CAMBRIDGE SYSTEMATICS

New York Best Practice Model 2019 Update

Summary Report

draft

report

prepared for

New York Metropolitan Transportation Council

prepared by

Cambridge Systematics, Inc.

with

Gallop Corporation
EA Harper Consulting

June 2023

www.camsys.com

draft report

New York Best Practice Model 2019 Update

Summary Report

prepared for

New York Metropolitan Transportation Council

prepared by

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

with

Gallop Corporation
EA Harper Consulting

date

June 2023

Table of Contents

List of Tables

List of Figures

1.0 Introduction

This report documents a project undertaken by the New York Metropolitan Transportation Council (NYMTC) to update their activity-based travel demand model for the greater New York region. This model, named the New York Best Practice Model (NYBPM), was recently updated to reflect a base year of 2019, by a team led by Cambridge Systematics, Inc. (CS). Other team members included EA Harper Consulting and Gallop Corporation.

1.1 Project Objectives

The NYBPM was developed to be capable of providing the necessary model-related information to determine whether performance targets are being met and the effects of the policies and investments involved with any planning scenarios on the achievement of these targets. With these requirements in mind, the 2019 NYBPM has the following capabilities:

- Production of measures of travel demand at aggregate and disaggregate levels. These include roadway traffic volumes and speeds/delays and transit boardings at stations and routes, as well as aggregate measures such as vehicle miles traveled.
- Sensitivity to the demographics of the traveling population and how they are changing over time.
- Sensitivity to changes in transportation level of service, due to both changes in regional land use and development over time (including changes in congestion levels) and changes in transportation service resulting from transportation policy changes and transportation investments.

The structure of the 2019 NYBPM is described in Chapter 2.0.

1.2 Products and Additional References

The main products of the 2019 NYBPM project include the following:

- The operational, validated model.
- The executable software to run the model, including source code, to run the activity-based demand modeling components.
- A model interface that uses NYMTC's chosen modeling software, TransCAD, to execute the model and provide key reports of model results.
- Scripts to run various model components, such as network skimming and highway and transit assignment, using TransCAD's GISDK programming platform.
- Socioeconomic data files at the transportation analysis zone (TAZ) level for the base year of 2019 and selected forecast years.
- Synthetic populations representing the model region's population for the base year of 2019 and selected forecast years.

- An integrated transportation network representing the major roadways in the region's highway system and the transit service operated in the region by public agencies and private operators.

- Compiled data for model validation, including traffic counts on screenlines and other key roadways and various measures of transit ridership.

- Data files for model validation, using information from the network and from data sources including NYMTC's Regional Household Travel Survey (RHTS), Regional Establishment Survey (RES), National Household Travel Survey (NHTS), the Public Use Microdata Sample (PUMS) from the U.S. Census Bureau's American Community Survey (ACS) and the LOCUS location-based services (LBS)data product.

- A variety of detailed reports documenting specific work items, including:

    - Model design report (Cambridge Systematics, Inc. et al., 2017)
    - Model validation plan (Cambridge Systematics, Inc. and EA Harper Consulting, 2021a)
    - Model validation report (Cambridge Systematics, Inc. et al., 2023)
    - Model users' guide (Cambridge Systematics, Inc., 2023)

- Various data files with model results, including TransCAD loaded networks with assignment results and a database containing activity-based demand model results.

A complete listing of references is provided in Chapter 7.0.

2.0 Model Structure

The 2012 NYBPM design is documented by Cambridge Systematics, Inc. et al (2017). A few modifications were made for the 2019 update; these are described later in this chapter. The structure of the 2019 NYBPM is illustrated in Figure 2-1.

2.1 Model Summary

2.1.1 Activity-Based Demand Components

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.

PopGen is an open source synthetic population generator developed by Arizona State University. PopGen Version 2.0 was used to generate synthetic populations for the NYMTC model region. The synthetic populations serve as input to subsequent microsimulation model components embedded within CEMSELTS and CEMDAP. PopGen is written in the open source programming language Python and can be seamlessly integrated with the activity-based microsimulation model system. The operational controls for the PopGen software have been integrated within the TourCast platform.

CEMSELTS (Comprehensive Econometric Microsimulator of Socioeconomics, Land use and Transportation Systems) is the component used to produce additional socioeconomic and demographic attributes for each person in the synthetic population with a view to develop a rich set of input data for the activity-based microsimulation model system. All of the variables that can be simulated by CEMSELTS are stripped away from the synthetic population generated by PopGen and replaced with simulated values from CEMSELTS. The resulting richer set of inputs is then fed to CEMDAP, to simulate complete daily activity-travel patterns for the population of the model region. The CEMSELTS components are shown in Table 2-1.

CEMDAP (Comprehensive Econometric Microsimulator for Daily Activity-travel Patterns) is a microsimulation implementation of a continuous-time activity-travel modeling system. It takes as input the disaggregate agent level socio-demographics, land use patterns, and transportation system level-of-service characteristics, and model parameters for the model region, to provide as outputs the detailed individual level daily activity-travel patterns for all the individuals in the study area. "Agents" in this case refer to the individuals who live in the model region, who are performing the activities that result in the travel being modeled. The CEMDAP components are further subdivided into segments based on travel type:

- Generation-allocation (GA)
- Worker (WSCH)
- Non-worker (NWSCH)
- Child (CSCH)
- Joint (JASCH)

Descriptions of the CEMDAP components in each of these five segments are provided in Table 2-2 through Table 2-6.

Figure 2-1. 2019 NYBPM Model Structure

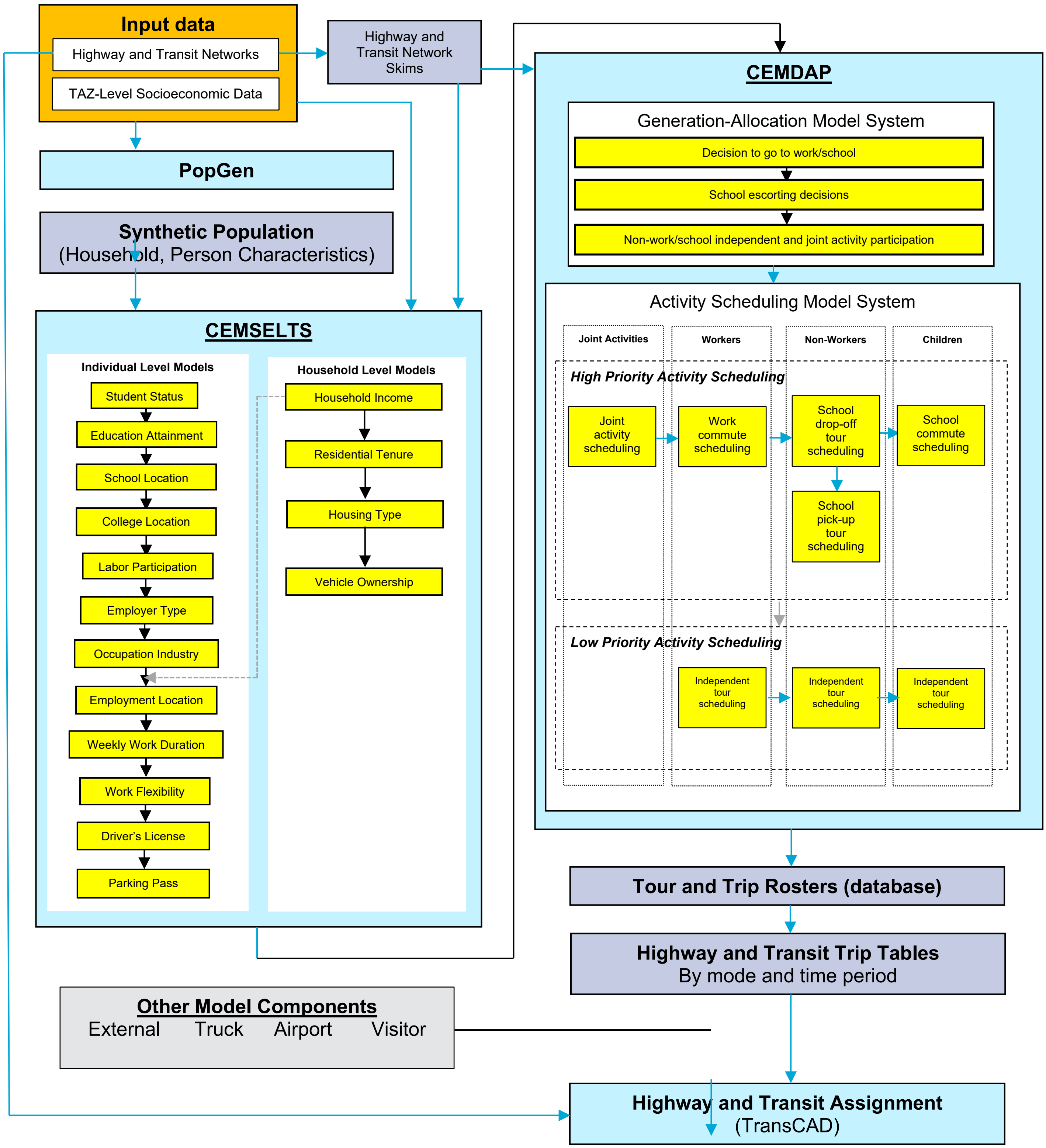


Table 2-1. CEMSELTS Components

Component	Description	Model Unit	Model Type	Data Source
Student status[1]	Student status - Grade level/college status for each person based on age	Person	Lookup tables	RHTS/PUMS
Education attainment	Less than high school/high school/some college/college graduate/any grad school	Person	MNL (5 alts)	RHTS/PUMS
School location	School location - TAZ for each K-12 student	Student	MNL (TAZ alts)	RHTS
College location	College location - TAZ for each college student	Student	MNL (TAZ alts)	RHTS
Labor force participation	Labor force participation - binary choice	Person	Binary logit	RHTS
Employer type	Employer type	Worker	MNL (5 alts)	RHTS
Occupation industry	Occupation industry	Worker	MNL (6 alts)	RHTS
Household income	Household income level	Household	ORL (8 alts)	RHTS
Residential tenure	Residential tenure - own/rent	Household	Binary logit	PUMS
Housing type	Housing unit type	Household	MNL (3/4 alts)	RHTS/PUMS
Employment location	Work location - Regular workplace TAZ for each worker	Worker	MNL (TAZ alts)	RHTS
Weekly work duration	Work duration - <35 hours, 35-45 hours, or >45 hours per week	Worker	MNL (3 alts)	RHTS
Work flexibility	Work flexibility - none, low, medium, and high	Worker	ORL (4 alts)	RHTS
Driver's license	Person holding of driver's license	Person	Binary logit	RHTS
Parking pass	Worker holding of parking pass	Worker	Binary logit	RHTS
Vehicle ownership[2]	Number of vehicles owned by the household	Household	MNL (5 alts)	RHTS

Model structure abbreviations: MNL – multinomial logit, ORL - ordered response logit, MDCEV – multiple discrete-continuous extreme value.

Data source abbreviations: RHTS – NYMTC Regional Household Travel Survey, PUMS – Public Use Microdata Sample from the U.S. Census Bureau's American Community Survey (ACS), NHTS – National Household Travel Survey.

Notes:

1. Lookup table obtained directly from RHTS/PUMS – no validation required
2. New component added after model design plan was completed

Table 2-2. CEMDAP Components – GA Series

Code	Component	What's Modeled	Unit	Model Type
GA1	Child's decision to go to school	Yes/no	Tour	Binary logit
GA2	Child's school start time	Continuous	Person	Hazard-duration
GA3	Child's school end time	Continuous	Person	Hazard-duration
GA4	Adult's decision to go to work	Yes/no	Person	Binary logit
GA5	Adult's work start and end times	32 periods	Tour	Multinomial logit
GA6	Adult's decision to go to school	Yes/no	Person	Binary logit
GA7	Adult's school start time	Continuous	Person	Log-linear regression
GA8	Adult's school end time	Continuous	Person	Log-linear regression
GA9	Child's travel mode to school	Modes	Trip	Multinomial logit
GA10	Child's travel mode from school	Modes	Trip	Multinomial logit
GA11	Allocation of drop off episode to parent	Mother/father	Household	Binary logit
GA12	Allocation of pick up episode to parent	Mother/father	Household	Binary logit
GA13	Determination of households with non-zero out-of-home duration	Out-of-home activities: yes/no	Household	Binary logit
GA14	Determination of total OH time of a household	% time in-home/% out-of-home/% travel	Household	Fractional split
GA15/ GA/16	Independent and joint activity participation	Activity purpose/ # of participants	Household	MDCEV
GA17	Decision of adult to undertake other serve-passenger activities	Yes/no	Person	Binary logit

Table 2-3. CEMDAP Components – WSCH Series

Code	Component	What's Modeled	Unit	Model Type
WSCH1	Worker commute mode	Modes	Tour	Nested logit
WSCH2	Number of before-work tours	0, 1, or 2+ tours	Person	Multinomial logit
WSCH3	Number of work-based tours	0, 1, or 2+ tours	Person	Multinomial logit
WSCH4	Number of after-work tours	0, 1, or 2+ tours	Person	Multinomial logit
WSCH5	Before-work tour mode	Modes	Tour	Multinomial logit
WSCH6	Work-based tour mode	Modes	Tour	Multinomial logit
WSCH7	After-work tour mode	Modes	Tour	Multinomial logit
WSCH8a	Worker number of stops on commute tour	0, 1, or 2 stops	Tour	Ordered probit
WSCH8b	Worker number of stops on before work/after work/at-work tour	1, 2, 3, 4, or 5 stops	Tour	Ordered probit
WSCH9	Worker home or work stay duration before tour	Minutes	Tour	Log-linear regression
WSCH10	Worker activity type at stop	Activity purpose	Trip	Multinomial logit
WSCH11	Worker activity duration at stop	Minutes	Trip	Log-linear regression
WSCH12	Worker travel distance to a stop	Miles	Trip	Log-linear regression
WSCH13	Worker location of a stop	Restricted set of 50 TAZs	Trip	Multinomial logit
WSCH14	Worker Commute Trip Mode Choice	Modes	Trip	Multinomial logit

