economic activity. Continuous geographic coverage and accurate data for all levels of detail make small areas an essential component. The issue of sampling errors and their impact on the estimates were recognized with the transitioning from the traditional method to a rolling sample approach. The Compass series published by the Census Bureau includes a set of instructions to guide users/researchers how to process errors in deriving new estimates using ACS data and its products.

In addition, custom applications have been developed by incorporating more reliable information from neighboring features also referred to as the Touch Method. These approaches produce measures that are aggregated to a different level of detail typically by merging nearby features. We outlined the basics of two experimental procedures that can maintain the desired level of geographical detail while limiting the variability of the estimates.

It should be noted that although analytical methods support the interpretation of data for small areas or areas with low response rates, local knowledge and professional judgement should be used in conjunction with analytical tools.

## 4.4 CTPP Data Access Guide

One of the objectives of this project was to supplement existing training materials and to provide less experienced users with an interface that can guide them to more easily access CTPP data. Inspired by the “Guidance for Data Users” for the ACS, we developed recommendations for a range of user profiles with different data needs using a troubleshooting guide format. These recommendations can be incorporated into the existing CTPP website as a stand-alone page or can be converted into a wizard format.

Additional instructions were given for downloading data using bulk selection and creating custom groups. Links for accessing historical CTPP data and raw data for the 2006-2010 CTPP were provided.

## 4.5 Comparison of 2006-2010 CTPP and LEHD Worker Flows

The LEHD database is another source widely used by transportation planners to evaluate employment levels and worker flows. The peer exchange panelists suggested that providing both CTPP and LEHD data in the same platform could be beneficial to users. This also brought up the differences in these data sources.

A set of small scale comparisons between 2006-2010 CTPP data and five LEHD datasets were developed to demonstrate the variability of the LEHD data and how CTPP data compare to the LEHD datasets. The following data items were chosen for these comparisons:

- Number of workers at their residence at the county level by industry type,
- Number of workers at their workplace at the county level by industry type,
- Magnitude of the intra-county flows by industry,
- Top five county-level flows,
- Relative size of the worker flows for the top 5, 10, and 25 flows,
- Number of matched census tract pairs in the top 5, 10, and 25 LEHD flows, and
- Frequency distribution of tract level flows by flow size categories.

The following counties were selected for analysis in this section to enable the project team to further analyze these issues:

- DuPage County, IL
- Dane County, WI
- Duval County, FL
- Saint Louis County, MN and
- Douglas County, NE.

The key findings are listed below, and a detailed discussion can be found in the Task 3 Report.

- Estimates of the number of workers at their residence from the CTPP data were higher than the estimates provided by the LEHD flows.

- Percent distributions of workers across industries were reasonable.
- Both the estimates for resident workers and the estimates of the total number of workers at the workplace significantly decrease starting in 2009 in LEHD data.
- CTPP data consistently provided higher estimates of intra-county worker flows. The discrepancies between the two data sources also highlight the effects of the recession that was reflected more accurately in the LEHD data.
- The top five county-level OD pairs were quite consistent between the CTPP and LEHD datasets.
- Analysis of the data at the tract level indicated that the LEHD flow patterns are generally more dispersed than the CTPP patterns.
- The level of agreement in the top ranked flows in the CTPP and LEHD datasets was analyzed by identifying the number of matched pairs in the 25 flows.
  - The number of matched Census tract pairs between the CTPP and LEHD varied across different counties and seemed to be inversely correlated with the size of the region.
- Analysis of worker flows at the tract level showed that LEHD data had very dispersed spatial patterns of origins. The frequency distribution of tract level flows in the LEHD data showed a higher frequency of small flows when compared to the CTPP data.

