CTPP Status Report

August 2015

U.S. Department of Transportation
Federal Highway Administration (FHWA)
Bureau of Transportation Statistics (BTS)
Federal Transit Administration (FTA)

AASHTO Standing Committee on Planning
TRB Census Subcommittee

Census Transportation Planning Product (CTPP) Update

Penelope Weinberger, AASHTO, Pweinberger@aashto.org

The CTPP Oversight Board will be meeting in August in Denver. Items on the agenda include: new directions for training; outreach and research; improving our web presence; planning a census data conference; finalizing the 2012 to 2016 special tabulation request to the Census Bureau; and of course, the budget.

Our goal is to submit a CTPP table list to the Census Bureau no later than January 2016. To help determine which tables to retain or eliminate in the 2012 to 2016 CTPP dataset, Beyond 20/20 Inc. recently developed a tool to evaluate the quality of the 2006 to 2010 CTPP data. This tool looped through the entire 2006 to 2010 data set, cell by cell; looking at the ratio of the margin of error (MoE) to the estimate. The tool reports a red flag for any cell with a MoE equaling more than 35 percent of the estimate, and again if it is more than 50 percent. For some cells, a low estimate and a high MoE make perfect sense, such as for cells in table A101202 Age (9) by School Enrollment (7). Since there are few people in older age cohorts enrolled in elementary school, high school, and college, a small estimate and a large MoE are not surprising. The overall score for cells in table A101202, therefore, ranges from 4.5 percent to 76 percent reliable. Cells at the TAZ level (217,526 TAZs) are the least reliable with only 4.5 percent of the cells contain an estimate where the MoE is less than 35 percent of the estimate. At larger geography, such as the national total and the geographic components, the cells are 76 percent reliable. We are currently analyzing the data for less intuitive anomalies.

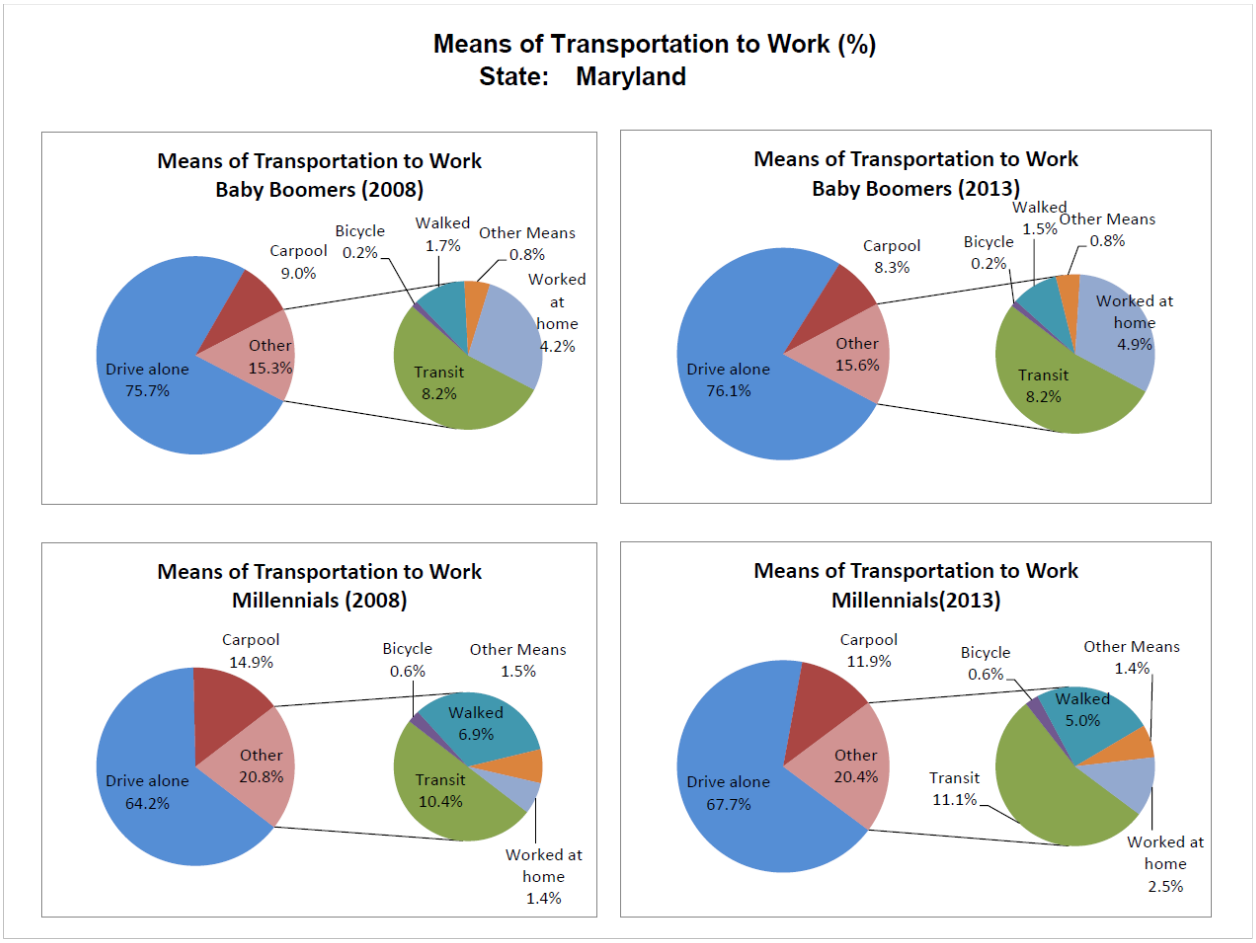
Please let me know if you are interested in the results, the outcomes of our annual meeting, or any other thing. As always, I am open to your comments and suggestions.