Table 2-4. CEMDAP Components – NWSCH Series

Code	Component	What's Modeled	Unit	Model Type
NWSCH1	Non-worker number of independent tours	1, 2, 3, or 4 tours	Person	Ordered probit
NWSCH2	Non-worker decision to undertake independent tour before pick-up/joint discretionary tour	Performs tour: yes/no	Tour	Binary logit
NWSCH3	Non-worker decision to undertake an independent tour after pick-up/joint discretionary tour	Performs tour: yes/no	Tour	Binary logit
NWSCH5	Non-worker number of stops in a tour	1, 2, 3, 4, 5, or 6 stops	Tour	Ordered probit
NWSCH6	Non-worker number of stops following pick-up/drop-off	0, 1, 2, or 3 stops	Tour	Ordered probit
NWSCH7	Non-worker home stay duration before tour	Minutes	Tour	Multinomial Logit
NWSCH8	Non-worker activity type at stop	Activity purpose	Trip	Multinomial logit
NWSCH9	Non-worker activity duration at stop	Minutes	Trip	Log-linear regression
NWSCH10	Non-worker travel distance to a stop	Miles	Trip	Log-linear regression
NWSCH11	Non-worker stop location	Restricted set of 50 TAZs	Trip	Multinomial logit
NWSCH4	Non-worker trip mode	Modes	Trip	Nested logit

Table 2-5. CEMDAP Components – JASCH Series

Code	Component	What's Modeled	Unit	Model Type
JASCH2	Joint activity start time	Minutes from 3:00 a.m.	Trip	Log-linear regression
JASCH3	Joint activity distance to stop	Miles	Trip	Log-linear regression
JASCH4	Joint Activity location	Restricted set of 50 TAZs	Trip	Multinomial logit
JASCH6	Joint discretionary trip mode choice	Modes	Trip	Nested logit

Table 2-6. CEMDAP Components – CSCH Series

Code	Component	What's Modeled	Unit	Model Type
CSCH4	Child departure time from home for independent discretionary tour	Minutes from 3:00 a.m.	Trip	Log-linear regression
CSCH5	Child activity duration at independent discretionary stop	Minutes	Trip	Log-linear regression
CSCH6	Child travel distance to independent discretionary stop	Miles	Trip	Log-linear regression
CSCH7	Child location of independent discretionary stop	Restricted set of 50 TAZs	Trip	Multinomial logit
CSCH3	Child mode for independent discretionary trip	Modes (see list)	Trip	Nested logit

The following set of modal alternatives is used in mode choice:

- Auto SOV
- Auto HOV (HOV2)– 2 occupants
- Auto HOV (HOV3+) – 3+ occupants
- Taxi

- Commuter rail/bus – auto access (includes commuter rail, zone based ferries such as NY Waterways, and commuter buses)
- Commuter rail/bus – walk access (includes commuter rail, zone based ferries, and commuter buses)
- Other rail – auto access (includes subway/el, PATH, LRT, and flat fare ferries such as the Staten Island ferry)
- Other rail – walk access (includes subway/el, PATH, LRT, and flat fare ferries)
- Local bus – auto access
- Local bus – walk access
- Walk
- Bicycle

Not all modes are available in every mode choice model. For example, auto SOV is unavailable for the child and joint mode choice models. Alternatives that were never or rarely chosen in the observed survey data sets were excluded from some models; for example, transit auto access modes are not available for work-based subtour mode choice.

It is important to note that differences among the types of transit services within each defined mode are considered in transit path building. As discussed in Section 0, a multipath transit assignment process is used, and therefore, for each zonal origin-destination pair, multiple paths are chosen during assignment.

2.1.2 Non-Activity-Based Model Components

The other demand components of the overall NYBPM consist of the non-activity based components shown at the lower left of Figure 2-1 as “Other Model Components.” These include the external travel, special generators, commercial vehicle travel, and visitor model components. The first three of these components were taken directly from the 2010 NYBPM, and their structures were not updated as part of the 2019 model development (although the 2019 model inputs (e.g., socioeconomic data) were used in running them as part of the 2019 NYBPM). The visitor model was a new component developed for the 2012 NYBPM, and while its structure did not change for the NYBPM 2019, the inputs—such as the visitor trip rates and destination choice coefficients—were updated to reflect 2019 conditions..

Air Passenger Models

Since air passengers traveling to and from airports use the same transportation infrastructure that other travelers use, their ground travel must be considered as part of the demand on infrastructure and are introduced into the NYBPM as “special generator” trips by mode. Air passengers who use highway modes are added to the TransCAD trip tables for highway assignment; however, air passengers using transit to access the airport are assigned as a separate mode, making it possible to account for differing values of time and different route and mode choices. The methodology in place comes from the 2010 model and was not updated for the 2019 model.

External Travel

External travel in the 2010 NYBPM was derived from cordon traffic counts. A gravity model was used to distribute trips within the model region, and growth factors were used to project external demand for future scenarios. These trips were added as separate cores to the trip tables by time period. In more recent updates, an effort was made to integrate the external model into the core modeling process, by accounting for jobs held by workers who resided outside the region and vice-versa. The Census Transportation Planning Products (CTPP) home to work survey and American Community Survey data were used to develop a seed matrix for County to County auto and total trips, which were adjusted using a Fratar process to produce future year matrices. Long distance non-work trips were derived from the NHTS. This methodology was maintained as part of this model update, using updated traffic counts to reflect 2019 conditions.

Truck/Commercial Vehicles

The NYBPM truck model underwent significant improvements during the 2010 update. The current methodology includes an explicit representation of external traffic at true origins and destinations rather than locking in external trips at external stations. FAF3 data were used for long distance trips and an enhanced Quick Response Freight Manual (QRFM) method was used to model shorter trips. A gravity model was used to distribute the trips. The model uses a form of generalized cost impedance, different from the generalized cost used in highway assignment. The structure of the truck/commercial vehicle component was not updated for the 2019 NYBPM.

Visitor Model

A new visitor model was developed as part of the 2012 NYBPM. The visitor model estimates the average weekday travel made within the model region by people who do not live in the region. This travel component is not covered by the RHTS, and since the travelers' residences are outside the region, the concept of home based tours used in the activity-based model is not relevant.

The visitor model is a person trip based component applied separately in TransCAD and is incorporated into the 2019 NYBPM interface. Transit person and auto vehicle trip tables are the outputs of the visitor model, and these tables are combined with the outputs of other model components for transit and highway assignment.

The data source for estimating the visitor model was the hotel component of the Regional Establishment Survey (RES). This is an intercept survey of visitors staying at hotels in the region that obtained information about all trips made by hotel guests within the past day, including trip purpose, time of day, origin/destination, and mode. The estimated models are applied to reflect the entire universe of visitors to the region, segmented by traveler type (e.g., business vs. leisure). While the inputs to the model were updated, the model was not re-estimated for NYBPM 2019, and LOCUS data for 2019 was the data source used to validate the model.

The visitor model consists of the following components:

- **Trip generation** – The number of trips generated per visitor is estimated. Trips are segmented by purpose (e.g., business, meal, other, non-hotel based). Trip production rates reflect average trip rates by purpose per traveler in each segment.
- **Time of day** – The generated daily trips are split into trips by time period using percentages derived from the LOCUS data . Factors vary by area type.
- **Destination choice** – Multinomial logit destination choice models were estimated for each traveler segment and trip purpose. The models estimate the probability of choosing each destination zone for each origin zone. Similar to (but simpler than) the destination choice models used for the

resident activity-based models, the utility variables include measures of impedance (e.g., mode choice logsum and distance) and size variables that include zone level measures of activity (e.g., employment by type).

- **Mode choice** – Mode choice is performed using multinomial logit models. Similar to (but simpler than) the mode choice models used in the resident activity-based models, these models estimate for each trip the probabilities of choosing the auto, rail, taxi, and walk modes. The models are segmented by traveler type and area type.

2.1.3 Trip Assignment

While the outputs of the activity-based model components are individual trip rosters, the trip rosters are combined into trip tables for use in aggregate highway and transit assignment processes. Walk and bicycle trips are not assigned in the 2019 NYBPM.

Highway Assignment

The NYBPM uses the TransCAD General User Equilibrium multi-modal multi-class highway assignment procedure. Highway assignment is performed for four time periods:

- A.M. Peak (6:00 -10:00 AM)
- Mid-Day (10:00 AM – 3:00 PM)
- P.M. Peak (3:00 – 7:00 PM)
- Night (7:00 PM – 6:00 AM)

Multiple internal iterations of equilibrium highway assignment are performed, within global iterations (feedback loops) of the entire model, with an averaging of trip tables in the intermediate iterations to promote convergence. The final iteration, which generates the scenario forecasts, is uses trip tables resulting from successive averages of previous iterations.

The NYBPM loads seven highway mode vehicle trip tables:

- Single Occupant Vehicle (SOV)
- High Occupant Vehicle – 2 person (HOV2)
- High Occupant Vehicle – 3 or more person (HOV3+)
- Taxi/TNC
- Heavy Truck
- Medium Truck
- Other Commercial

The equilibrium highway assignment procedure is applied in an iterative fashion, where travel times are updated after each iteration to reflect congestion occurring on the network. These updates to travel time are based on a volume-delay function for each link class. The volume-delay functions are modified versions of BPR functions with parameters varying by facility types. The free-flow times are based initially on the network data provided for each link and then updated in each iteration to represent the travel time resulting from the assigned traffic volumes from the last iteration. The method currently used is the minimization of generalized cost.

Turn penalties are included in the highway assignment to prohibit certain turn movements or penalize movements. These are included in the model by identifying specific turn movements by their node numbers, and then coding the penalty function that will apply to these turn movements.

Transit Assignment

Transit assignment is performed for the a.m. peak period. The mode choice component of the 2019 NYBPM has the following transit submodes:

- Commuter rail/bus – auto access
- Commuter rail/bus – walk access
- Other rail (except commuter) – auto access
- Other rail (except commuter) – walk access
- Local bus – auto access
- Local bus – walk access

The Staten Island Ferry is treated as part of the subway mode, and other ferry services are treated as part of the commuter rail mode. This is because the Staten Island Ferry—like subways—has a flat fare structure. It costs nothing to ride and is often used as part of multimodal trips that use other modes such as subway. Other ferry services have fares that are zone-based, similar to the way that commuter rail fares are collected.

The NYBPM uses the Pathfinder algorithm, which reflects crowding and transit line capacities. The multi-class capability of Pathfinder accounts for the fact that bus and subway lines are used by commuter rail riders in addition to other riders, and thus limited line capacities reflect both classes of riders. The assignment procedure is iterative and considers the volume/capacity ratio in the final assignment solution, with the objective of minimizing a travelers' generalized cost. Both route level capacities and link level capacities are used.

2.2 Incorporation of TNCs

The previous version of the NYBPM had a base year of 2012, before the widespread use of Transportation Network Companies (TNCs). TNCs include Uber, Lyft, Sidecar, and other ride-hailing companies that provide app-based, on-demand ride services. TNC drivers use their personal vehicles for commercial activities to transport passengers from one place to a set destination. TNC travel for forecast years was assumed to be part of the taxi mode.

By 2019, TNC travel constituted a significant portion of travel in the region, especially in New York City. NYMTC decided to add capabilities to the model to consider demand for and use of TNCs. Key objectives included:

- Accurately estimating levels of TNC demand that considers geographic location, travel purpose, trip characteristics, resident/visitor distinctions, and traveler characteristics;
- Optimal use of available data from a variety of sources;
- Keeping the process as practical and simple as possible within the NYBPM framework, maintaining user friendliness and minimizing model run time; and
- Considering that TNC demand could continue to grow beyond the new base year and ensuring that forecasts for future years consider this growth.

Available data sources for this effort included:

- Data on TNC and taxi demand maintained by the New York City Taxi & Limousine Commission (TLC), including publicly available summaries of TLC data (for example, https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/)
- NYMTC Regional Establishment Survey (though taxi and TNC were combined into a single mode response)
- 2017 NHTS
- LOCUS data
- Data on TNC use from other regions and from academic research

Two options for explicitly considering TNCs in the 2019 model were considered:

- Continuing to model taxi and TNC as a single mode and having a separate process to separate the two modes after mode choice
- Adding a TNC mode to the model in a nest with taxis

NYMTC decided that the first option better met the objectives. It was also decided that the use of TNCs for transit access and egress would be considered as part of auto access to transit.

2.2.1 TNC Mode Choice

Table 2-7 summarizes the taxi/TNC mode share data for resident trips from the 2017 NHTS and NYMTC's RHTS. The RHTS was conducted before the widespread use of TNCs, and no data was collected on TNC travel. Table 2-8 summarizes mode shares for visitors to the region from NYMTC's RES. The RES combined the taxi and TNC modes in its questionnaire.