This simple exercise confirmed that there are significant differences in worker flows between CTPP and LEHD data. Users need to review the methodologies and data patterns from these sources carefully. LEHD data portray a better picture of trends, however, it provides more dispersed patterns of worker flows compared to CTPP. The disclosure proofing methodology applied by the LEHD program, the smaller sample size of the ACS, and the resultant weights are probably contributing to the observed discrepancies in the spatial patterns of origins. Further research can evaluate these claims as well as investigate the feasibility of data fusion from these sources.





## 5.0 Summary and Looking into the Future

The activities and analyses summarized above and detailed in the Task reports provide an overall picture of the usefulness and value of the ACS-based CTPP data and dissemination tools. The list below summarizes many of the key findings of this effort:

- The size of the data user community is modest.
- The most common uses of the data are for travel demand modeling and the analysis of the commute market. Household travel survey sampling and expansion, corridor studies, environmental justice analysis, and evaluation of small travel markets such as bike to work flows and travel by self employed include some of the unique uses of CTPP.
- Stakeholders view CTPP as providing significant value over the ACS data.
- Most users are knowledgeable about the limitations and challenges of working with data derived from smaller sample sizes and accumulated over multiple years. Most users believe that the five-year data provide benefits compared to the decennial long form Census data.
- Most users prefer to use small geography data, particularly at traffic analysis zone level while recognizing the uncertainties and data quality at this level of geographic detail. Census tract and county level data are valued highly as well.
- Users acknowledge the sampling issues in small areas and resultant large MOEs. However, they rarely incorporate MOEs in their analysis. Some more experienced users rely on MOEs as a quality measure that is most often used qualitatively when interpreting the data.
- Perturbed data are welcomed by the community as a means of avoiding “missing” estimates and there is an overall trust towards the methods that are implemented. However, more seasoned users still would like to see comparisons before and after perturbation.
- The online software tool is considered to be capable and comprehensive by most users, but there is general consensus that the learning curve is steep. Some users still prefer to access data via more traditional ways, particularly when analyzing data from a large metro area. The role and functionality of the software may need to be reevaluated considering the substantial reductions in the CTPP content and the provision of most Part I and Part II tabulations by the Census.
- The great diversity of training opportunities was very appealing to the community. The respondents who had attended the in-person classes generally found them to be the most effective means of training, and some respondents pointed out the “marketing” value of these classes. Some Oversight Board interview respondents suggested that training be offered on what the CTPP data could be used for and how, instead of only providing instructions on how to access and download a particular tabulation.
- Both Peer Exchange participants and Oversight Board interview respondents cited the increasing importance of integrating CTPP with other common transportation planning datasets.

- The transportation data community also identified a set of potential value-added improvements particularly on data dissemination. The objective is to help users access multiple data sets and facilitate data fusion with other sources such as LEHD/LODES, and NHTS data.
- In addition, supplementing the data with travel distance information during the data compilation stage was suggested. This may require additional effort and coordination with national and local agencies, however, online travel time estimates can be used as a proof of concept.
- Potential long term ACS improvements worthy of exploration included second jobs, better information on telephone availability, sub modes such as access/egress modes, new modes such as shared ride arrangements, autonomous vehicles, and international commuting trips for border communities.
- CTPP data users identified a few potential ACS-based cross-tabulations and travel flow tabulations that would be helpful for specific analyses they recently performed or would like to perform. Driven by advanced travel demand modeling needs, these users aim to understand travel behavior at person level via additional tabulations such as age, gender, employment, occupation, and earnings; vehicle parity with workers, adults and/or household size; school enrollment and part-time arrangements to work; and internet access and use.
- The following research ideas are proposed for future exploration:
  - Compare CTPP data and alternative sources of data for model validation.
  - Develop guidance and/or a summary of “best practice” approaches on how to combine CTPP and other data sources, possibly including:
    - » LODES,
    - » National Household Travel Survey (NHTS) building upon ongoing research,
    - » Regional and statewide travel surveys,
    - » Private source employment data,
    - » Private source cellphone data, and
    - » Local agency employment and school data; and
  - Evaluate effects of household-based and person-based perturbation on the resulting estimates.