CTPP "Generations" Profile Using 2006-2008 and 2011-2013 ACS Public Use Microdata Sample (PUMS)

Elaine Murakami, FHWA, Elaine.Murakami@dot.gov
Jingjing Zang, Cambridge Systematics, jzang@camsys.com

Do Millennials travel differently than Baby Boomers? One way to answer this is to compare the commute (Journey-to-work) results from the American Community Survey (ACS). CTPP "Generations" profiles using data from the 2006 to 2008 and the 2011 to 2013 American Community Survey (ACS) Public Use Microdata Sample (PUMS) are now available at: http://www.fhwa.dot.gov/planning/census_issues/american_community_survey/products/2015_transportation_profiles.

In our example of the State of Maryland, shown in Figure 1, we can see that Baby Boomers have a much higher vehicle availability than Millennials. Baby Boomers are more likely to work at home, and Millennials are more like to walk, bike, and take transit to get to work.



Vehicles Avail Per Person	**Baby Boomers**		**Millennials**		**Total**	
Universe=Persons	2008	2013	2008	2013	2008	2013
0	88,196	92,417	49,723	80,602	137,919	331,796
0.10 to 0.49	62,301	70,160	66,844	105,030	129,145	275,053
0.50 to 0.99	300,445	312,073	226,259	360,036	526,704	1,151,982
1.00 and over	1,051,246	995,094	306,074	509,753	1,357,320	2,796,638
Total	1,502,188	1,469,744	648,900	1,055,421	2,151,088	4,555,469

Figure 1 Generation Profile Sample Tables and Charts—State of Maryland

Each profile includes information about Vehicles Available Per Person, Means of Transportation to Work and Travel Time to Work, as well as Race, Hispanic Origin, Personal Income, Housing Tenure, and Housing Structure Type. Two sets of pie charts are included in the Excel version of each profile.

While the Baby Boom generation has defined years of birth (1946 to 1964), the years of birth to define the Millennial generation are still fluid. For this tabulation of the ACS PUMS, we have used the years of birth between 1983 and 2000. Because some of our tabulations are of workers, in those tables, we restricted the age to people ages 16 and over. In 2008, many Millennials were still under age 16. This is the main reason that there are large differences between the total number of Millennials between 2008 and 2013 when looking at workers, and also for tables where the age of the “reference person” is used to classify households.

Usually the householder for each household is the reference person, or one of the people, in whose name the home is owned, being bought, or rented, and who is listed as “Person 1” on the survey questionnaire. If there is no such person in the household, any adult household member 15 and older can be designated. Table 1 provides an example of Baby Boomer Household/ Millennial Household.