Table 2-7. Resident Travel Taxi/TNC Mode Shares from Surveys

Tour type	TNC Shares from NHTS	Taxi Shares from RHTS		
		Manhattan	Other NYC	Other
Work	0.1%	4.8%	0.9%	0.7%
Non-worker	0.1% (All non-mandatory)	4.3%	2.0%	0.6%
Before work		0.0%	0.0%	0.0%
After work		5.1%	0.5%	0.0%
Child		0.3%	0.2%	0.0%
Joint	0.3%	9.1%	1.1%	0.2%
Work-based subtour	0.3%	0.3%	0.0%	0.0%

Table 2-8. Visitor Travel Taxi/TNC Mode Shares from RES

Visitor Trip Purpose	Taxi/TNC Mode Share
Hotel based business	18%
Hotel based meal	16%
Hotel based other	12%
Non-hotel based meal	13%
Non-hotel based other	6%

Table 2-9 presents data on for-hire vehicles (FHV) in New York City for 2012 (the NYBPM's previous base year), 2016 (the first year for which TNC demand data is available), 2017 (the year of the most recent NHTS), and 2019 (the base year for the updated NYBPM). The information includes publicly available data summarized by analyst Todd Schneider[1] and data from a 2019 report prepared for the New York City Taxi and Limousine Commission[2]. Data on national TNC demand growth over the past five years is also summarized by Dean[3].

Table 2-9. Taxi/TNC Trips, 2012-2019

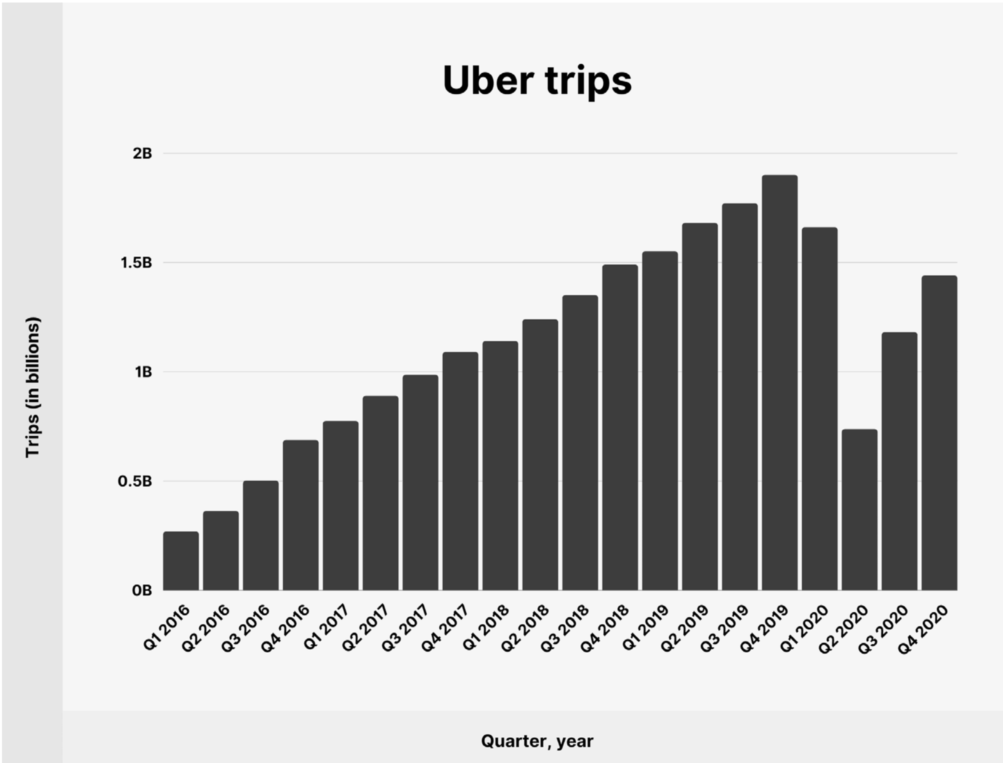
Measure (Source)	2012	2016	2017	2019
Taxi trips/day - NYC (Schneider)	515,848	397,780	334,865	247,742
TNC trips/day - NYC (Schneider)		217,147	393,918	724,388
Taxi+TNC trips/day - NYC (Schneider)	515,848	614,927	728,783	972,130
Total Uber trips (Dean)		350,000	900,000	1,650,000
Taxi trips - NYC (TLC)		13,300,000	11,000,000	8,500,000
TNC trips - NYC (TLC)		6,700,000	12,200,000	24,000,000
Taxi+TNC trips - NYC (TLC)	17,000,000	20,000,000	23,200,000	32,500,000
Taxi trips - Manhattan (TLC)		10,700,000	8,900,000	6,900,000
TNC trips - Manhattan (TLC)		3,800,000	5,400,000	9,700,000
Taxi+TNC trips - Manhattan (TLC)	11,700,000	14,500,000	14,300,000	16,600,000
Taxi trips - Other NYC (TLC)		2,600,000	2,100,000	1,600,000
TNC trips - Other NYC (TLC)		2,900,000	6,800,000	14,300,000
Taxi+TNC trips - Other NYC (TLC)	5,300,000	5,500,000	8,900,000	15,900,000

[1] https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/

[2] New York City Taxi and Limousine Commission and Department of Transportation. "Improving Efficiency and Managing Growth in New York's For-Hire Vehicle Sector." Final Report, June 2019.

[3] Dean, "Uber Statistics 2021: How Many People Ride with Uber?" Backlinko, https://backlinko.com/uber-users, Retrieved 11/18/21.

Figure 2-2. Growth in Uber Trips Nationally



Source: Dean (2021).

2019 mode shares for the combined taxi/TNC mode were estimated for three geographic regions: Manhattan, the rest of New York City, and the rest of the NYBPM model region. For the latter, the estimated mode shares for both 2012 (taxi) and 2019 (taxi plus TNC) are less than one percent, and so it was decided not to perform any mode choice revalidation outside New York City. Within New York City, estimated mode shares for both 2012 and 2019 are less than one percent for the following tour types: worker before work, child school, child non-school, and work-based subtours. It was therefore decided not to perform any mode choice revalidation for these tour types for the 2019 model.

Therefore, mode choice revalidation within New York City was limited to work commute, worker after work, non-worker, and joint tours. Table 1 shows the estimated 2012 and 2019 taxi/TNC mode shares for these tour types, for Manhattan and other NYC. Generally, mode shares in Manhattan increased by about 40 percent while mode shares elsewhere in the City tripled.

Table 2-10. Taxi/TNC Trips, 2012-2019

	2012 Taxi Share		2019 Taxi/TNC Estimate	
	Manhattan	Other NYC	Manhattan	Other NYC
Work	4.8%	0.9%	6.8%	2.7%
After work	5.1%	0.5%	7.2%	1.5%
Non-worker	4.3%	2.0%	6.1%	6.0%
Joint	9.1%	1.1%	12.9%	3.3%

For visitor trips, the original taxi/TNC mode shares used in model validation came from the RES. It makes sense to apply a growth rate to 2019. The best available source for a growth factor representing the combined taxi/TNC mode is the information compiled by Schneider for FHV. Applying this factor to the mode shares shown in results in the shares shown in Table 2-11.

Table 2-11. Estimated 2019 Visitor Travel Taxi/TNC Mode Shares

Visitor Trip Purpose	Taxi/TNC Mode Share
Hotel based business	28%
Hotel based meal	25%
Hotel based other	19%
Non-hotel based meal	21%
Non-hotel based other	9%

2.2.2 Inputs to Mode Choice

Since the mode choice models are not being restructured, the main question regarding how to incorporate TNCs into the NYBPM as a combined mode with taxi is how to represent the service characteristics. Generally, TNC and taxi in-vehicle travel times are similar, and so no changes were made to this variable, whose values are obtained from the highway network skims for the HOV modes.

Wait times can vary between TNCs and taxis though since 2012 many taxi services can also be summoned using apps, similar to TNCs. Of course, many cabs are hailed on-street, especially in Manhattan, which is not an option for TNCs. There is no available information to accurately estimate wait times and to distinguish them between taxis and TNCs. It was decided not to revise how wait time inputs are estimated for 2019 from the network.

3.0 Data Development

3.1 Zone Systems

Like all travel models, the NYBPM uses a system of transportation analysis zones (TAZ) to represent the locations of homes and workplaces, activities, and trip ends. The model region is divided into 5,418 internal TAZs and 111 TAZs representing external stations, where network roadways crossed the regional boundary. A summary of the TAZ numbering by geographic location is provided in Table 3-1. The TAZ system was unchanged from the 2012 NYBPM structure, which was documented in the 2012 model summary report (Cambridge Systematics, Inc. et al., 2021).

Table 3-1. TAZ Numbering

State	County #	County FIPS	County Name	District #	District	Sub-Region	TAZs Range (from)	TAZs Range (to)	# of TAZs
NY	1	36061	New York	1	CBD: Lower	CBD	1	14	14
				2	CBD: Valley		15	107	93
				3	CBD: Midtown		108	165	58
				4	Other Manhattan	Upper Manhattan	166	335	170
	2	36081	Queens	5	Queens	Other NYC	336	1004	669
	3	36005	Bronx	6	Bronx		1005	1343	339
	4	36047	Kings	7	Kings		1344	2103	760
	5	36085	Richmond	8	Richmond		2104	2212	109
	6	36059	Nassau	9	Nassau	Long Island	2213	2491	279
	7	36103	Suffolk	10	Suffolk		2492	2813	322
	8	36119	Westchester	11	Westchester	Mid-Hudson	2814	3036	223
	9	36087	Rockland	12	Rockland		3037	3101	65
	10	36079	Putnam	13	Putnam		3102	3120	19
	11	36071	Orange	14	Orange		3121	3200	80
	12	36027	Dutchess	15	Dutchess		3201	3279	79
NJ	13	34003	Bergen	16	Bergen	New Jersey NJTPA Core Area	3280	3467	188
	14	34031	Passaic	17	Passaic		3468	3568	101
	15	34017	Hudson	18	Hudson		3569	3754	186
	16	34013	Essex	19	Essex		3755	3983	229
	17	34039	Union	20	Union		3984	4098	115
	18	34027	Morris	21	Morris	New Jersey NJTPA Other	4099	4199	101
	19	34035	Somerset	22	Somerset		4200	4280	81
	20	34023	Middlesex	23	Middlesex		4281	4489	209
	21	34025	Monmouth	24	Monmouth		4490	4642	153
	22	34029	Ocean	25	Ocean		4643	4778	136
	23	34019	Hunterdon	26	Hunterdon		4779	4810	32
	24	34041	Warren	27	Warren		4811	4837	27
	25	34037	Sussex	28	Sussex		4838	4881	44
	26	34021	Mercer	29	Mercer	New Jersey (DVRPC)	4882	5005	124
CT	27	09001	Fairfield	30	Fairfield	Connecticut	5006	5215	210
	28	09009	New Haven	31	New Haven		5216	5404	189
Special Generator Zones							5405	5418	14
Total TAZs									5418

Note: Unused zone numbers – 5419 - 6000. External Zone numbers – 6001 - 6111

3.2 Socioeconomic Data

Socioeconomic data are used in numerous ways in the 2019 NYBPM, including the following:

- In the development of the synthetic populations using PopGen, as control totals
- To compute size variables for location choice models, to represent the amount of activity at potential destinations
- In the computation of accessibility variables used in some model components

NYMTC produced totals for a variety of socioeconomic data items at the TAZ level, for the base year of 2019 and for five-year increments from 2020 to 2055. Variables include the following:

- Persons/households
  - Total population
  - Total households
  - Population in households
  - Group quarters population
  - Persons by age group and gender
  - Persons by employment status
  - Labor force
- Employment by type
  - Office
  - Retail
  - School
  - Total

3.3 Networks

The CS team updated the 2012 NYBPM highway and transit networks to the 2019 base year and implemented various corrections and improvements. The network development process is documented in a separate report (Cambridge Systematics, Inc. and EA Harper Consulting, 2023), hereafter referred to as the network update report.

The development of the 2012 NYBPM included the integration of the highway and transit networks. The exercise of integrating the highway and transit networks was made possible by the availability of the General Transit Feed Specification (GTFS) files for the major transit operators. Highly accurate shapes (or lines) describing the bus network were conflated to the highway link layer and used to associate bus routes with highway links. Bus frequencies and speeds were also imported and processed to represent time period headways.

The integrated network link layer consists of three types of links:

- Highway network links representing the regular roadway segments, used for autos, trucks, taxis, and buses;

- Fixed guideway transit links exclusively used by transit modes like subway, rail, ferry, and trams; and
- Transit station connection links connecting the fixed guideway stations and road network links for station/stop access and egress, or for transferring between transit stations/stops.

The remainder of this section describes the work done to update the NYBPM 2012 highway and transit networks to a 2019 base year and implement other corrections and improvements to the 2012 network. More detail can be found in the network update report.

3.3.1 Highway Network

The network structure and attributes are retained from the 2012 base year model. A full data dictionary of all the integrated link layer fields can be found in the network update report.

The changes necessary to bring the NYBPM 2012 highway network up to a 2019 base year were provided by NYMTC, their consultants, and member agencies. In addition to the basic 2012-to-2019 changes, network coding of tolling, cashless tolling, and truck policies were updated.

Tolls

The NYBPM area includes five tolling agencies: the Port Authority (PA), Metropolitan Transit Authority (MTA), New York State Thruway Authority, New York State Bridge Authority (NYSBA), and the New Jersey Turnpike Authority. Section 2.1.1 of the network update report outlines the process used to develop the tolls for the NYBPM 2019 update, describes how the tolls are incorporated into the model, and describes how the process is different from what was used in the 2012 NYBPM.

One of the main changes to the 2012 NYBPM process was the direct incorporation of cashless tolling into the highway network. The 2012 NYBPM was developed with a base year that predated many of the toll facilities in the region being converted to “open road” tolling, where drivers with E-ZPass transponders do not have to slow down to pass through toll booths. With the update of the base year to 2019, open road tolling is more prevalent. Therefore, the coding of these toll facilities in the highway network was revised to reflect the increased speeds and capacities associated with open road tolling.

