CAMBRIDGE SYSTEMATICS

New York Best Practice Model 2019 Update

Model Validation Plan

final

report

prepared for

New York Metropolitan Transportation Council

prepared by

Cambridge Systematics, Inc.

with

EA Harper Consulting

final report

New York Best Practice Model 2019 Update

Model Validation Plan

prepared for

New York Metropolitan Transportation Council

prepared by

Cambridge Systematics, Inc.
101 Station Landing, Suite 410
Medford, MA 02155

with

EA Harper Consulting

date

July 13, 2021

Table of Contents

List of Tables

List of Figures

1.0 Overview

This document presents the model validation plan for the update to the New York Best Practice Model (NYBPM), an activity-based model that is being updated for the New York Metropolitan Transportation Council (NYMTC). The model update is being led by a consultant team consisting of Cambridge Systematics, Inc. (CS), EA Harper Consulting, Gallop Corporation, and Florida International University. This report presents the plan for validating the model that is being updated to a 2019 base year.

The development of this validation plan and the validation procedures and recommendations presented herein are based on guidance provided in the 2011 Travel Model Improvement Program Travel Model Validation and Reasonableness Checking Manual – Second Edition (Cambridge Systematics, Inc., 2010), hereafter referred to as the “Validation Manual.” This validation plan also builds on experience gained with other activity based models – including the 2012 NYBPM developed by the CS team.

The report is organized as follows. After the general validation process is summarized in this chapter, the data to be used for validation is described in Chapter 2. Chapter 3 describes the specific validation tests to be performed for the demand model components and for the highway and transit assignment. Chapter 4 briefly discusses the sensitivity testing process.

1.1 Validation Process

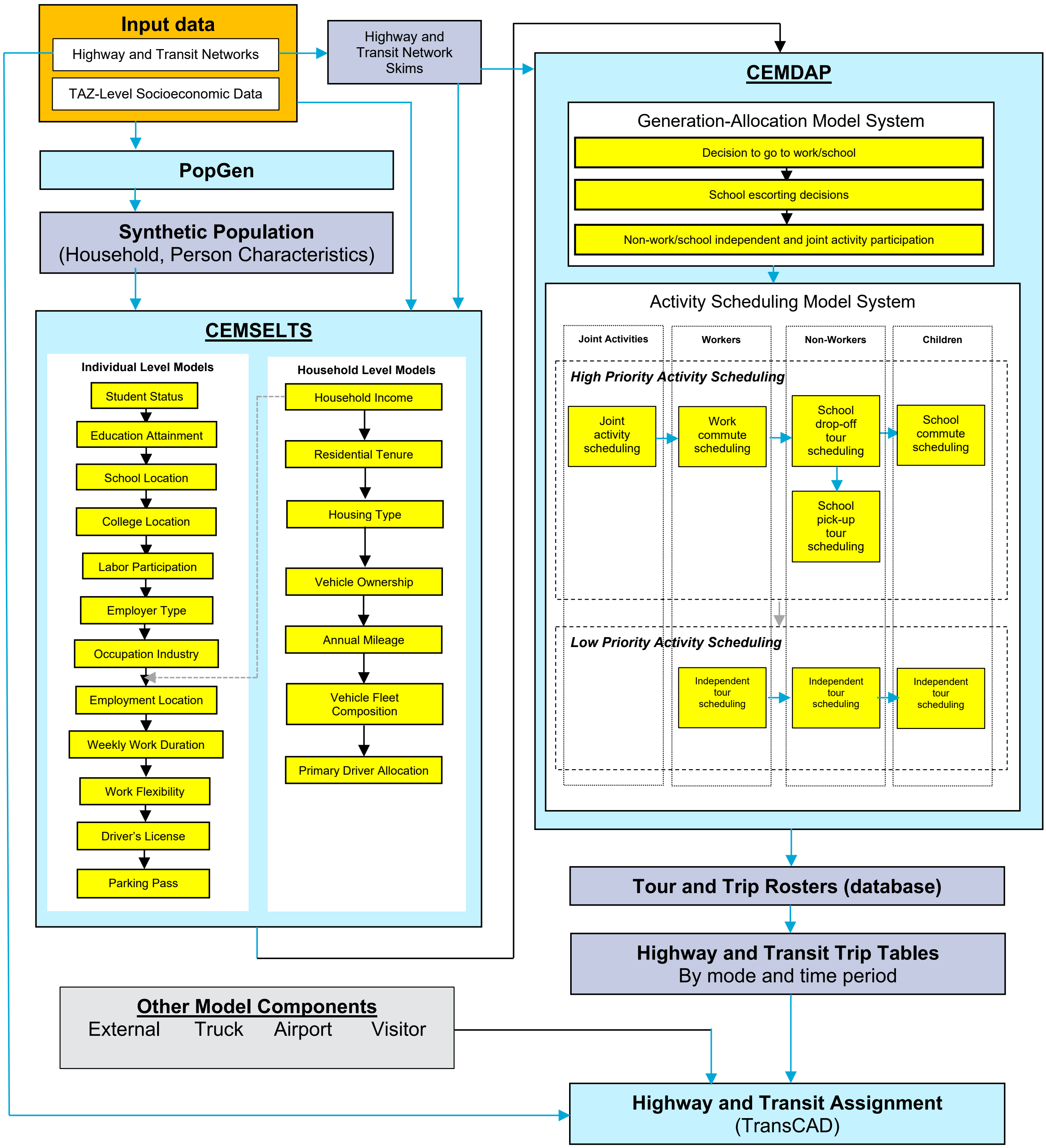
The overall activity based model system is defined by the integration of three key components:

- **PopGen**, the synthetic population generator;
- **CEMSELTS**, the socioeconomic modeling system; and
- **CEMDAP**, the activity-based modeling engine.

The other important components of the overall NYBPM include the highway and transit assignment processes, and the non-activity based components including external travel, special generators, and commercial vehicle travel. The model design is documented by Cambridge Systematics, Inc. et al (2017). The structure of the NYBPM is illustrated in Figure 1-1. (Please note that while the attributes of the joint activities predicted by the GA model system are simulated first in the scheduling system, joint activities are not actually scheduled before work or school trips.)