Baby Boomer Household			Millennial Household
Three-person household, with “reference person” age 56 in 2013			One-person household, with “reference person” age 23 in 2013
Born in 1957 (*Reference Person*)	Born in 1960	Born in 1990	Born in 1990 (*Reference Person*)

Table 1 Baby Boomer Household and Millennial Household

The profiles are available for 50 States and District of Columbia, and for 337 counties. County profiles are limited to counties which share no Public Use Microdata Areas (PUMA) with other counties (in both 2000 PUMA and 2012 PUMA geography).

We have not included margins of errors (MOE) in these profile sheets because we have limited the profiles to large geographies where MOEs are of less concern. Because the sample size of the American Community Survey is much smaller than the decennial census “long form,” it is more important to understand the potential errors in the tabulated results even they are not expressed on the profile sheets.

Please let us know if these profile sheets are useful to you. Please send your comments to CTPPSupport@camsys.com.

Ridesharing: the Easiest (and Hardest) Approach to Congestion Reduction: Mashing up CTPP with the American Community Survey to Visualize Ridesharing’s Potential across U.S. Neighborhoods

Peter Viechnicki, Deloitte Services LP,
pviechnicki@deloitte.com
Brad Gudzinas, Deloitte Advisory,
bgudzinas@deloitte.com
Mahesh Kelkar, Deloitte Services LP,
mkelkar@deloitte.com

As the cost of congestion continues to climb,[1] many cities and regional planning organizations are taking a fresh look at ridesharing (carpooling) as a way to help increase capacity on existing highways without starting expensive new capital projects. But there is no current consensus about which policy levers can finally reverse the decades-long decline in carpooling rates, or where such levers should be applied. To help inform this discussion, Deloitte recently released its Smart Mobility report—the first nationwide study estimating carpooling’s potential from CTPP and American Community Survey (ACS) data. It finds a surprisingly large untapped supply of potential ridesharers, with high concentrations in suburbs 10-15 miles outside the centers of large cities. It also shows how mashing up CTPP and ACS gives the policy-makers an insight into the true potential of carpooling in each neighborhood.

Deloitte’s new method is based on the 2006-2010 CTPP tract-tract work flow data linked to ACS data on commuting patterns by census tract id. The ACS tables used were B08301 and B08302, mode and departure time for journey to work. TIGER line files were used to obtain spatial geometries for each census tract.

Deloitte researchers developed a new algorithm combining data from these sources using open-source database tools, custom PERL code, and a GIS package.[2] The new algorithm then calculates ridesharing potentials for each census tract. Figure 2 shows details of how the maximum potential number of new ridesharers

1 The average American now wastes 34 hours a year sitting in traffic, according to the Texas Transportation Institute. The annual costs associated with this congestion are estimated at $121 billion, slightly more than 1 percent of all U.S. personal consumption spending.

2 Our methods work with either open-source GIS such as QGIS or COTS package such as ArcMAP 10.

and Vehicle Miles Traveled (VMT) saved was estimated for each census tract.

Several simplifying assumptions are made in the calculations based on knowledge of nationwide carpooling patterns:

- Commuters who live in the same tract and commute to the same tract for work could potentially carpool together;
- Commuters are not willing to alter their departure time by more than 30 minutes to accommodate a carpool;
- Commuters are also not willing to drive more than a mile out of their way to pick up a carpool partner;
- Commuters' homes are evenly distributed throughout the small, densely populated tracts in our study sample; and
- Sixteen percent of potential ridesharers would not choose ridesharing as they would need to make multiple stops on the way to or from work ("trip chaining").[3]