The approach used in the 2012 NYBPM used the code “TP” in the SPECIAL field in the highway network to represent toll plazas. Links were coded with as many lanes as there are toll gates, and toll plaza links were coded with lower capacities than the links before and after the toll plazas.

This approach was problematic, beyond the lack of representation of open road tolling. The volume-delay function used for highway assignment is not suitable for toll plazas, and in some cases, extremely high travel times were produced for toll plazas. In the 2012 NYBPM, it was necessary during highway assignment validation to make significant calibration adjustments to some toll plaza links to produce reasonable assigned volumes. Even with these changes, there were some toll facilities that were noticeably overassigned or underassigned.

A new approach used for the 2019 NYBPM was implemented. First, the CS team inventoried all toll facilities in the highway network (refer to Open Road Status.docx) and defined toll plazas by the type of toll collection taking place. The value in the SPECIAL field in the network was recoded with one of the following four values:

- **TP-C: Toll Plaza with cash option**. These toll plazas typically require vehicles with a transponder to slow to around 15-25 mph. A cash payment option is available.

- **TP-O: Open-road toll plaza**. These toll plazas have been modified and no longer have a cash payment option. Vehicles do not need to slow for toll collection.

- **TP-B**. Both cash and open road options are available.

- **TR: Toll road calculation point**. These links are present in the network to account for toll costs on closed systems. Links marked TR do not have any physical toll collection infrastructure. Tolls are collected via toll plazas (marked as TP-C, TP-O, or TP-B) when vehicles enter and exit the facility. From a modeling perspective, TR links function similarly to TP-O links. Vehicles do not need to slow down on these links to pay a toll. They are categorized separately to allow different treatment if necessary during future model calibration efforts.

The link travel times on open-road tolling links were modified to have lower delays. The coding convention with one lane per toll booth was eliminated.

The model scripts for highway assignment were modified to use these updated toll categories. Changes to the script with respect to open road tolling can be identified by searching the script for the term SPECIAL and the associated categories defined above.

Truck Routes and Restrictions

The network update work included a review and update of truck routes that were coded into the 2012 NYBPM highway network to reflect 2019 conditions. Section 2.1.2 of the network update report discusses details on data sources used, assumptions and approach adopted, to update the NYBPM 2019 truck routes.

Quality Checks and Reviews

NYMTC and their member agencies performed additional checks and provided a series of useful comments and corrections; these are provided in Appendix C of the network update report. As a result of these reviews, additional changes were made, including changes in numbers of lanes at specific locations and corrections to directionality on some links. All one-way links in the network use a Dir value of 1, indicating A to B travel, and so there are no links having a Dir value of -1. This was implemented in to make the Dir field clearer and less confusing, and to reduce opportunities for coding errors. Where necessary, link topology was reversed to accommodate B to A travel. Dir code 0 denotes a two way link.

3.3.2 Transit Network

In development of the 2012 NYBPM, the highway and transit link layers were integrated, enabling highly accurate GTFS route locations and schedules to be imported into the model network. This import process was repeated using 2019 GTFS files where available. The integration process aligns the GTFS shapes to the highway link layer (referred to here as conflation), and then converts the GTFS route data into the TransCAD format (referred to here as the route system development). The source for the route system update information can be found the “Source “ field of the route system database. In addition, the field “GTFS Filename” has the date of the GTFS file if GTFS was the data source. The GTFS conflation process is described in Section 3.1 of the network update report.

The transit route system was built on top of the integrated network link layer. There were three types of transit routes to be included: 1) bus routes for which GTFS data existed, 2) bus routes for which no GTFS data existed, and 3) fixed guideway routes. Routes with GTFS data were conflated into a separate route system and then merged with fixed guideway and non-GTFS routes from the 2012 route system. A data dictionary for the NYBPM 2019 route system can be found in Appendix B of the network update report.

Fares

The GTFS files available in the NYMTC region do not include the optional fare data. For the 2019 update the consulting team has performed a complete review and update of all transit fares, assuring that fares coded in the model represent actual transit travel costs as accurately as possible. Where 2019 observed data were available, fares were updated to represent 2019 values. Where no observed 2019 data could be obtained, factors representing 2012 to 2019 inflation rates were developed and applied. All fares are AM peak period fares; no discounts for passes, employee subsidy, ADA, seniors, or students were included to ensure that fares across routes were consistent. Given the large number of transit operators in the region with varying fare and discount policies, use of full fares in the transit system was deemed to be the most consistent and appropriate approach. More detail on transit fares in the network can be found in Section 3.4.1 of the network update report.

Transit Capacities

Bus and subway capacities were computed similarly to the 2012 NYBPM. Bus capacities are coded on the routes as total AM peak period capacity, as discussed in Section 3.4.2 of the network update report. Subway capacities were retained from the 2012 NYBPM.

Transit Access

The transit walk network (including all walk access, transfer, and connector links) were created from a combination of 1) walkable roadway links, and 2) additional connectors extracted from the 2012 NYBPM transit network and manually added and edited, as discussed in Section 3.4.4 of the network update report. Auto access is represented by the roadway and centroid connector links, as discussed in Section 3.4.5 of the network update report.

Travel Times

The GTFS files contain stop-to-stop travel times for each route. Since the GTFS stops do not always align with existing NYBPM roadway network nodes, GTFS travel times had be adjusted to correspond with the appropriate NYBPM roadway segment. The resulting bus travel time for a segment is stored on the TransCAD route stop at the beginning of the segment—the ‘from stop.’

Note that when multiple routes pass through a single intersection, there is a separate route stop for each route. Therefore, a single TransCAD node may be associated with many route stops. The bus travel times were calculated separately for each route passing through a given TransCAD node and stored on the appropriate route stop. In this way, speeds for express, limited stop, and local buses are represented independently.

Mode Definitions

The NYBPM 2012 included three separate mode identifiers. Two of these identifiers, MODE and MODE_1, contained slightly different mode groupings. The third field, MODE_ID, included numerous mode categories.

The MODE_ID categories were related to a fare implementation that is no longer used by the NYBPM. These three fields have been replaced with a single MODE field, defined in Table 3.2. Grouping of modes for use in the mode choice model has been retained from the NYBPM 2012 model system.

Note that there is no clearly defined attribute on the bus routes that distinguishes Commuter Buses from other buses. Both express and commuter buses capture transit riders that are commuting into Manhattan. However, in the NYBPM, Premium buses that run express, but that do not serve the region's core, are not included in this definition of Commuter Bus. As such, for all companies other than MTA, the presence of one primary terminus inside Manhattan designates a bus route as a Commuter Bus. For the MTA there are many routes with one terminus in Manhattan that are local buses. The additional criterion of express service/fare is used to eliminate non-premium services from the MTA Commuter Bus mode.

Also note that "Subway" includes MTA subways, PATH, and LRT such as Hudson Bergen Light Rail and Staten Island Railway (SIR) even though these services have different attributes. For example, some have at grade crossings while others don't. These were grouped together into one mode to make the skimming and assignment processes less onerous.

The TransCAD MODE table previously referenced MODE_ID and has been updated to align directly with the MODE field.

Table 3.2. Route System Mode Definitions

MODE	Description
1	Local Bus
2	Commuter Bus
3	Express Bus
4	Commuter Rail
5	Subway
6	Tram (RIT)
7	Ferry (Zone Fare)
8	Ferry (Flat Fare)
9	Ferry Bus

Transit Route-level Checks and Quality Control

A series of visual checks and tabulations were executed at multiple stages in the development of the route system. Adjustments to the conflation algorithms and route building methods were implemented accordingly. In addition, transit travel time skims were prepared by mode and time of day and reviewed for reasonableness. As shown in the network update report, thematic maps of the travel time skims were produced. The intention of this mapping was to verify that the route system generally operates as expected. Modifications to path-building parameters (e.g., maximum drive distances, weights on impedances, transfer penalties, etc.) were reviewed during model validation.

3.4 Data for Model Validation

Besides the socioeconomic and network data described earlier in this chapter, a variety of other data sources were used in the development of the 2019 NYBPM, for model validation and application. These sources are summarized below.

3.4.1 Survey Data

The following survey data sets were used in the 2019 NYBPM development process.

NYMTC Regional Household Travel Survey

The NYMTC Regional Household Travel Survey (RHTS) was conducted by NYMTC and 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, and 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. For a linked trip, the “From” place represents a trip Origin and the “To” place represents a trip destination. The linked trips may or may not have stops. Households were surveyed from each of the 28 counties within the NYBPM model region within New York, New Jersey, and Connecticut.

Model estimation was originally performed for the 2012 NYBPM. THE RHTS data set was the primary source for the estimation most of the CEMSELTS components (see Table 2-1) and all of the CEMDAP components. It served as the primary source for the validation of the CEMSELTS and CEMDAP components as it provided the most comprehensive data set that represented the travel behavior of the region’s residents. Since there were no more recent household survey data available for the 2019 NYBPM update, models were not reestimated, meaning that model parameters were not derived through a statistical estimation process. However, parameters were adjusted as necessary as part of model calibration.

NYMTC Regional Establishment Survey

The NYMTC Regional Household Travel Survey (RES) was conducted by NYMTC in 2014-2015. Like the RHTS, it covered the entire NYBPM model region. RES data was used as a supplemental source in the 2012 NYBPM estimation of some of the CEMDAP components, and the hotel sample was the primary data source used in the estimation of the visitor model component for the 2012 NYBPM. There was no newer RES data available for the 2019 update, and so the visitors’ model was not reestimated. However, the model parameters were adjusted as needed during model calibration.

Transit Rider Surveys

These surveys—which were the most recent available at the time—were valuable in validating the transit related model components such as mode choice and transit assignment. Transit survey data used for the 2019 NYBPM are summarized in Table 3.3. The survey data are described in more detail in the transit validation database report (Cambridge Systematics, Inc. and EA Harper Consulting, 2021b).

Table 3.3. Transit Surveys

Survey	Organization	Direction	Duration	Year	Growth Rate to 2019 (****)
Trans-Hudson Bus Surveys	NJ Transit	Inbound	All Day	2015	1.1511 (PABT) 0.5724 (GWBBT) 1.0966 (other NYC)
Bergen-Passaic Bus Survey	NJ Transit	I/O	All Day	2009	
Metro North OD Survey	MTA	Inbound	All Day	2017	1.0554
Long Island Railroad OD Survey	MTA	Inbound	All Day	2012	1.2028
Raritan Valley Commuter Rail Survey	NJ Transit	Inbound (*)	5:30AM - 8:00 PM (*)	2014	1.0541
Montclair Boonton Commuter Rail Survey	NJ Transit	Inbound (*)	5:30AM - 8:00 PM (*)	2015	1.0085
Morris and Essex Commuter Rail Survey	NJ Transit	Inbound (*)	5:30AM - 8:00 PM	2016	1.0138
Main Bergen and Pascack Valley Commuter Rail Survey	NJ Transit	Inbound (*)	5:30AM - 8:00 PM (*)	2013	1.1030
NEC and NJCL Commuter Rail Survey	NJ Transit	Inbound (*)	5:30AM - 8:00 PM (*)	2014	1.0541
Secaucus	NJ Transit	Inbound (*)	6:00 AM to 7:15 PM (*)	2013	1.1030
Newark Airport Commuter Rail Survey	NJ Transit	I/O	5:00 AM to midnight (*)	2014	1.0541
NJ Hudson Bergen Light Rail Survey	NJ Transit	I/O	5:00 to midnight (*)	2017	1.0285
Newark Light Rail Survey	NJ Transit	I/O	Primarily AM	2007	1.1850
PATH Survey	PANYNJ	I/O	5:00 AM to midnight (*)	2012	1.1462
NJ-Based Ferry Survey	NJ Transit	I(*)	5:00 AM TO 8:00 PM (*)	2013	1.1332

3.4.2 Traffic and Transit Counts

A variety of available traffic count data was used for highway assignment validation. Screenline count locations were identified where count data were available. Other count data, including classification counts, from New York City, New York State, river crossings, and major New Jersey screenlines from the NJTPA travel demand model, were incorporated into a database for highway assignment validation. The database development is documented in a report on the highway validation database (Cambridge Systematics, Inc. and EA Harper Consulting, 2021c).

A traffic count database was developed to represent the 2019 base year. The ability to validate the model is heavily dependent on the quality and coverage of the validation data. With that in mind, an approach was developed that widens the amount of validation data that can be accessed to populate the database and increases the reliability of the data. The approach relies on access to the universe of available geo-referenced count data. SPSS and ArcGIS were used to compile and develop a database that can be joined to the NYBPM network by the TransCAD Link ID field. A large amount of traffic volume data was obtained from a variety of data sources and systematically assembled. A geo-processing method was developed to automate and improve the accuracy of tagging the count locations (points) with NYBPM-2019 network links. In addition, improved methods for quality control were implemented.

Transit boarding counts were used for transit assignment validation. These were not uniformly available for all transit services due to the wide variety across operators and services of station entry, fare collection, and transferring types. The database development is documented in the transit validation database report (Cambridge Systematics, Inc. and EA Harper Consulting, 2021b). The data sources used are summarized in Table 3.4.

Table 3.4. Transit Count Data Sources

Dataset	Organization	Year
2019 Hub-Bound Report (**)	NYMTC	2019
NYCT 2019 Bus Boardings and Alightings (**)	MTA	2019
NYCT 2019 Station to Station Matrix (**)	MTA	2019
NYCT 2019 Subway Boardings & Alightings (**)	MTA	2019
East River Ferry Quarterly Report (**)	NYEDC	2019
National Transit Database 2019 (**) -local bus NY/CT & SI Railroad	Federal Transit Administration	2019

3.4.3 LOCUS Location-Based Services Data

The 2019 NYBPM validation took advantage of a "big data" source called LOCUS, a product of Cambridge Systematics. For the NYBPM validation, LOCUS data included origin-destination tour and trip tables by time of day derived from location-based services (LBS) data generated by smartphone applications where users have explicitly granted permissions for the application to track geolocations. A brief summary of the LOCUS data prepared for NYMTC is provided by Cambridge Systematics (2022). LOCUS data were used in the validation of the activity-based demand components where the data were relevant, as well as the visitor model.

