MEMORANDUM

To: NYMTC members

From: New York City Department of City Planning

Date: June 3, 2019

Re: Reliability considerations for CTPP 2012-16 in NY Metro Region

**What is CTPP?**

Census Transportation Planning Product (CTPP) is a special tabulation of U.S. Census Bureau American Community Survey (ACS) data designed for transportation planning, describing socioeconomic characteristics of the population by residence, workplace, and home-to-work (i.e. origin-destination) trips. The latest release is derived from the 2012-2016 ACS 5-Year Estimates, and the prior release was derived from 2006-2010 ACS 5-Year Estimates, each release representing a 5-year annual average across the indicated period.

CTPP datafiles are typically organized in three parts—**Part 1: Residence**; **Part 2: Workplace**; **and Part 3: Flows** (origin-destination pairs). Data are reported at geographic summary levels as granular as Census Tracts and Traffic Analysis Zones (TAZs), and as large as counties and states.

**What is the issue with the current CTPP release (2012-2016)?**

NYCDCP observed a significant undercount of workers and residents in Part 2 and Part 3 of the new CTPP release for data reported by geographic summary levels smaller than counties (e.g. tracts, TAZs).

For example, the 2012-2016 CTPP reported the total number of people who work in New York State (i.e. Part 2 total) is 9,515,770, when summing by county. However, the sum of reported workers by tract for New York State (NYS) is 7,257,862 (24% lower than the Part 2 county total). The reported total number of employed NYS residents by residence location (i.e. Part 1 total) is 9,148,028, which was equal to the county-level aggregation. However, the Part 3 reported total number of employed NYS residents aggregated by tract is 6,927,621, suggesting a known place of work for only 76% of employed NYS residents – representing a discrepancy of 2.2 million employed NYS residents when examining residence-workplace flows for geographies smaller than counties.

**What is the cause of the discrepancy between Part 1 and Parts 2 and 3?**

According to the U.S. Census Bureau and AASHTO, the entities that prepare and oversee the CTPP datasets, the primary reason for the undercount in Parts 2 and 3 is due to the large percentage of respondents (estimated 25% on average across the country) who do not provide any or enough information about their workplace location, which makes geocoding their places of work at smaller geographies (i.e. below county-level) difficult or not possible. Therefore, the reported totals for Parts 2 and 3 are for those respondents with a known place of work.

For the 2006-2010 CTPP (and earlier products), the Census Bureau adopted an extended allocation procedure, an imputation process intended to allocate workers with unknown workplace locations to a Census Block-level place of work based on other known/reported characteristics prior to publication of the datasets. Using that imputation process, the aggregate totals for smaller geographic summary levels (e.g. tracts) in the Parts 2 and 3 data represented nearly 100 percent coverage in New York City and surrounding metro region.

However, that extended allocation procedure was not applied to the 2012-2016 CTPP following methodological revisions and questions surrounding the accuracy of prior results. Though the current release may provide more accurate information, missing cases at small summary levels remain missing in the new CTPP, resulting in the significant difference between county-level and tract-level figures uncovered in Parts 2 and 3.

**How does this affect the way we use CTPP data?**

Part 1 data for all geographic summary levels and Parts 2 and 3 data reported for county or larger geographies are not affected by the change in methodology. However, CTPP users should take extra caution when using tract-or TAZ-level Part 2 and Part 3 totals and distributions, such that:

* **Part 2 and Part 3 published totals for sub-county geographic summary levels are not complete**. Reported totals and distributions for these geographies solely represent known cases. Reported totals will not add to larger geographies (e.g. tracts aggregated to counties), nor will it equal totals reported in Part 1.
* **Cases with missing place-of-work data are not evenly distributed across geographies and may bias towards certain population groups**. DCP’s preliminary analysis shows that the NYC neighborhoods (parks and cemeteries excluded) are impacted by the missing place-of-work data to various degrees. Residents with a reported place-of-work by NYC neighborhood[[1]](#footnote-1) range from 37% to as high as 81%. From DCP’s observation, neighborhoods with lower Census survey response rates tend to be more impacted by the undercounting issue.
* **Missing cases may carry a different distribution of characteristics from reported cases**. Based on DCP’s preliminary analysis, reported cases in the published data available via AASHTO may not be representative of all cases (i.e. the knowns may not accurately represent the unknowns). For example, the reported Part 2 and Part 3 “Means of Transportation to Work” files may show that 60% of residents from Census Tract A travel by car to work in Census Tract B. However, only 20% of all residents from Census Tract A are reported for Census Tract B, and the missing 80% of reported Census Tract A residents might not show the same tendency to commute by car.
* **2006**-**2010 CTPP cannot be compared to 2012-2016 CTPP for geographies smaller than counties for Part 2 and 3 data.** Any attempt to compare the new place-of-work counts with the previously reported CTPP totals at such geographic summary levels will likely lead to a questionable result, given the significant change in methodology between the two releases.

**What should we do now?**

The methodological change resulting in the 2012-2016 CTPP will not be modified to have an imputed “complete” count for Parts 2 and 3 data at smaller geographies that will sum to county/larger geographies. Preliminary conversations with the U.S. Census Bureau and AASHTO have explored the release of a comparable data product limited to reported knowns for the 2006-2010 CTPP, which would enable longitudinal comparisons. However, such a product would take time to prepare and there is no immediate solution for comparing Part 2 and 3 smaller geographies across the years, or analyze differences in the representative distributions. We are actively working with our colleagues at the Census Bureau to better understand data reported for smaller geographies in both the current and prior releases.

Given that many planning professionals and transportation agencies use CTPP for various purposes, in the meantime we request your attention and caution when using the new CTPP release. We advise that you please use discretion in interpreting the data and analytic results based on local knowledge and information from other federal data sources, regional and local travel surveys, transit agency reports, etc. Please advise data users in your agency and consultants working on your projects to do the same.

1. Neighborhood Tabulation Area [↑](#footnote-ref-1)