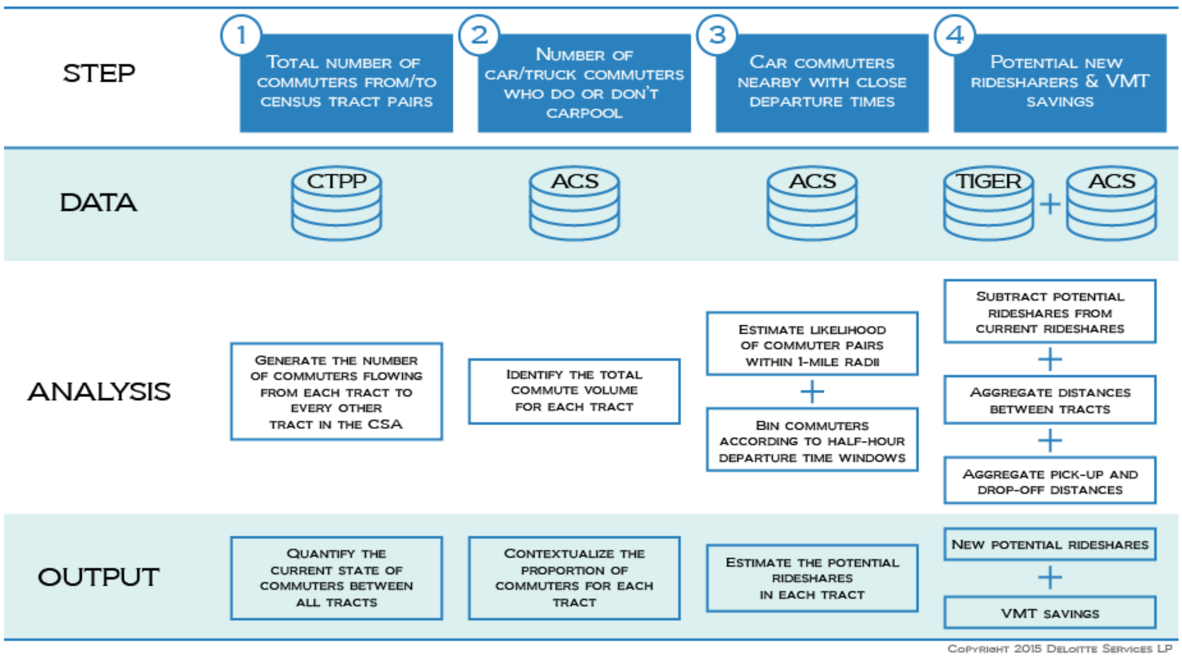


Figure 2 How We Mash up CTPP and ACS to Estimate Ridesharing Potential

Because this new method also estimates reductions in vehicle miles traveled for new ridesharers, it allows forecasting potential economic savings, safety improvements, congestion reductions, and lowered carbon dioxide emissions for a geography.[4]

Deloitte's new algorithm yields estimates of the number of potential ridesharers/carpoolers for all census tracts that are contained within "combined statistical area" metro areas in the U.S. The results can be aggregated to the metro area level for planning purposes. Figure 3 shows a sample map of potential ridesharing growth neighborhoods in Indianapolis, Indiana and Jacksonville, Florida. The "ring of ridesharing potential" shows up in both Indianapolis and Jacksonville, 10 to 15 miles outside the city center.

---

[3] The most recent National Household Travel Survey (NHTS) found that 16 percent of home-to-work tours involved multiple stops, while 84 percent of home-to-work tours had no stops.

[4] For details of how we estimate economic benefits, safety improvements, congestion reduction, and carbon emissions savings, please see Peter Viechnicki, Tiffany Fishman, Abhijit Khuperkar, and William Eggers, "Smart Mobility: Reducing Congestion and Fostering Faster, Greener, Cheaper Transportation Options," Deloitte University Press, 19 May 2015, DUPress.com.