In model validation, LOCUS has three major advantages over survey data:

- LOCUS data from 2019, the model base year, were available for use; the available survey data were collected earlier, including the RHTS, which was conducted in 2010-2011, and transit surveys conducted mainly in the mid-2010s (see the model validation plan for a complete list of validation data sources).
- Tour-level data are provided by LOCUS; this is also true for the RHTS, but not for other data sources such as transit surveys and the ACS.
- Sample sizes are much larger than for the household surveys, reducing sampling error.
- Since LBS data are passively collected, there is no response bias.

The major shortcoming of LBS data is that demographic information about the travelers is not available. In addition, travel modes in the LBS data are not known, and so the LOCUS data set is not used in mode choice validation. It is also important to recognize that LOCUS data represents mobile devices, not persons. There are population segments (for example, young children) who have lower mobile device penetration rates and who may be underrepresented.

Tour and trip purpose information is also limited in LOCUS since, unlike in surveys, travelers cannot be asked the reasons for their travel. However, limited travel purpose information can be inferred. Specifically, algorithms are used to infer the home locations and regular (work and school) locations. Note that LOCUS does not distinguish between "work" and "school" locations; while there is some available information in LOCUS that is correlated with the distinction between work and school (e.g., tour start/end times and duration at the regular location), other key information is unavailable, most notably traveler age.

Consequently, tours in the LOCUS data set are classified as follows:

- **Regular tours** – These are tours that begin and end at home where one of the stops is at a regular location (work or school). These are identified by the combination of home and regular location. They correspond to the worker commute, adult school, and child school tours in the NYBPM.

- **Non-regular tours** – These are tours that begin and end at home where none of the stops is at a regular location. These are identified by home location only since rules would be required to define a "primary" non-home location for each tour, and this is not done for the non-work/non-school tours in the NYBPM. They correspond to the non-worker, child (non-school), and joint tours in the NYBPM.

- **Non-home based tours** – These are tours that begin and end at a location other than the traveler's inferred home location. These are identified by the start/end location of the tour. They roughly correspond to the work-based tours in the NYBPM, which does not generate tours that begin and end at locations other than home or work.

The LOCUS data set also includes the individual trips that comprise each tour. In total, therefore, there are four LOCUS data sets used in model validation:

- Regular tours – identified by the combination of home and regular locations
- Non-regular tours – identified by the home location
- Non-home based tours – identified by the tour start/end location
- Trips – identified by the combination of origin and destination

The tour data sets are segmented by tour start and end hours. Each record in the trip data set is classified by time period, corresponding to those used in highway assignment:

- AM peak (6:00 AM – 10:00 AM)
- Mid-day (10:00 AM – 3:00 PM)
- PM peak (3:00 PM – 7:00 PM)
- Night (7:00 PM – 6:00 AM)

4.0 Model Estimation

Model estimation is the process of obtaining statistical estimates for the parameters of the mathematical functions that are part of the activity-based demand model components in CEMSELTS and CEMSELTS. It consists of using statistical processes to estimate the most likely values of the parameters of the mathematical formulations used in the model to relate the model outputs to the input data, i.e., the activity-based demand model components in CEMSELTS and CEMSELTS. Different types of these formulations are used, depending on the specific component; these are specified in Table 2-1 through Table 2-6. The estimation procedure performed for the 2012 NYBPM used observed data from several sources, including:

- Regional Household Travel Survey (RHTS)
- Regional Establishment Survey (RES)
- The 2010-2012 three-year Public Use Microdata Sample (PUMS), part of the American Community Survey (ACS)
- The 2009 National Household Travel Survey (NHTS)

The original model estimation process and results are described in detail in the 2012 NYBPM estimation report (Cambridge Systematics, Inc. et al., 2018b).

As previously noted, since there were no more recent household survey data available for the 2019 NYBPM update, models were not reestimated. However, some were restructured with new mathematical formulations. The following models were restructured as part of the 2019 update:

- Independent and joint activity participation (GA15/GA16)
- Adult's work start and end times (GA5)
- Non-worker home stay duration before tour (NWSCH7)
- The various mode choice models, including school (GA9/GA10), worker (WSCH1, WSCH14), non-worker (NWSCH4), joint travel (JASCH6), and child models (CSCH3)

The main issues to be addressed by these changes included:

- Inability to reflect activity type distributions properly;
- Time of day results for some models that did not match observed or reasonable temporal distributions patterns, and were difficult to validate in their original forms; and
- Mode choice results that were off by time of day—specifically, off-peak transit shares that were too high, resulting in too much ridership–based on counts–in the mid-day and, especially, the nighttime period.

The changes to each component were made by:

- Respecifying the component in its new form (restructuring), if necessary;
- Specifying the values for any new model parameters;
- Revising the code to run the revised component;
- Revalidating/recalibrating the components, revising parameters as needed.

The details for the revisions are documented by Cambridge Systematics, Inc. (2023a). A brief summary is provided below.

4.1 Independent and Joint Activity Participation

The original GA15/GA16 models jointly estimated individual and joint activities for all members of a household, with separate models for households with five or fewer members (GA15) and households with more than five members (GA16). The original models were formulated as multiple-discrete continuous extreme value (MDCEV) models.

These original models proved difficult to calibrate, particularly with regard to simulating reasonable numbers of activities for certain purposes such as work-based. The MDCEV models were computationally complex and posed challenges for code debugging, maintenance, and run times. It was therefore decided to replace these models with a new set of multinomial logit models, simulating for each household:

- Joint activity generation
- Joint activity participation
- Individual activity generation
- Household activity duration (time allocation)

The initial model parameters were obtained from the original MDCEV models and reassigned to the four new models based on their type. These parameters were calibrated to produce reasonable comparisons to data from the Regional Household Travel Survey (RHTS), and validation results were reasonable, both for the revised models and downstream models directly affected by the results of these models.

The existing GA15 and GA16 models were replaced by a set of four models:

- **Joint activity generation** – Multinomial logit model that simulates the number of joint activities performed by a household by activity type (eat-out, maintenance, other, recreational, shopping, social)

- **Joint activity participation** – Multinomial logit model that simulates which household members participate in each joint activity simulated in the previous model

- **Independent activity generation** – Multinomial logit model that simulates the number of individual (non-joint) activities performed by each household member by activity type

- **Activity time budget** – Proportional shares model that simulates the percentage of time for each simulated activity.

4.2 Adult's Work Start And End Times

The original model (GA5) was challenging to calibrate to reflect observed diurnal distributions from the RHTS. It included variables with trigonometric functions (e.g., $sin(2\pi ta/24)$) that were very difficult to directly relate to specific start and end times.

The revised model is a multinomial logit model that instead uses constants for specific work start and end time periods and for durations. It proved much simpler to calibrate to match the observed distributions of start times, end times, and durations quite well. Besides the time period constants (for both start and end times for the work activity), model parameters included constants for work duration in hours (ranging from 0-1 hours to greater than 10 hours), and constants retained from the original GA5 model specific to departure times if the worker is a mother. Also retained from the original model were two work flexibility variables and variables representing the expected home to work in-vehicle travel time and cost.

The constants for the new model were not estimated from the RHTS data but rather were calibrated to match the observed diurnal distribution of work start and end times. The coefficients for the work flexibility variables and the variables representing the expected home to work in-vehicle travel time and cost were retained from the original GA5 model.

4.3 Non-Worker Home Stay Duration Before Tour

This model was originally formulated as a log-linear regression model. It did not accurately reflect temporal distributions of non-worker activity start times and was prone to producing values outside of the acceptable time of day range (i.e., after the end of the travel day) under certain combinations of inputs, due to constraints imposed from earlier models (e.g., previously scheduled joint travel).

This model was reformulated as a multinomial logit model of activity start time for non-workers, with a form similar to the start time component of the revised work start and end time model, which proved much easier to calibrate to reflect observed diurnal distributions of activity start times. There are 48 alternatives, representing half-hour periods.

4.4 Mode Choice Models

The mode choice models employ the nested logit structure. The original validation results indicated that they did not capture observed variations in transit use by time of day.

The model structures were not changed for these models. Rather, mode-specific variables specific to time of day were added to the worker and non-worker trip mode choice models, and the models were recalibrated to better reflect these temporal differences. Specifically, variables were added to each modal utility function with a value of one if the (majority of the) trip takes place in each of the four assignment periods defined for the NYBPM (a.m. peak, mid-day, p.m. peak, night) and zero otherwise.

Initially the coefficient of each variable was set to zero (meaning that the modal utilities did not vary by time of day, as in the original models), and the values of the coefficients were calibrated in an attempt to better reflect differences in transit use by time period while retaining the already reasonable distributions of auto trips by time of day. The inclusion of these variables met with limited success. Commuter rail trips in the two off-peak (mid-day and night) periods remained higher than observed as indicated in the transit validation database, though some improvement was achieved. However, additional calibration was not resulting in further improvement. As the model is further refined, it will make sense to look at other model components, including the various tour and trip time of day models, the commuter rail skims by time periods, and others, as well as the observed data to determine how further improvement might be attained.

4.5 CEMSELTS Models

Finally, it should also be noted that as part of this task, revisions to several CEMSELTS models were made:

- The following models were eliminated because the results are not used downstream, and the results from these models do not change the validation outcome: parking pass, annual mileage, vehicle fleet composition, and primary driver allocation.
- Intrazonal variables were added to the school location model.
- Household size and geographic dummy variables were added to the school location model.
- Geographic dummy variables were added to the housing type model.

5.0 Model Implementation

The 2019 NYBPM is implemented using a combination of proprietary, open source, and custom software and code. These include the following:

- TransCAD, proprietary travel modeling software from Caliper Corporation
- TransCAD GISDK scripts from the 2012 NYBPM, adapted to reflect the specifics of the 2019 NYBPM
- Newly written (by CS) TransCAD GISDK scripts
- PopGen version 2 open source software (programmed in Python)
- CEMSELTS and CEMDAP C++ application code, originally written by UTA, adapted by CS and UTA to reflect the specifics of the 2012 NYBPM, and revised by CS to accommodate model changes for the 2019 NYBPM.
- Python scripts that connect TransCAD, the PostgreSQL database, and CEMSELTS/CEMDAP code.

The 2019 NYBPM User Guide (Cambridge Systematics, Inc., 2023b) provides specific information to model users on the model structure, hardware requirements, software structure, interface, and data files. A separate memorandum (Cambridge Systematics, Inc., 2021) documented the work done to develop the model implementation for the 2012 NYBPM, including the TransCAD interface development, the adaptation of the CEMSELTS/CEMDAP code, the PostgreSQL database, and reporting functions; most of which remains relevant for the 2019 NYBPM.

5.1 TransCAD Interface and Components

NYMTC continues to use the TransCAD software platform, from Caliper Corporation, and its flowchart Interface for running the 2019 NYBPM. This approach streamlined model application and analysis. The platform was updated to use TransCAD version 9, including Caliper's most current flowchart interface and scenario management system.

CS developed the model implementation using TransCAD as the model interface and to perform many of the model functions, including network skimming and assignment, application of non-activity-based model components, and reporting. The TransCAD interface calls the applications for the activity-based NYBPM components that run using separate code, including the synthetic population generation using PopGen and the activity-based demand application (CEMSELTS and CEMDAP). Scenario and file management is achieved through a scenario management system integrated into the user interface.

The model includes a combination of steps run in TransCAD (e.g., networks, path building, assignment, visitor model), Python (e.g., PopGen2), and in the custom CEMSELTS and CEMDAP programs. Much of the model data is stored in a PostgreSQL database, with Python scripts used to help tie the different components together.

The structures for most of the non-activity-based demand components, including visitor travel, truck trips, airport trips, and external travel, were not revised as part of the 2019 NYBPM development. The truck and external models were not revised, but the trip tables from the 2012 NYBPM were updated for 2019 and forecast years using a factoring process. The external trip component of the model is based on 2019 traffic counts at all external stations along with disaggregated external trip tables, which are used as a seed and expanded using an iterative proportional fitting (IPF) procedure to match traffic volumes at all external stations. In forecast year models, updated external volumes are input to the process. The model scripts expand the external seed matrices for consistency with forecasted external volumes.

Truck trip tables for the 2019 NYBPM are based on disaggregated truck trip tables from the 2012 model, factored based on results of initial assignment of trucks. Resulting truck VMT on links with counts was compared to count VMT, resulting in adjustment factors. These adjusted truck trip tables are input directly into the model. Forecast year truck trip tables were developed by inflating the 2019 truck tables using TAZ level socioeconomic data. The inflation process used the IPF procedure to increase trip activity-based on employment growth rates in each TAZ. The resulting forecast year truck tables have been provided with the travel model dataset and are input directly into the model.

The TransCAD flowchart interface provides a streamlined approach to running the travel model and managing multiple scenarios. CS implemented the 2019 NYBPM through the following steps:

- Updated the model to work with the scenario manager toolbar available in TransCAD 9.
- Made updates to the model flowchart structure, including:
  - Move the time period network process from a utility to a step within the model flow.
  - Remove the *Trip Based Skims* and *Short Dist Trucks and Commercial Vans* steps from the model flowchart. *Trip Based Skims* are not needed because all necessary skims are created during the H*ighway Skims* step. The *Short Dist Trucks and Commercial Vans* step is now addressed during the *Combine OD Tables* model step.
- Implemented the method for storing and updating parameters within the Flowchart Scenario Manger.
- Defined a folder structure used for storing model input, output, and other files associated with running a scenario.
- Used utilities originally developed for the 2012 NYBPM to aid in managing and running the model, to aid in management and creation of transportation networks, and to create separate time period networks for the time periods used in the 2019 NYBPM.