The parameters for the CEMSELTS and CEMDAP components were originally estimated for the 2012 NYBPM as documented by Cambridge Systematics, Inc. et al. (2018). The main model estimation data set used data from the Regional Household Travel Survey (RHTS), conducted by NYMTC and North Jersey Transportation Planning Authority (NJTPA) in 2010-2011. Other data sources, such as the National Household Travel Survey (NHTS), were used for a few CEMSELTS components, where the RHTS data was not sufficient for estimation. Generally, the RHTS and these other data sources are also used for model validation, as discussed below. Note that the 2010-2011 RHTS remains the most recent comprehensive travel survey in the NYMTC model region, and so new model estimation is not being preformed for the 2019 NYBPM update (though selected models may be reestimated later if 2019 application results show that estimation may be warranted).

Figure 1-1. 2012 NYBPM Model Structure



The Validation Manual recommends that a validation plan be developed in conjunction with the model design plan whenever a travel model is estimated or updated. The validation plan should assess the available validation data, determine what can be validated, and specify any appropriate guidelines.

The validation plan should also set the stage for quality model validation documentation. It should be noted that where numeric targets for comparisons between model results and observed data are provided, they should be interpreted as guidelines, not pass/fail tests for model validation. Comparisons to observed data are helpful in helping to understand how well the model reflects actual travel behavior, but matching "standards" is neither necessary nor sufficient to prove model validity. It must be remembered that the observed data being used for the validation might be a source of some of the error as there is always some error associated with observed data.

The Validation Manual stresses that all components of a model should be validated, including the model input data. Based on experience gained with other activity-based model validation efforts, individual model component validations should focus on how well the component reproduces distributions and data that will be passed to subsequent modeling components. While other checks may also be useful, it is quite possible for the amount of information produced for the various tests to be overwhelming. Specific checks are noted later in this document.

The focus of the 2019 NYBPM validation process includes comparisons of base year model component results to observed data and model sensitivity tests.

Sources of Error

As documented in the Validation Manual, there are several types of error that can affect models and model validation, including:

- Model specification error;
- Model estimation error;
- Model aggregation error;
- Input data error; and
- Model validation data error.

The model development process seeks to minimize the first four types of error listed above and the model validation process seeks to measure the success of the model development process in minimizing the errors. However, it is often overlooked that model validation data also are subject to error. For example, "observed" average daily traffic counts are, in effect, based on surveys of traffic. Most are estimated from actual traffic counts performed on one or two days over the course of a year and factored to "average daily traffic."

All of the above types of error will affect validation results. At some levels of aggregation, the impacts might be significant. For example, at a disaggregate validation of individual mode choice behavior, the impact might be substantial and a rho-square value of 0.2 might represent a very reasonable model. For other, more aggregate comparisons, an R^2 value in that range might suggest that the model is not reasonably reproducing the observed data. However, in either case, it must be remembered that the observed data being used for the validation might be a source of some of the error.

2.0 Data for Validation

This section identifies the data available for validation of the 2019 NYBPM. A variety of data are needed to perform the validation tests described in Chapter 5.0, but those tests are limited by the available data.

2.1 Travel and Demographic Data

2.1.1 American Community Survey Data

The American Community Survey (ACS) is conducted continuously by the Census Bureau and provides a great deal of information that can be used for validation. Because the ACS is conducted continuously, the Census Bureau can make data available every year rather than every 10 years like the decennial census (though for a smaller samples of the population). The most recent ACS data release is for the period 2015 to 2019; the Public Use Microdata Sample (PUMS) has also been released for this period.

The ACS provides data on housing and population not available from the 2010 Census. Information includes:

- Population characteristics
    - Age
    - Sex
    - Relationship to head of household
    - Income
    - Employment information including labor force status, industry, and occupation
- Household characteristics
    - Vehicles available
    - Income
    - Tenure
    - Housing value
    - Rent

The Census Transportation Planning Products (CTPP) are based on data from the ACS. In addition to providing information on the place of work of residents and their journeys to work, the CTPP also provide zonal level information in the form of cross-classifications across variables. For instance, one could examine households by size, number of vehicles, and income rather than univariate distributions provided in the data sources described above.

The Public Use Microdata Sample (PUMS) from the ACS show the full range of population and housing unit responses collected on individual ACS questionnaires, aggregated to Public Use Microdata Areas (PUMA) of about 100,000 population. For example, they show how respondents answered questions on occupation, place of work, and so forth. The records contain information from the completed ACS questionnaires for most questions for the selected subsample of housing units and group quarters persons including questions on age, sex, tenure, income, education, language spoken at home, journey to work, occupation, condominium status, shelter costs, vehicles available, and other subjects. Many multi-variate customized tabulations can be summarized from the PUMS files. Some of the CEMSELTS model components for which RHTS data were unavailable were estimated using PUMS data.

2.1.2 Travel Survey Data

Regional Household Travel Survey

The RHTS is the most recent source of observed choice preference data used for estimation of most model components. It is also an important source of model validation data since much of the information on specific travel behavior of specific travel types is unavailable elsewhere.

The RHTS was conducted by NYMTC and North Jersey Transportation Planning Authority (NJTPA) in 2010-2011. Like most household travel surveys, the survey collected information specific to each household, including information related to each person living in the household and each vehicle owned by the household. In addition, each household was assigned a travel day, when household members were asked to record all travel and the characteristics of that travel for a 24-hour period. In total, nearly 19,000 households completed the travel diary information and made nearly 144,000 (linked) trips during their designated travel days. Household were surveyed from each of the 28 counties within the New York Metro area, including 12 New York counties, 14 New Jersey counties, and two Connecticut counties.

The RHTS data are split into several categories.

1. **Basic person and household variables**. These variables, such as the age and gender of the respondent and the household size, are taken directly from the survey responses.
2. **Derived person and household variables**. These variables are derived from the data. Most of these variables, such as the amount of time spent by persons in households performing activities, deal with the activity participation of individuals in the household.
3. **Person/household location variables**. These variables describe the locations of each household and the key locations of household members, such as the regular workplace for workers and regular school location for students.
4. **Vehicle variables**. These variables, such as the number of vehicles owned, are drawn directly from the survey responses and describe the information related to vehicles owned by a household.
5. **Tour level variables**. These variables describe information at the tour level. The variables include tour mode, stay duration previous to the tour, number of stops, vehicle information, and information related to work-based subtour formation.
6. **Stop level variables**. These variables include activity purpose, duration, and location.