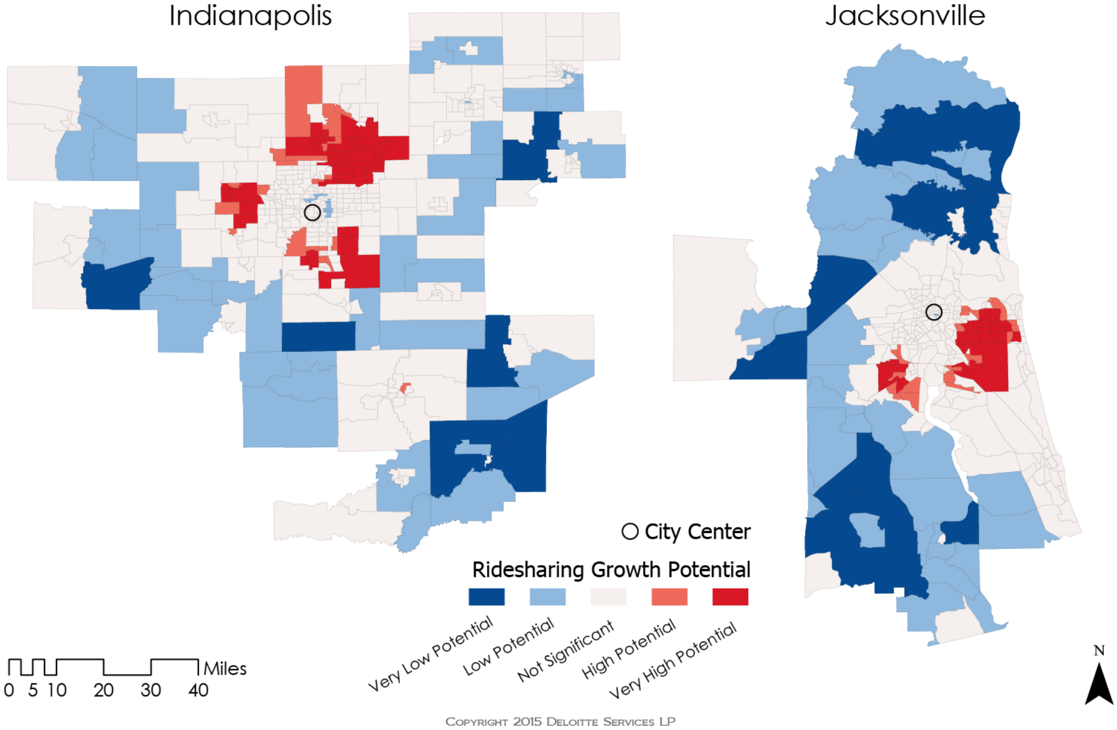


Figure 3 Exploring Ridesharing Potential in Jacksonville and Indianapolis

To see this phenomenon in action, access Deloitte's interactive map. Click on the city of Indianapolis and zoom in. Look for the second map in the ridesharing theme—"Neighborhoods Projected New Ridesharers." The highest number of potential ridesharers are concentrated in the range of 10-15 miles from the city center. These are the suburbs of Carmel, Fishers, Greenwood, Zionsville, and Brownsburg (marked in a darker shade). This is where the potential of ridesharing may lie and possibly one that the transportation planners can tap into.

To facilitate additional discussion about ridesharing potential, the results of the Smart Mobility analyses are freely available in a variety of formats. City-level savings tables and mobility snapshots, and an interactive map allow users to zoom in and out on a neighborhood. An ArcGIS REST API endpoint for power users to pull down raw data is also available, inviting a deeper dive into the interactive data to investigate ridesharing, carsharing, and bike commuting's potential for neighborhoods and cities.[5]

[5] See the ArcGIS REST API resources page for details of how to request data via the API: http://arcg.is/1GRnwRH.

National Synthetic Population Generation Using Census Data

Ram M. Pendyala, Georgia Institute of Technology, Atlanta,
rpendyala6@mail.gatech.edu
Daehyun You, Georgia Institute of Technology, Atlanta
Venu Garikapati, Georgia Institute of Technology, Atlanta

The generation of a representative synthetic population of households and persons is critical to the deployment of microsimulation models of activity-travel demand that aim to simulate the activity-travel patterns of individual agents in a metropolitan area. Recently, a large national-scale synthetic population generation effort was undertaken as part of a Federal Highway Administration (FHWA) Exploratory Advanced Research project on the development of a Long Distance Passenger Travel Demand Modeling Framework.

The goal of the project was to develop a comprehensive microsimulation modeling framework capable of forecasting long-distance passenger travel demand by all modes of transportation for the entire nation. This required generating a nationally representative synthetic population of more than 300 million agents (individuals). The modeling framework includes model components capable of predicting all aspects of long-distance passenger trips, including generation, destination choice, intermediate stops, duration and length of stay, mode use, and accompaniment. This article offers a synopsis of the national synthetic population generation effort that utilized American Community Survey (ACS) data. For this project, the national synthetic population generation effort was limited to the 50 States plus the District of Columbia. As of the 2010 census count, this represented a population of 308.7 million people. Of this population, 300.8 million people resided in 116.7 million households, while the remaining 8 million people lived in group quarters. The nation had 3,143 counties, 73,057 census tracts, and 217,740 block groups in the 50 States plus District of Columbia.