CS developed a customized reporting system that outputs an html file (NYMTC_Summary.html) for each scenario/alternative containing key model results and validation statistics. This file includes the following information:

- Summary of key highway network statistics
- Origin-destination trip summary by mode at the district level
- Vehicle-miles traveled and vehicle-hours traveled by facility type and area type at the district level, for each time period
- Transit boarding summary by transit mode and access mode

- Transit origin-destination trip summary by transit mode and access mode at the district level by time period
- Various summaries used in highway and transit assignment validation (usable only for base year validation since the observed data represent the base year)

5.2 PopGen

PopGen is an open source synthetic population generator. Arizona State University (ASU), where PopGen was developed, originally adapted PopGen for use in the 2012 NYBPM and revised the PopGen application to produce the synthetic populations needed to run the NYBPM. ASU prepared complete documentation (Arizona State University and Cambridge Systematics, 2021) of the NYBPM PopGen application, which describes the setup of PopGen, data and input file preparation, configuration files, scenario generation, post processing steps, output files.

While PopGen can be run in a standalone manner, it is really meant to be run directly from the TransCAD user interface of the NYBPM, as described in the user guide. Running PopGen from the interface is a user option since some applications do not require generation of a new synthetic population. PopGen only needs to be run when the socioeconomic dataset has changed or been revised.

5.3 CEMSELTS/CEMDAP Code

For the 2012 NYBPM, the CEMSELTS/CEMDAP code was substantially revised and updated by CS and UTA to run the newly designed and validated model. The code changes accomplished the following:

- Code revisions to accommodate the CEMDAP revisions that replaced the various “travel time to stop” models for workers, non-workers, children, and joint tours in the original CEMDAP structure with “travel distance to stop” models.
- Code revisions to accommodate revisions to the worker mode choice models that allow for mixed-mode tours for workers.
- New code to apply a new model added to CEMSELTS, the vehicle ownership model.
- Code changes to incorporate new variables that were not included in the original CEMSELTS/CEMDAP models. Some of these new variables were included in model estimation while others were added during model validation. Most of these were geographic-specific variables relevant for the NYMTC model region, such as indicator variables for specific origin-destination regions like Manhattan.

The CEMSELTS/CEMDAP code in some cases pulls the necessary outputs from previous steps from the database to use as input to subsequent components. In other cases, the legacy code that passes information directly from one component to the next was unchanged. As previously noted, the results of the CEMSELTS and CEMDAP components are written to the database by the code.

For the 2019 NYBPM, the CEMSELTS/CEMDAP code was updated to reflect changes form the 2012 NYBPM, including the changes described in Section 4.0.

6.0 Model Validation

This chapter summarizes the validation of the NYBPM. The model validation report (Cambridge Systematics, Inc. et al., 2023) provides complete details of the validation process and results. Presented here are summaries of the validation process and some of the key validation results.

A model validation plan (Cambridge Systematics, Inc. and EA Harper Consulting, 2021a) was developed prior to model development. This plan laid out the process for the model validation and specified the tests that were performed. The tests in the plan included checks of the results of all model components compared to the observed data, checks of the highway and transit assignment, and tests of the sensitivity of the model to changes in input data.

A major component of the validation process was the comparison of model results for the base year of 2019 to observed data (see Section 3.4 for a summary of the observed data used in validation). This consisted of the validation of the demand components (i.e., CEMSELTS and CEMDAP), which is summarized in Section 6.1, and the validation of highway and transit assignment, discussed in Section 6.2. The other major validation component was sensitivity testing, where the model results for scenarios where key inputs are changed are compared to the base year scenario results. This was done to test the sensitivity of the model to various key inputs related to the types of analyses the NYBPM will be used for. Sensitivity testing is summarized in Section 6.4.

6.1 Validation of Activity-Based Demand Components

To validate the CEMSELTS and CEMDAP components, the estimated models were applied in sequence in a "single pass" using the congested speed network data used for the model estimation process. Model parameters and constants were adjusted as necessary to better replicate expanded data from the observed survey data. The single-pass process involved the following steps:

- **Apply each estimated model component** using the skims and socioeconomic data used for model estimation. The applied model results were compared to the validation targets.
- **Calibrate individual CEMSELTS and CEMDAP components**. Based on the above step, each component was calibrated by adjusting parameter values and as necessary. In a few cases, models were re-estimated to include new variables.
- **Examine error propagations**. As all of the activity-based model components are linked to one another and applied in sequence, each subsequent model component is affected by models upstream. Doing a single-pass validation therefore helped to understand the magnitude and direction of error propagation through the model system.

The comparisons of model results to observed data were performed for market segments relevant to the particular component such as subregions, household characteristics such as income, and personal characteristics such as age and gender. These comparisons were done using Excel spreadsheet files. R scripts were used to export data from the model database and process the data to be imported into the Excel spreadsheets, which were populated in advance with the observed data summaries.

In some cases, model parameters were adjusted ("calibrated") to produce more reasonable results. There was not, however, a universal attempt to match all results from the observed for all market segments by adjusting model constants or other parameters. Calibration adjustments were made only when the uncalibrated model results appeared unreasonable and the survey data results were based on a substantial number of observations. The specific calibration adjustments are documented in the Excel files, which were provided to NYMTC along with the model validation report.

The purpose of the base year comparisons was to verify, to the extent possible, that the model produces reasonable estimates of travel behavior. While it is desirable for a model's base year scenario to reasonably reflect the observed data, a more important objective is for the model to react correctly when run for scenarios representing transportation system, policy, or land use changes that planners wish to study. It is usually possible to improve the match between model results and observed data by adding or making changes to the values of parameters pertaining to various travel market segments, but while such parameters are added for better prediction of variables that obviously need correction, increasing the effects of such parameters—such as constants—can make the model less sensitive to factors that affect travel in these scenarios.

The remainder of this section summarizes the validation results for CEMSELTS and CEMDAP. A few selected tables from the Excel spreadsheets for key models are presented as examples of the validation results.

6.1.1 CEMSELTS Components

The following is a summary of the base year model comparisons for CEMSELTS components. For these models, the observed data used for comparison come from the RHTS, except where noted.

- Education attainment – Regional model results are within one percentage point of the observed, and comparisons for all segments are within seven percent of observed.
- School location – Average modeled home-school distance is within 0.4 miles of observed; the coincidence ratio for the distance frequency distribution is 82 percent. Average modeled home-school distance comparisons by subregion are shown in Table 6-1.
- College location – Average modeled home-college distance is within one mile of observed; the coincidence ratio for the distance frequency distribution is 83 percent.
- Labor force participation – Regional model results are essentially the same as observed; results by subregion, age group, and gender are all within five percent. As a reminder, this model simulates whether a person is employed or not, rather than a labor force participation rate as used by the Bureau of Labor Statistics (BLS).
- Employer type – Regional model results are essentially the same as observed; results by subregion, age group, and gender are all within 12 percent (all but three within five percent).
- Occupation industry – Regional model results are with five percent of observed; results by subregion, age group, and gender are all within 10 percent (most within five percent).
- Household income – Regional model results for all income groups are very close to observed (there is a slight shift from $150K-$200K to $100K-$150K in the model). Results by subregion and other segments are all close.
- Residential tenure – Regional model results are essentially the same as observed; results by subregion and other segments are mostly within five percent of observed.
- Housing type – Regional model results are low for apartments and high for "other"; results by age group, and gender track similarly. Given the separate models for own vs. rent and inside/outside the city, it is likely that a population density variable would probably reduce the impacts of geographic-

specific variables without changing the overall results much. As such, a population density variable is not included in this model.

- Employment location – Regional results show that the modeled average home-work distances for the states of New York and New Jersey are essentially the same as observed. The modeled averages for subregions (shown in Table 6-2) are generally within a mile or so of observed but are farther off for a few of the more remote subregions. The coincidence ratio for the distance frequency distribution is 78 percent. An additional comparison was performed by comparing the modeled subregion to subregion home-work distribution to the distribution from the ACS; the comparison between home and modeled work locations showed a very close match with the ACS data.

  An additional comparison was performed by comparing the modeled subregion to subregion home-work distribution to the distribution from the ACS. It should be noted that the home-work location information in the ACS data is not completely consistent with the definition of "regular workplace" in the NYBPM, which was estimated from RHTS responses regarding regular work location. Nevertheless, the comparison between home and modeled work locations showed a very close match with the ACS data.

- Weekly work duration – Regional model results are within three percent of observed; results by subregion, age group, and gender are mostly within five percent. Note that this model underestimates the share working in Manhattan for more than 45 hours/week by about 14 percent. Also playing a factor here is that the survey—or any observed data—has inherent error and that the survey data is expanded differently than the synthetic population used to generate the CEMSELTS results.

- Work flexibility – Regional model results are within about two percent of observed; results by subregion, age group, and gender are mostly within five percent. However, the difference for some segments by age group is greater. Related to the zero coefficient for the 18–25 age category (calibration tab of the template for this model), the impacts of the age group variables are all relative to one another.

- Driver's license – Regional model results are within one percent of observed; results by subregion, age group, and gender are all within five percent.

- Vehicle ownership – The modeled regional percentages of households by number of vehicles match the observed regional within one percent. With a few exceptions, the percentages by subregion, household size, income level, and number of workers are within five percentage points of observed.

Table 6-1. Average Modeled and Observed Home-School Distances (miles)

Subregion	Expanded RHTS data	Model	Difference (Model – Survey)
Manhattan	2.4	1.9	-0.5
Other NYC	2.7	2.5	-0.2
Long Island	3.1	2.8	-0.3
Westchester-Putnam-Dutchess	3.4	2.9	-0.5
Rockland-Orange	7.2	2.6	-4.6
Bergen-Passaic	2.2	2.5	0.3
Essex-Hudson-Union	2.1	2.3	0.2
Middlesex-Morris-Somerset-Mercer	2.3	2.6	0.3
Monmouth-Ocean	2.8	2.6	-0.2
Hunterdon-Sussex-Warren	4.6	3.3	-1.4
Connecticut	3.3	2.5	-0.9
Region	2.9	2.5	-0.4

Table 6-2. Average Modeled and Observed Home-Work Distances (miles)

Subregion	Expanded RHTS data	Model	Percent Difference (Model - Survey)
Manhattan	5.1	3.8	-1.3
Other NYC	8.3	8.6	0.3
Long Island	15.4	12.9	-2.5
Westchester-Putnam-Dutchess	15.2	16.4	1.1
Rockland-Orange	20.9	17.9	-3.0
Bergen-Passaic	11.9	10.2	-1.7
Essex-Hudson-Union	9.8	8.9	-1.0
Middlesex-Morris-Somerset-Mercer	14.0	13.9	-0.0
Monmouth-Ocean	18.0	18.7	0.8
Hunterdon-Sussex-Warren	21.2	22.4	1.2
Connecticut	12.0	12.1	0.1
Region	11.7	11.2	-0.6
New York	10.7	10.1	-0.6
New Jersey	13.4	13.0	-0.5

6.1.2 CEMDAP Components

The observed data for the CEMDAP comparisons come from NYMTC's RHTS. The GA series was validated first, and the four remaining series were validated in parallel. The following is a summary of the base year model comparisons.

GA Series

- GA1 – Child's decision to go to school – Regional shares of children who did and did not go to school on the travel day are the same as observed, and comparisons by grade levels are close. The RHTS data shows some variation by subregion which is not captured by the model (though it is unclear why attendance rates among subregion should vary much).
- GA2/GA3 – Child's school start and end times – The modeled average school activity duration is about 6.9 hours, compared to just over 7.0 hours in the expanded RHTS data set. The coincidence ratios between the modeled and RHTS diurnal distributions at the hourly level are 71 percent for start times and 62 percent for end times.
- GA4 – Adult's decision to go to work – Regional model results are within five percentage points of observed, and comparisons for subregions, age levels, and work durations are close.
- GA5 – Adult's work start and end times – The modeled average work activity duration is 8.0 hours, compared to 7.6 hours in the expanded RHTS data set. The coincidence ratios between the modeled and RHTS diurnal distributions at the hourly level are 98 percent for start times and 98 percent for end times. See the WorkStartEnd (hours) tab on the validation sheet for this model for more details. The modeled and RHTS percentages of a.m. and p.m. peak period start and end times are shown in Table 6-3.
- GA6 – Adult's decision to go to school – Regional model results essentially the same as the observed, and comparisons by subregions and gender are within 10 percentage points.
- GA7/GA8 – Adult's school start and end times – The modeled average school activity duration is 5.9 hours, compared to 5.7 hours in the expanded RHTS data set. The coincidence ratios between the

modeled and RHTS diurnal distributions at the hourly level are 43 percent for start times and 52 percent for end times.