National Household Travel Survey (NHTS)

The National Household Travel Survey (NHTS) is the primary source of the nation's information about travel by US residents. The inventory of travel behavior includes trips made by all modes of travel and for various purposes. The most current survey (2017 NHTS) is the eighth in this series of surveys; data from the previous NHTS, in 2009, was used for the estimation of a small number of CEMSELTS components for the 2012 NYBPM. The same data will be used for validation of these components, where the RHTS does not include the necessary data to validate the specific component. For example, the RHTS did not collect data on educational attainment, and so the CEMSELTS educational attainment model was validated based on PUMS data.

Citywide Mobility Survey

Each year, the New York City Department of Transportation (NYC DOT) conducts a travel survey called the Citywide Mobility Survey (CMS). Launched in 2017, it seeks to assess the travel behavior, preferences, and attitudes of New York City residents. The data from this survey will be reviewed to determine how it could best be used for this model update.

2.1.3 Transit Rider Surveys

The region has many transit operators that collect survey data for their systems. The NYBPM includes the services of dozens of transit operators and thousands of routes. Many of these operators are smaller companies and/or private operators with no obligation to report information about their operations or passengers.

However, the major carriers are public agencies, and many have collected data on patronage for planning and customer service purposes. The extent of data collected varies considerably across operators, but many of the major operators have recent, usable, full on-board surveys that are sufficient for validation at the boarding station, access mode, and time-of-day level. Major operators/systems that have available survey data sets include the following (year of survey shown in parentheses):

- Port Authority of New York and New Jersey (PANYNJ) Trans-Hudson Bus Surveys (2015)
- PANYNJ Bergen-Passaic Bus Survey (2009)
- New Jersey Transit (NJ Transit) New Jersey Local Bus (2013)
- NJ Transit Newark Local Bus Survey (2013)
- Metro North Origin-Destination Survey (2017)
- Long Island Railroad Origin-Destination Survey (2012)
- Raritan Valley Commuter Rail Survey (2014)
- Montclair Boonton Commuter Rail Survey (2015)
- Morris and Essex Commuter Rail Survey (2016)
- Main Bergen and Pascack Valley Commuter Rail Survey (2013)
- Northeast Corridor and New Jersey Coast Commuter Rail Survey (2014)
- Secaucus Origin-Destination Survey (2013)
- Newark Airport Commuter Rail Survey (2014)
- NJ Hudson Bergen Light Rail Survey (2017)
- Newark Light Rail Survey (2007)
- Port Authority Trans-Hudson (PATH) Survey (2012) (checking on availability of newer version for possible use in validation)
- NJ-based Ferry Survey (2013)

These surveys provide boarding times, boarding station or terminal, and access modes and have been expanded to counts at the boarding location for the year of the survey. These are the most recent datasets available for each survey at this time.

Finally, the MTA household travel survey is a sample of residents in the five boroughs and is expanded to households and not to station boardings. Therefore, boardings will be under-counted since many passengers from outside the five boroughs board MTA buses and subways. While the MTA household travel survey is a very good source for information on travel behavior in the five boroughs, to validate the transit assignment we need more accurate boarding estimates. This is why we have chosen to use the MTA boarding data.

2.1.4 Passively Collected Travel Data

LOCUS

LOCUS is a data product that uses anonymized smartphone location-based services (LBS) data to produce travel data, expanded to represent travel within the NYMTC model region. It is discussed as part of task 10A in the scope of work for this project. The LOCUS data obtained by NYMTC will include a representative sample of complete daily activity patterns using one year's worth of data obtained for 2019, summarized to provide tour and trip level information on travel flows by time of day. This will include the origin, destination, and time of day of each trip. While complete demographic information for travelers is not available, LOCUS is able to impute the home location of each traveler as well as locations traveled to on a regular, frequent basis, generally the work or school location. Since the home locations of the smartphones are known, travel made by visitors can be kept separate from travel made by residents of the model region.

LOCUS can be used as a validation source for aggregate origin-destination flows by time of day. Since information about detailed trip/tour/activity purpose, travel mode, or user demographics is not available from LOCUS, model results by demographic segments or disaggregate tour purposes cannot be compared. However, because of its much larger sample size than available survey data, LOCUS can be used for validation of any spatial or temporal segmentation of travel flows. Specifically, LOCUS data can be used to validate overall O-D person flows across all modes, as well as for the visitor model.

Traffic Counts

Traffic counts have been assembled from a variety of credible sources, including the New York State Department of Transportation (NYSDOT), New York City Department of Transportation (NYCDOT), NYMTC's Hub-bound database (2019), 2016 Bridge Report, New Jersey Department of Transportation (NJDOT), New Jersey Turnpike Authority, and Connecticut DOT. Point data were requested from each source, and a significant amount of traffic count data was obtained for the NYBPM model region. The 2019 Brooklyn Bridge data were revised in April 2021. Please refer to the following addendum: ADDENDUM-to-HB2019_April_2021.pdf (nymtc.org)

The emphasis was on ensuring that these point data are tagged to the right model network link. The large volume of point data allowed the team to use data that were consistent and to discard those that appeared problematic. The emphasis has been on the screenline links, which have been thoroughly checked and tested including checks for representation by functional class, direction, and geography. On critical links where count data were not available, such as crossings and external stations, data were pulled in from older sources and factored to represent 2019 conditions.

Where possible, counts have been segmented by vehicle class (auto, heavy truck, medium truck, taxi/for-hire vehicles, and bus) from the NYSDOT and NYCDOT databases. Similarly, taxis have been separated from other vehicles in counts for New York City and were already separated in the data we were provided with.

Transit Boarding Counts

Several sources of transit boarding counts are being used. Please note that this data is over and above the data from the Transit Rider Surveys described in Section 2.1.3:

- **2019 Hub-Bound Report**—The Hub-bound report provides transit person trips for a fall business day in 2019 by mode and time of day across five screenlines. A summary of the 2019 a.m. peak period Hub-bound trips is shown in Table 2-1. To check the consistency of Hub Bound with other data sources, boardings from the transit rider surveys, noted above, will be reconciled to the totals in this table. (Note that trips into the HUB do not mirror trips out of the HUB and do not necessarily represent origins & destinations.) To maintain consistency with the NYBPM mode definitions, the modes shown in the Hub-bound report may be combined (and bicycle will be removed since the model does not assign bicycle trips).