A synthetic population may be generated at a variety of geographic resolutions, including, for example, county, census tract, traffic analysis zone, or census block group. A delicate balance must be struck between representativeness of the synthetic population and computational complexity and burden. In choosing the appropriate geographical unit or spatial resolution for the synthetic population generation effort, it was felt that the "county" is less than ideal due to its very aggregate and coarse nature. Ideally, it would be desirable to perform synthetic population generation at the level of the block group; however, using the block group as the spatial unit for a national synthetic population generation exercise is computationally prohibitive. As a compromise between these two extremes, the census tract was chosen as the basis for the national synthetic population generation effort. The tract-level synthesis involves generating a population for just over 73,000 census tracts in the country, and even the deployment of a modest parallel computing architecture offers reasonable computational time.

In this study, the procedures embedded in the PopGen (Ye et al, 2009)[6] software package were used to generate a nationwide synthetic population. The PopGen system is a robust synthetic population generation software capable of controlling for both household- and person-level attributes of interest. The key input data sets are as follows:

- A sample file that includes disaggregate household and person records for a sample of the population. This sample file serves two key purposes:
  1. It provides the multidimensional joint distribution (seed matrix) among attributes of interest; and
  2. Households included in the synthetic population are drawn from the sample file.
- A marginal control file that includes aggregate household- and person-level control totals and distributions at the desired level of geographic resolution (census tract in this study). This file provides the control totals that must be matched in the synthetic population generation process.
- A geographic correspondence file that maps individual geographies (such as census tracts) to larger geographic areas, namely, the Public Use Microdata Area (PUMA). The joint distribution (seed matrix) of attributes of interest for a specific PUMA is applied to all census tracts that belong (map) to that particular PUMA.

The 2007-2011 five-year ACS data was used for the national synthetic population generation effort. Thus, the marginal control data for a variety of household and person attributes is derived from the ACS 2007-2011 data compilation, and the sample file corresponds to the ACS 2007-2011 PUMS data—a five percent sample of the U.S. population. Also, the latest 2010 decennial census version of the Mable Geocorr geographic correspondence files,

[6] Ye, X., K. Konduri, R.M. Pendyala, B. Sana, and P. Waddell (2009) A Methodology to Match Distributions of Both Household and Person Attributes in the Generation of Synthetic Populations. *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, National Research Council, Washington, D.C.

developed by the Missouri Census Data Center, are used to map census tracts to PUMAs.

PopGen follows a three-step process. First, the joint distribution of the attributes of interest is determined for each geography (census tract). The marginal control totals from the census files are used to expand this joint distribution matrix so that the marginal control totals are matched exactly. This procedure, known as iterative proportional fitting (IPF), is applied to both the household- and person-level attribute joint distributions. As a result of the first step, the total number of households or persons that need to be generated in each cell of the joint distribution matrix is determined.

In the second step, every household in the sample is given a weight such that the weighted total of households (persons) matches the total number of households (persons) in each cell of the joint distribution as calculated through the IPF procedure. This step is referred to as the Iterative Proportional Updating (IPU) algorithm wherein the weights associated with households are iteratively updated such that the weighted frequencies of households and persons match the expanded joint distribution totals at both the household and person levels.

In the third step, households are drawn through a Monte Carlo simulation procedure using the weights computed in the second step. This completes the synthetic population generation process.

PopGen is able to use any combination of control variables for synthesizing a population for the nation. While the use of many control variables may sound appealing, the use of a large number of control variables can increase computational burden and lead to sparse multidimensional joint distribution matrices. At the household level, control variables included presence or absence of children, household size, age of householder, household income, number of workers in household, and type of household. At the person level, control variables included age, gender, employment status, and race. The synthetic population also includes group-quarter residents, distinguishing between individuals in institutional and noninstitutional settings.