- GA9/GA10 – Child's travel mode to school and from school – At the regional level, the modeled shares for all modes are within one percentage point of the observed. In the subregional summaries, there are some differences between modeled and observed mode shares. The largest of these are in the splits between school bus and walk mode shares in Connecticut and most of New Jersey. These mode shifts do not affect trip assignment since neither walk nor school bus person trips are assigned.
- GA11/GA12 – Allocation of drop-off and pickup episodes to parent – Regional model results are within one percent of observed for drop-off and within three percent for pickup.
- GA13 – Determination of households with non-zero out-of-home duration – The regional percentage of households with non-zero out-of-home activities is within one percent of observed. Comparisons by subregion, income level, and household size are mostly within five percent.
- GA14 – Determination of total OH time of a household – The modeled percentage of time spent in home is 60 percent, compared to 68 percent in the RHTS data. However, it is believed that the aggregate percentage of time spent inside the home as reported in the RHTS is too high to be consistent with observed regional travel counts. The lower percentage of time spent inside the home was therefore deemed acceptable.
- GA15/GA16 – Independent and joint activity participation for households – The modeled percentages of time spent by activity type are within one percentage point for both individual and joint activities. Modeled numbers of participants in joint activities are very close to observed.
- GA17 – Decision of adult to undertake other serve-passenger activities – Regional model results very close to those observed in the expanded RHTS data set, as were model results by gender and employment status. The model reflected that serve-passenger activity participation was lower in Manhattan (though not quite as low as in the observed data).

Table 6-3. Modeled and Observed Percentages of Work Start and End Times in Peak Periods

Peak Period	Expanded RHTS data		Model Results	
	Start	End	Start	End
AM (6:00-10:00)	76.8%	1.8%	77.4%	1.7%
PM (3:00-7:00)	4.7%	69.9%	4.7%	70.2%

WSCH Series

- WSCH1 – Worker commute mode – As one of the key components of the entire model, significant attention was paid to the validation and calibration of this component. This included revisiting the validation of this component after the initial highway and transit assignment results to better reflect observed travel conditions, including observed data from sources such as transit rider surveys and the hub-bound report. Based on lessons learned from the validation of the 2012 NYBPM, changes were made that caused some mode choice results to differ from those observed in the RHTS data set; for example, based on commuter rail survey data and the trip level data from the RHTS, the auto access and walk access splits for the commuter rail/bus mode from the RHTS appear to be incorrect. Therefore, the relative shares of auto access and walk access for the commuter rail/bus mode were calibrated to match observed shares from commuter rail survey data rather than the RHTS.

  Table 6-4 compares the regional model results to the targets. The model is producing about the correct number of trips by mode though walk trips are overestimated by the model compared to the RHTS data. Generally, the trends in the model results track those in the RHTS data. For example:

  - For Manhattan residents, auto mode shares are very low (less than 10 percent) while transit shares exceed 60 percent, and non-motorized mode shares are around 20 percent.
  - The transit shares are slightly lower for residents of the other New York City boroughs while the auto shares are around 30 percent, and non-motorized mode shares are under 10 percent.
  - In the rest of the region, auto shares are in the 80 to 90 percent range, except in Essex/Hudson/Union Counties in New Jersey, where the auto share is under 70 percent. The highest auto shares are in the subregions farthest from New York City. Transit shares are in the 15 to 25 percent range in the nearest subregions to the city and are under 10 percent in the rest of the region. The non-motorized shares are under five percent outside New York City and are lower the farther from the city.
  - Travelers from households with annual incomes below $30,000 have auto mode shares of around 40 percent, transit mode shares around 40 percent, and non-motorized mode shares around 15 percent.
  - Travelers from households who own zero vehicles have transit shares around 70 percent and non-motorized mode shares around 20 percent.

- WSCH2 – Number of before-work tours – The model closely matches the observed percentages of workers with zero, one, and two or more before-work tours.
- WSCH3 – Number of work-based subtours – The model closely matches the observed percentages of workers with zero, one, and two or more work-based subtours.
- WSCH4 – Number of after-work tours – The model closely matches the observed percentages of workers with zero, one, and two or more after-work tours.
- WSCH5/WSCH6/WSCH7 – Before-work/work-based/after-work tour mode – The models closely match the observed regional mode shares for these tours made by workers. Modeled mode shares for the various geographic and demographic segments are generally consistent with observed mode shares, for segments with large enough sample sizes to make worthwhile comparisons.
- WSCH8 – Worker number of stops on commute/before work/after work/at-work tours – The models estimate fewer stops on the way to work and more stops on the way home from work than the RHTS data indicate. However, the model results seem reasonable given the higher propensity to stop while going home than going to work. Evidence from validation counts by time period supports this.
- WSCH9 – Worker home or work stay duration before tour – The model simulates longer average durations than observed for before-work and at-work tours and similar durations for after-work tours.
- WSCH10 – Worker activity type at stop – The model estimates similar percentages of activities for all purposes to observed data.
- WSCH11 – Worker activity duration at stop – The model estimates the average activity duration at a stop at 65 minutes, compared to 43 minutes in the RHTS data. The modeled average duration is within a few minutes of the observed for most activity purposes (30 minutes different for the activity with the longest duration—work-related).
- WSCH13 – Worker location of a stop – The modeled average trip distance is very close to the observed average of 5.6 miles. The modeled averages for subregions are all close to the observed. The modeled averages for subregions are all close to the observed (shown in Table 6-5). The coincidence ratio for the distance frequency distribution is 78 percent.
- WSCH14 – Worker commute trip mode choice – The worker commute trip mode choice is closely related to the mode choice for the commute tour (model WSCH1, discussed above). The regional modeled and observed trip mode shares are similar to those shown for model WSCH1 in Table 6-4. The model results show that, as is observed, trips on commute tours tend to use the same modes as the tour mode. Model results for geographic and demographic segments also match the observed mode shares well.

Table 6-4. Modeled and Observed Regional Worker Commute Mode Shares

Tour Mode	Commute to Work		Commute from Work	
	Observed*	Model	Observed*	Model
SOV	45.9%	44.2%	46.1%	43.4%
HOV 2	11.1%	11.1%	10.9%	10.7%
HOV 3+	5.4%	6.2%	5.2%	5.7%
Taxi	0.9%	0.7%	1.2%	0.6%
Commuter rail/bus – auto access	0.7%	4.1%	1.1%	4.4%
Commuter rail/bus – walk access	7.0%	3.4%	6.7%	3.5%
Subway/ferry – auto access	0.3%	0.4%	0.6%	0.3%
Subway/ferry – walk access	16.4%	14.8%	16.0%	12.8%
Local bus – walk access	5.1%	2.3%	4.5%	5.4%
Walk	5.1%	10.8%	5.7%	11.3%
Bike	0.6%	0.4%	0.6%	0.4%
School bus	1.5%	1.7%	1.6%	1.5%

* - From RHTS

Table 6-5. Average Modeled and Observed Worker Stop Distances (miles)

Subregion	Expanded RHTS data	Model	Difference (Model - Survey)
Manhattan	1.9	1.8	-0.1
Other NYC	4.1	4.3	0.3
Long Island	7.1	7.4	0.3
Westchester-Putnam-Dutchess	6.4	6.5	0.1
Rockland-Orange	8.7	8.4	-0.3
Bergen-Passaic	5.2	5.4	0.2
Essex-Hudson-Union	4.2	4.3	0.0
Middlesex-Morris-Somerset-Mercer	6.7	6.7	0.0
Monmouth-Ocean	8.5	8.5	0.0
Hunterdon-Sussex-Warren	13.2	13.2	0.0
Connecticut	6.2	6.2	0.0
Region	5.6	5.6	0.0

NWSCH Series

- NWSCH1 – Non-worker number of independent tours – The model closely matches the observed percentages of non-workers with zero, one, and two or more independent tours.
- NWSCH2/NWSCH3 – Non-worker decision to undertake independent tour before/after pick-up or joint discretionary tour – The model matches the observed percentages of non-workers who choose to undertake an independent tour before a pick-up or joint discretionary tour within four percent and after a pick-up or joint discretionary tour within one percent..
- NWSCH5 – Non-worker number of stops in a tour – The model closely matches the observed percentages of non-workers with one, two, three, four, five, and six stops on tours.
- NWSCH6 – Non-worker number of stops following pick-up/drop-off – The model slightly overestimates the percentage of non-workers who make four or more stops and somewhat underestimates the percentage who make three stops.

- NWSCH7 – Non-worker home stay duration before tour – The modeled distribution of tour arrival time closely matches the observed tour arrival time distribution.
- NWSCH8 – Non-worker activity type at stop – The model matches the observed percentages for all activity types except serve passenger and work-related within one percentage point. Since work-related stops are special cases for non-workers, it was difficult to simulate many of these types of stops.
- NWSCH9 – Non-worker activity duration at stop – The model underestimates the average activity duration at a stop by less than 13 minutes for every activity purpose. The model does underestimate durations for younger adults.
- NWSCH11 – Non-worker stop location – The modeled average trip distance is 5.7 miles, compared to the observed average of 5.7 miles. The modeled averages for subregions are all close to the observed (shown in Table 6-6). The coincidence ratio for the distance frequency distribution is 88 percent.
- NWSCH4 – Non-worker trip mode – Table 6-7 shows the regional modeled and observed non-worker trip mode shares. The modeled walk shares are slightly low.

Table 6-6. Average Modeled and Observed Non-Worker Trip Distances (miles)

Subregion	Expanded RHTS data	Model	Difference
Manhattan	2.7	2.7	0.0
Other NYC	4.3	4.6	0.3
Long Island	7.4	7.5	0.1
Westchester-Putnam-Dutchess	5.6	5.7	0.1
Rockland-Orange	9.4	9.0	-0.4
Bergen-Passaic	5.8	5.9	0.1
Essex-Hudson-Union	5.2	5.3	0.1
Middlesex-Morris-Somerset-Mercer	6.4	6.5	0.1
Monmouth-Ocean	8.3	8.3	0.1
Hunterdon-Sussex-Warren	12.5	12.5	0.0
Connecticut	5.4	5.4	0.1
Region	5.7	5.7	0.0

Table 6-7. Modeled and Observed Regional Non-Worker Trip Mode Shares

	Observed*	Model
SOV	45.5%	46.4%
HOV 2	15.9%	17.6%
HOV 3+	7.6%	8.7%
Taxi	1.2%	0.3%
Commuter rail/bus – auto access	0.5%	0.4%
Commuter rail/bus – walk access	0.6%	0.4%
Subway/ferry – auto access	0.2%	0.7%
Subway/ferry – walk access	4.4%	4.6%
Local bus – auto access	0.0%	0.0%
Local bus – walk access	3.3%	5.2%
Walk	19.8%	15.3%
Bike	0.8%	0.4%

* - From RHTS

JASCH Series

- JASCH2 – Joint activity start time – The overall coincidence ratio between the modeled and observed temporal distributions of joint activity start times is 67 percent. Please refer to the Start Time (hours) tab on the validation sheet for this model for more details. The main difference is that the model form tends to overestimate start times in the final period of the day (8:00 p.m. to 3:00 a.m.). This reflects the difficulty of scheduling joint activities among household members who have various other commitments such as work and school that must be accommodated. The model estimates the percentage of joint activities in the a.m. peak period well but does tend to underestimate p.m. peak period activity start times.

- JASCH4 – Joint activity location – The modeled average trip distance is 4.8 miles, compared to the observed average of 4.8 miles. The model underestimates very short trips of less than one mile and overestimates trips of two to five miles. The coincidence ratio for the distance frequency distribution is 0.57.

- JASCH6 – Joint discretionary trip mode choice – The regional modeled and observed joint trip mode shares are close for all modes.

CSCH Series

- CSCH4 – Child departure time from home for independent discretionary tour – The overall coincidence ratio between the modeled and observed temporal distributions of child (non-school) activity start times is 66 percent. While peak departure times occur at the observed times of day, the model tends to underestimate a.m. peak period activity start times and overestimate p.m. peak period activity start times.

- CSCH5 – Child activity duration at independent discretionary stop – The model simulates the average activity duration at a stop within 25 minutes of the observed.

- CSCH7 – Child location of independent discretionary stop – The modeled average trip distance is 3.0 miles, compared to the observed average of 3.0 miles. The modeled averages for subregions are all very close to the observed. The coincidence ratio for the distance frequency distribution is 57 percent.

- CSCH3 – Child mode for independent discretionary trip – There is some underestimation in the model of transit and walk trips and a corresponding overestimation of auto trips.

6.2 Highway and Transit Assignment Validation

6.2.1 Highway Assignment

The highway validation focused on three main classes of measures:

- Summaries of vehicle-miles of travel (VMT);
- Summaries of individual link traffic volumes; and
- Intra-regional traffic flows as defined by screenlines.

All of these measures are based on comparisons of assigned volumes from the model to observed traffic counts. Due to the large number of jurisdictions that maintain the roads in the network and the variety of roadway types, the counts are assembled from several sources. Generally, the highway assignment results match observed data reasonably well, with no major high or low biases compared to traffic counts.

VMT Checks

For the region, the modeled VMT on links with traffic counts is about half a percent lower than the observed VMT computed from the counts. Table 6-8 shows the modeled and observed VMT by facility type, with the percentage difference compared to the targets from the model validation plan. All targets are met. There are no targets for the (generally low volume) local streets and ramp facility types, which comprise less than one percent of the VMT.

Table 6-8. Modeled and Observed Daily VMT by Facility Type

	Model VMT	Count VMT	Total	Target
Interstate/Freeway/Tollway	21,452,040	21,576,781	-0.6%	7%
Principal Arterial	3,473,982	3,587,811	-3.2%	10%
Minor Arterial	2,573,652	2,546,392	1.1%	10%
Major Collector	669,231	600,967	11.4%	15%
Minor Collector	135,377	146,440	-7.6%	15%
Local Street	42,442	31,915	33.0%	
Ramp	16,907	27,754	-39.1%	
Total	**28,363,631**	**28,518,060**	**-0.5%**	**1%**

The percentage differences between modeled and observed VMT for the four time periods used in highway assignment are:

- AM peak (6:00 AM - 10:00 AM): 0.4%
- Mid-day (10:00 AM – 3:00 PM): -12.9%
- PM peak (3:00 PM – 7:00 PM): -2.4%
- Night (7:00 PM – 6:00 AM): -5.2%

Note that this summary does not include all links included in the summary shown in Table 6-8; there are some links with daily counts but not counts by time of day.