Table 2-1. Hub-Bound 2019 AM Peak Period Transit Trips by Mode

Inbound	Local/ Express Bus	Private Bus	Subway	Commuter Rail	Private Ferry	Staten Island Ferry	Bicycle	Tram	Total
Queens	9,377		246,172	85,218	2,058		1,134	1,154	345,113
Brooklyn	14,774		401,874		1,151		3,544		421,343
Staten Island						14,328	202		14,530
New Jersey		116,186	84,317	62,451	14,414				277,368
60th Street Sector	12,510		314,619	79,721			4,719		411,569
Inbound Total	**36,661**	**116,186**	**1,046,982**	**227,390**	**17,623**	**14,328**	**9,599**	**1,154**	**1,469,923**
Outbound	**Local/ Express Bus**	**Private Bus**	**Subway**	**Commuter Rail**	**Private Ferry**	**Staten Island Ferry**	**Bicycle**	**Tram**	**Total**
Queens	209		47,480	5,618	298		188	304	54,097
Brooklyn	278		86,563		200		572		87,613
Staten Island						2,835	17		2,852
New Jersey		32,178	16,198	10,233	1,036				59,645
60th Street Sector	5,257		161,053	6,685			2,705		175,700
Outbound Total	**5,744**	**32,178**	**311,294**	**22,536**	**1,534**	**2,835**	**3,482**	**304**	**379,907**
Grand Total	**42,405**	**148,364**	**1,358,276**	**249,926**	**19,157**	**17,163**	**13,081**	**1,458**	**1,849,830**

- **2019 NYCT Subway Boardings**—The MTA has provided detailed boarding data by station complex, station entry, line, and time of day. In addition, a synthesized station-boarding to station-alighting matrix by hour has also been provided. This data can be aggregated to groups of stations and groups of lines for direct comparison to model results.

- **2019 NYCT Bus Boardings and Alightings by Stop**—The MTA has provided detailed data on bus boardings. This data can be aggregated to daily bus boardings by county or any other geographical aggregation.

- **National Transit Database**—The National Transit Database provides estimated unlinked transit trips at an aggregate, system level. For non-commuter, non-MTA, local bus providers, total transit trips will be estimated from this federal repository.

Please note that while commuter rail and PATH boardings are not included in this section, PATH Boardings are listed under section 2.1.3, Transit Rider Surveys. These boardings are also shown as checks to be done for transit assignment in Section 3.6.

3.0 Validation of Model Components

This chapter discusses the specific components of the validation process. These include tests of the socioeconomic and network input data used, tests for activity-based demand model components, highway and transit assignment checks, and sensitivity tests.

The validation checks for model input data are discussed in 3.1. Section 3.2 discusses validation tests specifically for the activity-based demand model components. Highway assignment checks are described in Section 3.3, and assignment checks are described in Section 3.6.

3.1 Input Data Checks

The two main types of model input data are socioeconomic data and transportation network data.

3.1.1 Socioeconomic Data

The zonal-level data available as inputs into the model include:

- Number of households
- Population
- Employment by type (retail, office, other)
- K-12 educational enrollment
- University enrollment

The CS team will continue to rely on the accuracy of inputs provided by NYMTC, as the socioeconomic data goes through an extensive QA/QC process before it is provided to the CS team. This includes comparison with other data sources, such as Census data. However, given that the household and population data form the control totals for PopGen, and the employment-based and educational enrollment data are instrumental in several model components, the CS team will work with NYMTC if issues arise as the data is used in the model.

Table 3-1 presents a list of validation checks by level of aggregation that should be considered. No specific criteria have been specified for the validation tests; rather, the reasonableness of the data should be gauged by potential impacts on model results.

Table 3-1. Socioeconomic Data Validation Tests

AGGREGATION LEVEL	DATA ITEM	VALIDATION TEST

County	• Number of households
	• Population in households
	• Employment by Type
	  • Retail
	  • Office
	  • Other
		• Compare households and persons totals to original NYMTC data –should match at county level
		• Compare employment by type to other sources, including previous NYMTC estimates – results should be “close”

Synthetic Population

The NYBPM uses the PopGen synthetic population generator. Table 3-2 summarizes the reasonableness tests for household data that will be produced for base year 2019. The primary reasonableness tests will be socioeconomic distributions stratified by various geographic strata.

Table 3-2. Synthetic Population Generator Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOME
County	• Total population • Group quarters population • Total households • Employed labor force • Distribution of households by size • Distribution of persons by age • Distribution of persons by gender • Distribution of persons by age and gender	• PopGen outputs should closely match control totals • Comparisons to ACS data where feasible

In some cases, such as the comparisons described in Table 3-5, numeric targets are provided as guidelines for the comparisons. (As noted in Section 1.1, hitting these targets should not be treated as pass/fail tests.) For other tests, such as those described in Table 3-2, no numeric targets are provided. In these cases, the error levels for the observed data are unknown and may be high (due to small survey sample sizes for certain segments, for instance). The recommendations are therefore to check for "major differences" or "reasonable patterns." Judgment on the part of NYMTC and the model development team will be needed to assess the impacts of differences from observed data on model results.

It should be noted that some person and household characteristics are estimated later by CEMSELTS. The tests of the model components that estimate these characteristics are discussed in Section 3.2.

3.1.2 Transportation Network Data

The transportation network data and the path-building procedures that use these data will be checked. As with the socioeconomic data tests, there will not be any specific criteria by which these network tests will be measured. Reasonableness of the data will be gauged by potential impacts on model results.

Highway Network Checks

The highway network was checked as part of the 2012 NYBPM development, and the network was updated to 2019 in this project. This process will be documented as part of task 5: Update Base Year Integrated Network. The new highway network will be spot checked to confirm accuracy. This will include checks of the coding for the Manhattan side of the Queensboro Bridge to ensure accurate connections between the bridge roadways, ramps, and local streets. Estimated highway travel times for a series of origin-destination pairs will be compared to observed travel times. This will help to confirm the appropriateness of the entire network processing procedures including assignment of free flow speeds and link capacities, and the how volume-delay functions relate traffic to reduced (congested) travel speeds.

There were a number of network-related issues that were identified as part of the 2012 NYBPM highway assignment validation. These included some network coding problems, centroid connector locations, and stub links. Some of these issues were not able to be addressed during the 2012 validation process; the network will be checked with regard to the remaining issues and corrected as appropriate.