The synthetic population generation process was executed at the level of the census tract using ACS 2007-2011 data for the entire nation. The synthetic population files for each State were assessed to ensure that the population synthesized for each census tract closely mirrored that in the marginal control data sets from the census. Comparisons were performed on total households/persons generated for each State, and the results show a strong agreement between the synthetic population and true population in each census tract for all States.

In addition to ensuring that the count of number of households and persons is correct, the attributes of those households and persons must also be compared. These comparisons can be performed at various geographic levels, including State-, county-, and census tract-level. The set of graphs in Figure 4 shows a comparison of household and person attributes for one randomly chosen census tract in Maricopa County (Greater Phoenix metropolitan region) in Arizona. The comparisons demonstrate the very close match between actual population characteristics and synthetic population characteristics, and this pattern of consistency was found to hold nationwide.

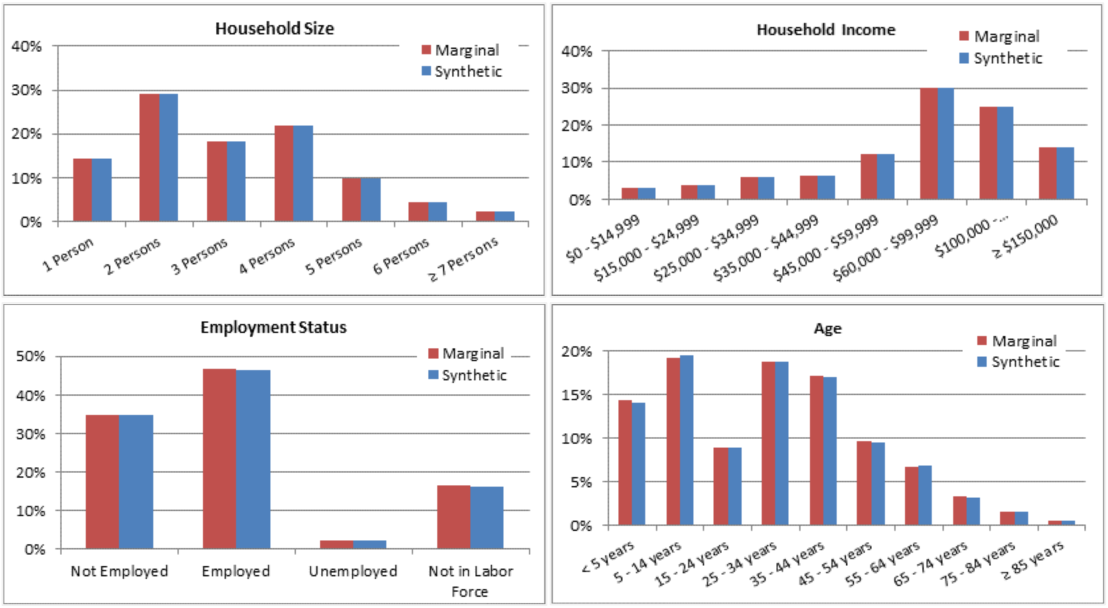


Figure 4 Comparison of Control Distributions Between Actual and Synthetic Populations (Census Tract 522745 in Maricopa County, Arizona)

In summary, the national synthetic population generation effort proved to be successful, yielding a representative nationwide synthetic population suitable for long-distance travel demand modeling and forecasting. The 2007-2011 ACS data sets offered rich and consistent aggregate and disaggregate information about the U.S. population that proved critical to the development of a national synthetic population. With ACS data sets updated on an annual basis, it is possible to synthesize a population for any region of interest for any model base year, and this has proven critical to the deployment of advanced activity-based microsimulation model systems around the country.

Acknowledgments

The FHWA Exploratory Advanced Research Project "Foundational Knowledge to Support a Long-Distance Passenger Travel Demand Modeling Framework" was conducted under contract DTFH61-10-R-00036 by a team led by Resource Systems Group, Inc. Maren Outwater served as the Principal Investigator. PopGen was developed while the authors were at Arizona State University. Karthik Konduri, University of Connecticut, is the primary author of the PopGen software package.

CTPP Program Update

Elaine Murakami is retiring on September 30, 2015, after 22 years at Federal Highway Administration, and 9 previous years at the Puget Sound Regional Council.