Transit Network Checks

Transit networks will be examined using the following tests:

- Review rules for defining walk and auto access
- Spot compare coded fixed guideway running time estimates by time of day to schedules
- Spot compare coded headways by time-of-day for routes to timetable values
- Compare travel times for a series of origin-destination pairs to observed travel times

3.2 Activity-Based Component Checks

A systematic procedure was developed to validate the activity-based demand model components during the 2012 NYBPM project. This process involved developing Excel templates for the components, summarizing the relevant observed data from sources such as the NYMTC Regional Household Travel Survey (RHTS), and reading the model results from the PostGreSQL database into the Excel files. The comparisons were automatically displayed, and a tab in each file was used to track the parameter changes made during the calibration process.

These Excel templates will be reused for the 2019 NYBPM validation, with revisions made as needed to reflect changes from the 2012 NYBPM.

3.2.1 CEMSELTS Components

The following are the main components of CEMSELTS to be validated:

- Person level characteristics
    - Education attainment
    - Labor force participation
    - Employer type
    - Occupation Industry
- Location choices
    - School location
    - College location
    - Work location
- Work activity characteristics
    - Work duration
    - Work flexibility
    - Mobility choices
- Household characteristics
    - Household income
    - Residential tenure
    - Housing type
    - Annual household mileage
    - Vehicle fleet and primary driver allocation
    - Vehicle availability

Table 3-3 presents the validation tests for the person level characteristics. The school and college location choice tests are shown in Table 3-4, and Table 3-5 presents the tests for workplace location choice. Table 3-6 shows the validation tests for the work activity characteristics, and the tests for the household characteristics are presented in Table 3-7.

Table 3-3. CEMSELTS Person Level Characteristics Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOME
Regional	• Percentage of persons for each choice alternative ○ Total ○ By subregion ○ By gender ○ By age group	• Compare to RHTS/PUMS data summaries and look for major differences

Table 3-4. CEMSELTS School/College Location Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES
Region	• Modeled versus observed (from RHTS) home-to-school distance histogram • Modeled versus observed (from household survey) average impedances (same as above) stratified by: • State • Subregion	• Modeled to observed averages should be ±1 to 2 miles • Modeled to observed distance histogram coincidence ratios[1] > 0.7 (school)/> 0.6 (college)
Region	• Modeled versus observed (from RHTS) percentages of intrazonal flows for: • Region • Subregions	• Most should be within five percentage points

Note:

1. Coincidence ratio is a measure of fit between two distributions, in this case the observed and modeled distance frequency distributions. Mathematically, it is the area under both curves divided by the area under either curve.

Table 3-5. CEMSELTS Workplace Location Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES
Region	• Modeled versus observed (from RHTS) home-to-work distance histogram • Modeled versus observed (from household survey) average impedances (same as above) stratified by: • State • Subregion	• Modeled to observed averages should be ±1 to 2 miles • Modeled to observed distance histogram coincidence ratios[1] > 0.7
Region	• Modeled versus observed (from RHTS) percentages of intrazonal flows for: • Region • Subregions	• Most should be within five percentage points
Subregion	• Modeled versus ACS home-to-work flows by subregion	• Modeled percentage of regional flows for each subregion-to-subregion pair should be within 1 to 2 percentage points
Region/Subregion	• Modeled workplaces versus employment by subregion	• Ratio of workplaces to employment should be 0.80 to 0.95 for the region • Ratios of workplaces to employment should be 0.75 to 1.00 for each subregion

Table 3-6. CEMSELTS Work Activity Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOME
Regional	• Percentage of persons for each choice alternative ○ Total ○ By Subregion ○ By Gender ○ By Age Group	• Compare to RHTS data summaries and look for major differences

Table 3-7. CEMSELTS Household Level Characteristics Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOME
Regional	• Percentage of households for each choice alternative ○ Total ○ By subregion ○ By number of workers (where appropriate) ○ By income level (where appropriate)	• Compare to RHTS/PUMS/NHTS data summaries and look for major differences

3.2.2 CEMDAP Components

There are many components of CEMDAP to be validated. These components are (with component codes from the model design plan shown):

- Generation-allocation model system
  - GA1, GA6 – Decision to go to school
  - GA2, GA3, GA7, GA8 – School start/end times
  - GA4 – Decision to go to work
  - GA5 – Work start and end times
  - GA9, GA10 – Travel mode to/from school
  - GA11, GA12 – Allocation of school drop-off/pick-up
  - GA13, GA14 – Out-of-home duration
  - GA15, GA16 – Independent/joint activity participation
  - GA17 – Decision to undertake serve-passenger activities
- Activity scheduling model system – workers
  - WSCH1 – Commute mode
  - WSCH2, WSCH3, WSCH4 – Number of before-work, work-based, after work tours
  - WSCH5, WSCH6, WSCH7– Before-work, work-based, after work tour modes
  - WSCH8 – Number of stops in a tour
  - WSCH9 – Home or work stay duration before the tour
  - WSCH10 – Activity type at a stop
  - WSCH11 – Activity duration at stop
  - WSCH12 – Travel time to a stop
  - WSCH13 – Location of a stop
  - WSCH14 – Worker trip mode choice
- Activity Scheduling Model System – Non-Workers
  - NWSCH1 – Number of independent tours
  - NWSCH2, NWSCH3 – Decision to undertake an independent tour before/after pick-up or joint discretionary tour
  - NWSCH4 – Tour mode
  - NWSCH5 – Number of stops in a tour
  - NWSCH6 – Number of stops following a pick-up/drop-off
  - NWSCH7 – Home stay duration before a tour
  - NWSCH8 – Activity type at stop
  - NWSCH9 – Activity duration at stop
  - NWSCH10 – Travel time to stop
  - NWSCH11 – Stop location
- Joint discretionary tour scheduling model system
  - JASCH1 – Decision of joint or separate travel
  - JASCH2 – Joint activity start time
  - JASCH3 – Joint activity travel time to stop
  - JASCH4 – Joint activity location
  - JASCH5 – Vehicle used for joint home-based tour
- Activity scheduling model system – children
  - CSCH1 – School to home commute time
  - CSCH2 – Home to school commute time
  - CSCH3 – Mode for independent discretionary tour
  - CSCH4 – Departure time from home for independent discretionary tour (time from 3 a.m.)
  - CSCH5 – Activity duration at independent discretionary stop
  - CSCH6 – Travel time to independent discretionary stop
  - CSCH7 – Location of independent discretionary stop

We have organized the types of validation tests to reflect the focus of each particular model component and to group components by the types of outputs they produce. The following typology is used to describe the validation tests:

- **Activity choices**
  - GA1, GA6 – Decision to go to school
  - GA4 – Decision to go to work
  - GA13, GA14 – Out-of-home duration
  - GA15, GA16 – Independent/joint activity participation
  - GA17 – Decision to undertake serve-passenger activities
  - GA11, GA12 – Allocation of school drop-off/pick-up
  - WSCH2, WSCH3, WSCH4 – Number of before-work, work-based, after-work tours
  - WSCH8 – Number of stops in a tour (worker)
  - WSCH10 – Activity type at a stop (worker)
  - NWSCH1 – Number of independent tours
  - NWSCH2, NWSCH3 – Decision to undertake an independent tour before/after pick-up or joint discretionary tour
  - NWSCH5 – Number of stops in a tour (non-worker)
  - NWSCH6 – Number of stops following a pick-up/drop-off
  - NWSCH8 – Activity type at stop (non-worker)
  - JASCH1 – Decision of joint or separate travel
- **Timing/scheduling choices**
  - GA2, GA3, GA7, GA8 – School start/end times
  - GA5 – Work start and end times
  - WSCH9 – Home or work stay duration before the tour (worker)
  - WSCH11 – Activity duration at stop (worker)
  - NWSCH7 – Home stay duration before a tour (non-worker)
  - NWSCH9 – Activity duration at stop (non-worker)
  - JASCH2 – Joint activity start time
  - CSCH4 – Departure time from home for independent discretionary tour (time from 3 a.m.)
  - CSCH5 – Activity duration at independent discretionary stop
- **Location choice/travel time models** (note that work and school locations have already been simulated in CEMSELTS)
  - WSCH13 – Location of a stop (worker)
  - NWSCH11 – Stop location (non-worker)
  - JASCH4 – Joint activity location
  - CSCH7 – Location of independent discretionary stop
  - WSCH12 – Travel time to a stop (worker)
  - NWSCH10 – Travel time to stop (non-worker)
  - CSCH1 – School to home commute time
  - CSCH2 – Home to school commute time
  - CSCH6 – Travel time to independent discretionary stop
  - JASCH3 – Joint activity travel time to stop
- **Mode choices**
  - GA9, GA10 – Travel mode to/from school
  - WSCH1 – Commute mode
  - WSCH5, WSCH6, WSCH7– Before-work, work-based, after work tour modes

- WSCH14 – Worker trip mode choice
- NWSCH4 – Tour mode (non-worker)
- JASCH5 – Vehicle used for joint home-based tour
- CSCH3 – Mode for independent discretionary tour

Table 3-8 presents the validation tests for the activity choices group. The tests for the timing/scheduling choices group are presented in Table 3-9. Table 3-10 presents the location choice tests while Table 3-11 shows the mode choice tests.

Table 3-8. CEMDAP Activity Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES
Regional	• Percentage of persons/households for each choice alternative (activity participation, escorting choice, activity type), segmented by (as appropriate for the specific choice) o Total region o Subregion o Gender o Age group/grade level o Employment status o Work duration o Household size o Income level o Number of vehicles	• Compare to RHTS data summaries and look for major differences
Regional	• Percentages of persons making tours/stops—as indicated by the specific choice--segmented by (as appropriate for the specific choice): o Total region o Household size o Income level o Age group o Employment status o Tour mode	• Compare to RHTS data summaries and look for major differences

Table 3-9. CEMDAP Timing/Scheduling Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES
Regional (tour or stop level, depending on the specific choice model)	• Histograms of tour arrival/departure times to/from primary tour destination by tour purpose:	• Compare modeled to observed temporal distributions from RHTS • Review for reasonable patterns
Regional (tour or stop level, depending on the specific choice model)	• Average activity durations by (as appropriate for the specific choice): • Gender • Age group • Household income level	• Compare modeled to expanded observed temporal distributions • Review for reasonable patterns

Table 3-10. CEMDAP Location Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES
Aggregate (tour/trip level)	• Modeled versus observed (from RHTS) distance histograms • Modeled versus observed (from RHTS) average distances segmented by subregion • Modeled versus observed (from RHTS) intrazonal percentages	• Modeled average distances should be within about half a mile of observed • Modeled to observed impedance histogram coincidence ratios > 0.6 (>0.7 for work) •

Table 3-11. CEMDAP Mode Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES

Aggregate (tour/trip level)	• Mode shares segmented by:
	  • Entire region
	  • Subregion
	  • Household size
	  • Vehicles available
	  • Income level
	  • Age level
	  • Gender
		• Compare modeled to observed mode shares
		• Review for reasonable patterns

3.3 Validation of Origin-Destination Flows

The main outputs of CEMSELTS/CEMDAP are the rosters of trips that represent all travel performed by the synthetic population. These rosters include the following information:

- Home location of the traveler
- Demographic information about the traveler and their household
- Tour type (work, school, non-mandatory, joint, work-based subtour, etc.)
- Trip origin and destination (TAZ)
- Trip time of day (segmented per the specifications of the appropriate submodel)
- Trip mode

The LOCUS data set provides a means for checking overall origin-destination flows by time of day and aggregate trip purpose (home-work/school, home-other, and non-home based) that was not available for the 2012 model validation. While some of the checks described in Section 3.2, such as the comparison of modeled to ACS home-to-work flows by subregion (Table 3-5) involve comparison of modeled origin-destination flows to observed data for specific segments, the RHTS does not provide a sufficient source to compare origin-destination flows for all person trips (and it is almost a decade old compared to the model base year).

While LOCUS does not provide information on traveler demographics, detailed travel purposes, or mode choice, it does provide detailed information on trip start and end locations and time of day of travel. The LOCUS data also segments trips by whether they are part of home based work/school tours, work-based subtours, or other tours.