Elaine has been involved with the CTPP developed from the 1990 and 2000 decennial censuses, and the 2006-2010 CTPP using the American Community Survey. One thing Elaine counts among her many accomplishments was using a GIS-approach (and supplying the software) for TAZ delineation of CTPP 2000. In addition to the TAZ program Elaine also encouraged transportation agencies to use color maps to promote the use of GIS, and she produced the "Census Mapbook for Transportation Planning" which contains mapping applications of CTPP data and is still relevant today. The Mapbook was completed in 1994.

Elaine is very proud of the recognition she received from TRB with emeritus status for her work with many committees, including: Travel Survey Methods (ABJ40), Urban Data (ABJ30), Transportation Planning in Small- and Medium-Sized Communities (ADA30), and the Women's Issues in Transportation (ABE70). Research activities at FHWA have included the first GPS and "hand-held" computer for travel surveys conducted in 1995, an SBIR-sponsored project on using web-based GIS for household travel surveys, and most recently, testing a Smartphone app for a multiday, multimodal travel survey. In her work at the Puget Sound Regional Council, Elaine began the Puget Sound Transportation Panel, a longitudinal survey with choice-based samples (transit and carpool). Her work on travel surveys has led to data archiving projects, with GPS-based travel surveys now being housed with virtual access at the Transportation Secure Data Center.

Elaine has also been very active with women's issues in transportation as well as mentoring women working in transportation, both formally through Women's Transportation Seminar (WTS) and informally. Some of her work on the history of women in transportation is still available. http://www.fhwa.dot.gov/wit/.

As Elaine retires she plans to spend more time quilting, folk dancing, traveling, and teaching teenagers to sew (Figure 5). Her email address is ermurakami@gmail.com

Figure 5 Elaine Murakami and Her Quilt

CTPP Contact List

Email: CTPPSupport@camsys.com
CTPP 2006-2010 Data: http://ctpp.transportation.org/Pages/5-Year-Data.aspx
CTPP website: http://www.fhwa.dot.gov/planning/census_issues/ctpp/
FHWA website for Census issues: http://www.fhwa.dot.gov/planning/census_issues
AASHTO website for CTPP: http://ctpp.transportation.org
1990 and 2000 CTPP data downloadable via Transtats: http://transtats.bts.gov/
TRB Subcommittee on census data: http://www.trbcensus.com

AASHTO
Penelope Weinberger
Phone: 202/624-3556
Email: pweinberger@aashto.org

Tracy Larkin Thomason, NVDOT
Chair, CTPP Oversight Board
Phone: 702/385-6500
Email: Tlarkin@dot.state.nv.us

Guy Rousseau, Atlanta Regional Commission
Vice Chair, CTPP Oversight Board
Phone: 404/463-3274
Email: GRousseau@atlantaregional.com

Census Bureau: Social, Economic and Housing Statistics Division
Brian McKenzie
Phone: 301/763-6532
Email: brian.mckenzie@census.gov

FTA
Ken Cervenka
Phone: 202/493-0512
Email: ken.cervenka@dot.gov

BTS
Julie Parker
Phone: 202/366-6373
Email: Julie.parker@dot.gov

FHWA
Elaine Murakami
Phone: 206/220-4460
Email: elaine.murakami@dot.gov

Ed Christopher
Phone: 708/283-3534
Email: ed.christopher@dot.gov

TRB Committees
Stacey Bricka
ETC Institute (as of October 4, 2015)
Chair, TRB Urban Data Committee
Email: sbricka@etcinstitute.com

Clara Reschovsky
TRB Census Subcommittee Co-Chair
Phone: 202/962-3332
Email: creschovsky@mwcog.org

Mara Kaminowitz
TRB Census Subcommittee Co-Chair
Phone: 410/732-0500
Email: mkaminowitz@baltometro.org

CTPP Technical Support
Jingjing Zang
CTPPSupport@camsys.com

CTPP Listserv

The CTPP Listserv serves as a web-forum for posting questions, and sharing information on Census and ACS. Currently, more than 700 users are subscribed to the listserv. To subscribe, please register by completing a form posted at: http://www.chrispy.net/mailman/listinfo/ctpp-news.

On the form, you can indicate if you want emails to be batched in a daily digest. The website also includes an archive of past emails posted to the listserv.