The comparisons between the trip rosters and the LOCUS data set will be performed using the following process:

1. The LOCUS data set will be segmented into the following:
    - Trips made on home based tours to work/school
    - Work-based subtours
    - All other tours
    - Trips made by non-residents of the region (these will be removed from the LOCUS data set and used for validation of the visitor model, as described in Section 3.4).
2. The NYBPM trip rosters will be segmented the same way
3. Both the segmented LOCUS data set and the NYBPM trip rosters will be split into the four time periods used for highway assignment (a.m. peak (6:00 AM – 10:00 AM), mid-day (10:00 AM - 3:00 PM), p.m. peak (3:00 PM – 7:00 PM), and night (7:00 PM – 6:00 AM).
4. The following comparisons will be performed between the LOCUS data set and trip rosters by time period:
    - Total trips by subregion by tour type
    - Total trips between subregions

These comparisons will be used to determine whether any of the CEMSELTS and CEMDAP submodels should be further calibrated to achieve better matches of origin-destination flows.

3.4 Visitor Model

The visitor model is a separate trip-based component of the NYBPM. For the 2012 model validation, no additional information on visitor travel was available beyond the NYMTC Regional Establishment Survey (RES), which was used for visitor model estimation. The RES was limited in terms of sample size and by the fact that its sample frame was a set of hotels. The RES therefore did not capture travel made by visitors who did not stay in hotels in the model region. However, surveying at hotels was the best means of capturing travel by visitors from outside the New York Region because, except for extended stay hotels, nearly all guests are visitors from outside the region. The usefulness of a data source for validation of the visitor model depends on what data is provided, in what form, and at what geographic level. For the 2019 model validation, the LOCUS data set provides a more robust and recent data source for visitor model validation because it is a consistent source for movements by visitors (people whose imputed home locations are outside the region). The subset of the LOCUS data set related to non-residents of the model region will be compared to the trip tables created by the visitor model. Non resident trip flows can be separated from resident trip flows in the LOCUS data because the home locations of devices are able to be imputed based on the locations/durations by time of day. Any devices whose homes are located outside the model region are therefore visitors, and their travel within the region would be covered by the visitor model.

It should be noted that the model trip tables include information about visitor type (business or leisure), trip purpose, modes used, and user demographics that will not be available in the LOCUS data. The comparison therefore will be done for the overall visitor person trip table to the trip table of the LOCUS data for visitors. The comparison will be at the subregion-to-subregion level for the four time periods used in the visitor model (a.m. peak, mid-day, p.m. peak, and night).

3.5 Highway Assignment Checks

Highway assignment validation is focused on the comparing modeled link traffic volumes to observed volumes. It should be recognized that even a poorly specified model can be made to reproduce observed traffic volumes for a base year. The validation of the individual modeling procedures described in Section 3.2 are intended to help ensure that the regional travel model is, in fact, reasonable. However, if the individual model components are all deemed to be reasonable but the modeled traffic volumes do not reasonably reflect observed traffic volumes for a base year, the model is not valid.

The highway assignment validation will focus on several classes of measures:

- Vehicle-miles of travel (VMT);
- Individual link traffic volumes; and
- Intra-regional traffic flows as defined by approximately 65 screenlines and eight corridors

There are several types of individual checks within each of these classes.

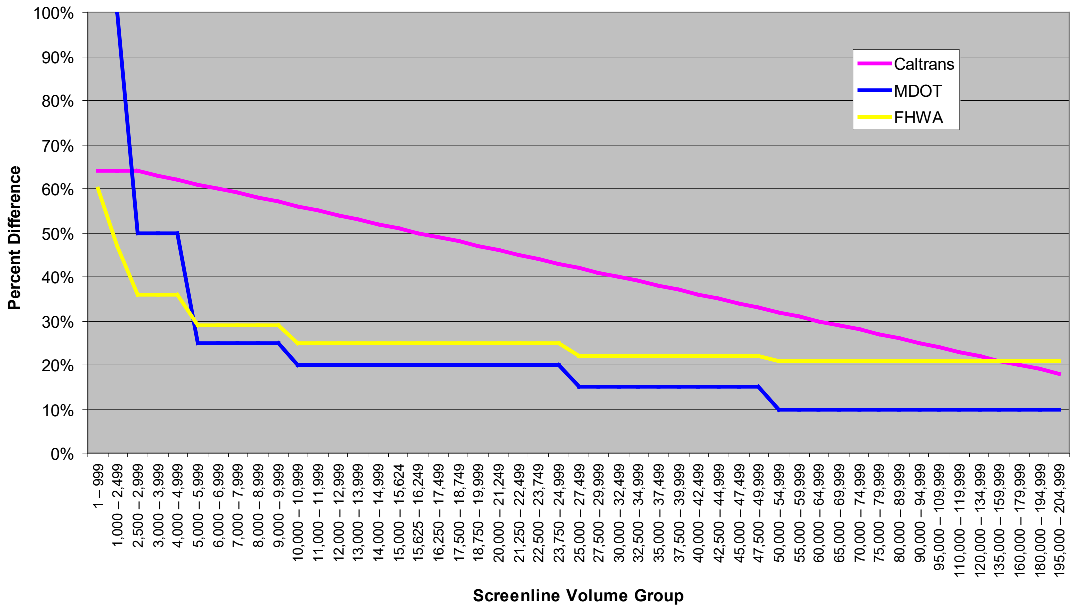
Table 3-12 summarizes the highway assignment validation measures, which focus much more specifically on numerical guidelines than the validation of the individual model components. A primary reason for this is the availability of independently collected data such as traffic counts for the system level validation.

Table 3-12. Highway Assignment Validation Tests

VALIDATION FOCUS	VALIDATION MEASURES	EXPECTED OUTCOMES
Vehicle-miles of travel	• Comparison of modeled VMT to VMT estimated from traffic counts by: • Region • Subregion • Facility type • Time period	• Modeled regional VMT should be within the following percentages of estimated VMT: • ±1 percent for the region • Percentages shown in Table 3-13 by facility type on links with counts
Individual link volumes	• Plot of absolute and percent difference between modeled and observed volumes for links with counts • Scatterplot of modeled versus observed daily traffic volumes by link • Percentage root mean square error (%RMSE) by: • Region • Volume group • Anomalous links • Links with 0 volumes • Links with very high v/c ratios	• Visual inspection for large errors in modeled link volumes or for general trends in errors • %RMSE by volume group should be within the targets shown in Table 3-14
Screenlines	• Percent deviation by screenline	• Percent deviation should be within the targets shown in Figure 3-1

Table 3-13 presents the targets for VMT by functional class and area type, which is the state of practice guidance. Figure 3-1 shows the targets for percentage difference by volume level for screenlines, and Table 3-14 presents the assignment results by RMSE% by volume group.

Figure 3-1. Screenline Crossing Percent Difference by Volume Group



Source: Adapted from *Las Vegas Travel Demand Model Guidelines for Estimation, Calibration, & Validation*, prepared for Regional Transportation Commission of Southern Nevada, prepared by Fehr & Peers Transportation Consultants, March 25, 2005, page 28.

Table 3-13. Traffic Assignment VMT Targets by Functional Class

Stratification	VMT Target
Functional Class	
Freeways	±7%
Expressways	±7%
Principal Arterials	±10%
Minor Arterials	±10%
Collectors	±15%
All Links	±1%

Source: Adapted from Giaimo, Gregory, Travel Demand Forecasting Manual 1 – Traffic Assignment Procedures, Ohio Department of Transportation, Division of Planning, Office of Technical Services, August 2001.

Table 3-14. %RMSE Targets by Functional Class

Volume Group	Target Values
0 - 1,000	<200%
1,000 - 5,000	<100%
5,000 - 10,000	<45%
10,000 - 20,000	<35%
20,000 - 30,000	<26%
30,000 - 50,000	<24%
50,000 - 100,000	<21%
100,000 and up	<12%
All Links	<40%

3.6 Transit Assignment Checks

While transit trips for an average weekday are simulated by CEMDAP, the existing NYBPM transit assignment process focuses on a.m. peak period trips. The transit assignment validation effort therefore focuses on this time period.

The transit assignment validation is perhaps the most challenging part of the model validation process. The challenges are due not only to the complexity of the transit system in the region—with its numerous operators, wide variety of modes, and significant transfer activity—but also to the difficulty in collecting data and the inconsistencies among data collected at different times by different methods.

Building on the experience of validating the 2012 NYBPM, significant attention is being paid to resolving data inconsistencies to the extent possible prior to the beginning of the validation effort. In addition, transit assignment validation tests are being prioritized based on the reliability of the data sources used and the importance of the corresponding model results to understanding travel within the region, both by transit and other modes.

With this in mind, the following priorities have been set for the transit assignment validation:

1. **Comparison of assigned transit trips to the 2019 Hub-bound report**. As noted in Section 2.1.4, the numbers in the Hub-bound report (shown in Table 2-1) will be combined to reflect a more aggregate set of modes (e.g., bus, commuter rail, ferry, and subway/PATH), and the assignment results will be compared. We have worked closely with the PANYNJ to understand and address the discrepancy in NJ bus ridership between the hub-bound report and PA's report.

2. **Commuter rail boarding-alighting summary by direction by geographic segment**. Some initial thoughts regarding the segmentation to be used are as follows:

   - Brooklyn/Queens to Manhattan
   - Bronx to Manhattan
   - Nassau/Suffolk Counties to Manhattan
   - Nassau/Suffolk Counties to Brooklyn/Queens
   - Nassau/Suffolk Counties to Nassau/Suffolk Counties
   - Westchester/Putnam/Dutchess/Bronx Counties plus Connecticut to Manhattan

- Westchester/Putnam/Dutchess/Bronx Counties plus Connecticut to Westchester/Putnam/Dutchess/Bronx Counties plus Connecticut
- Rockland/Orange Counties plus New Jersey to Manhattan
- Rockland/Orange Counties plus New Jersey to Rockland/Orange Counties plus New Jersey

As explained in more detail below, targets must be based on known research of observed data. Such research does not exist for commuter rail alightings, and so targets are not used here.

3. **Subway "first boardings" by borough, and PATH "first boardings."** A "first boarding" refers to the first time a rider enters the subway/PATH system, as opposed to inter-subway transfers. Transfer passengers from NJT to PATH at Newark Penn Station are considered "first" boardings on the PATH system.

4. **Local bus boardings by subregion**. Subregions will be defined in consultation with NYMTC but may include the New York City boroughs and one to three subregions outside the City.

5. **Commuter/express bus boardings by subregion**. Subregions will be defined in consultation with NYMTC based on commuter/express bus service areas (e.g., New Jersey, Rockland/Orange, Westchester/Putnam/Dutchess/Connecticut, and Long Island).

No specific numeric targets have been defined for these checks due to the uncertainties about the error associated with the observed data sources. Numeric targets must be based on research across the industry regarding error levels in observed data. We do not have such research or guidance regarding the error levels for the various data sources that will be used for transit assignment validation (indeed, the different sources undoubtedly have different error levels), and such targets would be arbitrary and a matter of opinion. We believe that the use of the term "standards" is not appropriate for travel model validation as it implies that matching base year observed data is the most important measure of how well a travel model produces necessary information for planning and forecasting.

The priorities shown above will be followed; for example, the first priority will be to compare to the Hub-bound report, and larger differences in subsequent tests may be deemed acceptable if calibration adjustments to make improvements would result in worse comparisons to the Hub-bound report.

4.0 Sensitivity Tests

One goal of activity-based models is an increased sensitivity to model inputs that are known or believed to affect travel behavior. Because of the nature of activity-based models, a richer set of such input variables can be considered, and the sensitivity of the model results to these inputs can be checked. Sensitivity testing involves revising key factors and observing the effects on forecasted travel. These revisions can be made to model parameter values (e.g., the mode choice cost coefficient) or to model inputs (e.g., land use variables, socioeconomic conditions, parking costs, etc.).

Sensitivity tests can be performed for any of the validation measures described in the previous chapters. Typically, however, data is not available to compare the results of sensitivity tests to. Instead, as with forecasts beyond the model's base year, sensitivity tests should be reviewed for reasonableness, with expected outcomes of the tests noted beforehand. Any unexpected outcomes observed from the tests should be explainable. It is easiest to perform sensitivity testing using the validated base year model scenario as the basis for comparison (even though, in some cases, it might be unrealistic for the particular changes to apply to the short-term).

The model sensitivity tests performed for the 2012 NYBPM were:

- Changes in parking costs
- Changes in roadway tolls (perhaps varying by time of day)
- Adding a major development in a specific location in the region
- "Autonomous vehicle package," where a set of changes to the transportation system and travel behavior that might be expected when autonomous vehicles become commonplace.

The specific sensitivity tests to be defined for the 2019 NYBPM should reflect the uses of the model for planning purposes. As was the case for the 2012 NYBPM validation, the definitions for the sensitivity tests will be developed in consultation with NYMTC staff. Along with NYMTC staff, the Model Advisory Committee will have the opportunity to review the results of the sensitivity tests.