Table 6-9 shows the modeled and count VMT for a set of districts that comprise the entire region.

Table 6-9. Modeled and Observed Daily VMT by Subregion

	Model VMT	Count VMT	% Difference
Manhattan CBD	525,384	456,193	15.2%
Upper Manhattan	1,843,952	1,680,290	9.7%
Other NYC	3,031,361	3,493,830	-13.2%
Long Island	3,930,448	4,349,865	-9.6%
Mid-Hudson	125,697	140,581	-10.6%
NJTPA Core	5,113,496	5,445,481	-6.1%
NJTPA Other	8,859,255	8,123,192	9.1%
Connecticut	3,841,561	3,886,160	-1.1%
Mercer County, NJ	1,092,477	942,466	15.9%
Total	**28,363,631**	**28,518,058**	**-0.5%**

Link Volume Checks

The overall fit between individual modeled and observed link volumes was examined using the percentage root mean square error (%RMSE) measure. Table 6-10 shows the %RMSE grouped by volume group.

While the %RMSE meets the target for the entire set of all links with counts, the %RMSE for the individual segments do not meet most of the targets from the validation plan. This may be due to the issues with some count locations and network loading points. For example, the modeled volumes on roadways where zone centroid connectors meet the highway network may be high if actual network loading points for the zone are more dispersed; conversely, modeled volumes on roads where trips from a zone are actually loading may be low if the zone's centroid connectors are not nearby.

Table 6-10. %RMSE by Volume Group

Volume Group	Links	% RMSE	Target
0 - 1,000	23	557%	100%-200%
1,000 - 5,000	176	219%	45%-100%
5,000 - 10,000	277	76%	36%-45%
10,000 - 20,000	328	57%	28%-34%
20,000 - 30,000	152	41%	24%-26%
30,000 - 50,000	160	41%	21%-24%
50,000 - 100,000	236	24%	12%-21%
100,000 and up	69	19%	12%
All Links	**1,421**	**40%**	**40%**

Table 6-11 shows the VMT, as estimated by the model and observed through traffic counts on eight major routes that have at least 75,000 VMT on links with counts. Most of these routes had modeled VMT within 10 percent of observed. The model most notably overestimates volumes on the West Side Highway.

Table 6-11. Modeled and Observed VMT on Major Routes

	Model VMT	Count VMT	% Diff.
Garden State Parkway	9,548,206	9,425,144	1%
New Jersey Turnpike	2,429,373	2,892,386	-16%
New York State Thruway	224,968	248,585	-10%
Long Island Expressway	339,908	316,400	7%
I-95	189,257	179,039	6%
I-287	153,029	157,901	-3%
West Side Highway	120,308	75,730	59%
FDR Drive	181,010	175,073	3%

Table 6-12 shows a comparison of volumes on the major crossings into and within New York City, grouped by waterway and location. With the exceptions of the crossings that represent single bridges, each group's modeled volume is within 10 percent of the traffic counts, except the Hudson River, with has a 14 percent difference.

Table 6-12. Modeled and Observed Volumes on Major Crossings

	Links	Model	Count	% Diff.
Outerbridge Crossing	2	78,158	79,194	-1.3%
Goethals Bridge	2	71,296	82,852	-13.9%
Arthur Kill Subtotal	**4**	**149,454**	**162,046**	**-7.8%**
Bayonne Bridge	2	42,956	30,754	39.7%
Kill Van Kull Subtotal	**2**	**42,956**	**30,754**	**39.7%**
Holland Tunnel	2	94,271	87,985	7.1%
Lincoln Tunnel	4	105,385	95,582	10.3%
G Washington Bridge	4	349,897	292,421	19.7%
Tappan Zee Bridge	2	183,328	166,172	10.3%
Mountain Bridge Rd	1	37,251	21,552	72.8%
Newburgh Beacon Bridge	2	78,664	82,472	-4.6%
Hudson River Subtotal	**15**	**848,795**	**746,184**	**13.8%**
Verrazano Bridge	6	256,495	205,540	24.8%
The Narrows Subtotal	**6**	**256,495**	**205,540**	**24.8%**
Hugh L Carey Tunnel	4	82,815	57,599	43.8%
Brooklyn Bridge	2	82,525	121,503	-32.1%
Manhattan Bridge	5	84,693	69,245	22.3%
Williamsburg Bridge	2	128,547	128,214	0.3%
Queens Midtown Tunnel	2	84,548	89,352	-5.4%
Ed Koch Queensboro Bridge	5	172,063	161,881	6.3%
R.F. Kennedy Bridge (Queens)	2	79,522	54,141	46.9%
Bronx Whitestone Bridge	2	135,748	126,187	7.6%
Throgs Neck Bridge	2	110,398	121,024	-8.8%
East River Subtotal	**26**	**960,859**	**929,146**	**3.4%**
R.F. Kennedy Bridge (Manhattan)	2	96,619	101,592	-4.9%
Willis Avenue Bridge – NB	**1**	**63,425**	**62,293**	**1.8%**
3rd Ave Bridge – SB	1	44,401	54,955	-19.2%
Madison Avenue Bridge	1	38,648	44,134	-12.4%
145th St Bridge	1	24,629	30,521	-19.3%
Macombs Dam Bridge	1	51,462	38,007	35.4%
Cross Bronx Exp Bridge	2	184,119	187,868	-2.0%
Washington Bridge	2	58,747	57,458	2.2%
W 207th St Bridge	1	42,267	40,078	5.5%
Broadway Ave Bridge	1	55,450	35,861	54.6%
Henry Hudson Pkwy Bridge	2	86,418	92,355	-6.4%
Harlem River Subtotal	**15**	**746,185**	**745,122**	**0.1%**
R.F. Kennedy Bridge (Bronx)	2	121,634	83,053	46.5%
Bronx Kill Subtotal	**2**	121,634	83,053	46.5%

Screenlines

To examine how well the model reflects intra-regional traffic flows, a set of 29 screenlines was defined. The validation plan defined target percentages for the difference between the summed volumes and traffic counts based on the daily traffic across the screenline. Table 6-13 shows the modeled volumes and counts for both directions for these screenlines. The volume difference meets the targets for 25 of the 29 screenlines.

Some of the major regional trip movements were examined by summing volumes for multiple screenlines. This summary is shown in Table 6-14.

Table 6-13. Modeled and Observed Volumes on Screenlines

	Links	Model NB/EB	Count NB/BE	% Diff NB/EB	Model SB/WB	Count SB/WB	Diff SB/WB	Model Total	Count Total	% Diff Total	Target
EW Border between Brooklyn & Queens	29	415,750	376,033	10.6%	431,720	429,860	0.4%	847,470	805,893	**5.2%**	29
NS Border between Bronx & Westchester	24	387,371	344,433	12.5%	238,542	220,875	8.0%	625,913	565,308	**10.7%**	24
EW Border of Dutchess NY & Litchfield CT	5	5,459	7,144	-23.6%	5,367	6,835	-21.5%	10,826	13,979	**-22.6%**	5
NS Border of Dutchess NY & Columbia NY	11	16,189	15,513	4.4%	16,321	15,751	3.6%	32,510	31,264	**4.0%**	11
NS Border between Manhattan & Brooklyn	12	197,223	204,110	-3.4%	179,967	136,203	32.1%	377,190	340,313	**10.8%**	12
EW Border between Manhattan & Queens	11	199,998	205,222	-2.5%	184,807	199,510	-7.4%	384,805	404,732	**-4.9%**	11
EW Border between Manhattan & Bronx	13	352,611	330,078	6.8%	295,508	313,452	-5.7%	648,119	643,530	**0.7%**	13
NS Broder between CBD & upper Manhattan	18	341,758	304,490	12.2%	284,145	291,179	-2.4%	625,903	595,669	**5.1%**	18
EW Border between Nassau & Suffolk	28	393,572	360,027	9.3%	372,161	356,164	4.5%	765,733	716,191	**6.9%**	28
NS Border between Nassau and Long Beach- Jones Beach	8	51,942	52,075	-0.3%	55,412	54,132	2.4%	107,354	106,207	**1.1%**	8
EW Border between NJ & Manhattan (PANYNJ Crossings)	10	268,743	233,254	15.2%	277,953	242,734	14.5%	546,696	475,988	**14.9%**	10
EW Border between NJ & Staten Island	8	137,730	103,453	33.1%	95,674	89,347	7.1%	233,404	192,800	**21.1%**	8
NS Border between Sussex NJ & Orange NY	7	35,771	33,427	7.0%	40,360	34,356	17.5%	76,131	67,783	**12.3%**	7
NS Border between Bergen NJ & Rockland NY	23	152,150	191,085	-20.4%	172,304	187,099	-7.9%	324,454	378,184	**-14.2%**	23
EW Border between Ulster NY & Duchess NY	4	29,546	30,705	-3.8%	29,404	30,938	-5.0%	58,950	61,643	**-4.4%**	4
EW US202 Bridge between Westchester & Orange	1	17,196	11,983	43.5%	19,912	9,569	108.1%	37,108	21,552	**72.2%**	1
EW I84 Bridge between Dutchess & Orange	2	38,023	41,236	-7.8%	40,424	41,236	-2.0%	78,447	82,472	**-4.9%**	2
NS Border between Orange, Sullivan, and Ulster	12	58,020	55,084	5.3%	56,213	53,860	4.4%	114,233	108,944	**4.9%**	12
EW Border between Putnam & Fairfield CT	5	51,719	44,039	17.4%	44,744	44,711	0.1%	96,463	88,750	**8.7%**	5
NS Border between Putnam & Dutchess	9	50,183	52,518	-4.4%	49,175	51,278	-4.1%	99,358	103,796	**-4.3%**	9
EW Border Between Queens & Nassau	31	588,966	517,034	13.9%	574,334	508,129	13.0%	1,163,300	1,025,163	**13.5%**	31
NS Cross Bay Blvd Between Queens & Rockaway	4	21,066	24,071	-12.5%	21,779	22,371	-2.6%	42,845	46,442	**-7.7%**	4
NS Border between Queens & Bronx	9	217,692	187,239	16.3%	253,549	216,085	17.3%	471,241	403,324	**16.8%**	9
EW Border between Westchester & Rockland (Cuomo Br.)	2	83,205	82,852	0.4%	99,127	83,320	19.0%	182,332	166,172	**9.7%**	2
NS Border Between Rockland & Orange	6	31,724	31,055	2.2%	35,094	30,777	14.0%	66,818	61,832	**8.1%**	6
EW Border between Staten Island & Brooklyn	6	120,247	107,496	11.9%	134,908	98,044	37.6%	255,155	205,540	**24.1%**	6
EW Border between Westchester & Putnam	14	108,577	63,933	69.8%	104,191	64,779	60.8%	212,768	128,712	**65.3%**	14
EW Border between Westchester & Fairfield CT	21	87,406	110,482	-20.9%	87,414	105,414	-17.1%	174,820	215,896	**-19.0%**	21

Table 6-14. Aggregate Screenline Summary

	Links	Model NB/EB*	Count NB/EB*	% Diff NB/EB*	Model SB/WB*	Count SB/WB*	Diff SB/WB*	Model Total	Count Total	% Diff Total
From/to Manhattan	46	1,027,785	982,144	4.6%	929,025	882,419	5.3%	1,956,810	1,864,563	**4.9%**
Intra-Manhattan	18	341,758	304,490	12.2%	284,145	291,179	-2.4%	625,903	595,669	**5.1%**
Other Intra-NYC	48	774,755	694,839	11.5%	841,956	766,360	9.9%	1,616,711	1,461,199	**10.6%**
Other Cross-Hudson	13	276,154	239,524	15.3%	255,137	223,472	14.2%	531,291	462,996	**14.8%**
Other Intra-NYS	124	1,633,401	1,445,146	13.0%	1,450,688	1,308,505	10.9%	3,084,089	2,753,651	**12.0%**
Other NY-NJ	30	187,921	224,512	-16.3%	212,664	221,455	-4.0%	400,585	445,967	**-10.2%**
NY-CT	26	139,125	154,521	-10.0%	132,158	150,125	-12.0%	271,283	304,646	**-11.0%**
Regional cordon	32	109,214	108,446	0.7%	107,305	107,384	-0.1%	216,519	215,830	**0.3%**

* - For "From/to Manhattan," regardless of orientation, "NB/EB" represents to Manhattan and "SB/WB" represents to Manhattan.

6.2.2 Transit Assignment

Transit assignment is performed in the NYBPM for the AM peak period. The transit assignment validation was less straightforward because of gaps in and inconsistencies among observed data sources. While attempts were made to ensure as much consistency as possible with the information from disparate sources, it was sometimes necessary to choose which measures to prioritize. In general, the goal was to make sure that total transit demand is reasonable and is consistent with areas of highest ridership. The specific checks discussed below provide some information on the results of some of these choices.

Station Groups

Four sets of station groups were defined for rail transit assignment validation. The groups are subsets of transit modes: commuter rail, PATH, subway, and light rail. Nine major commuter rail terminals are defined as individual station groups of one station only. Other commuter rail groups are aggregations of established branches or lines. Subway station groups represent the four New York City boroughs that have subway service.

Table 6-15 shows the comparison of modeled and observed a.m. peak period boardings for aggregate station group segments for which observed data are available—namely, for the three commuter rail operators (MNR, LIRR, and NJT) and PATH, for the two Midtown commuter rail stations, for other NJT services, for the MTA subway lines, and for the major stations in Brooklyn/Queens. The results indicate that:

- The model overestimates MTA subway boardings by about a third;
- Overall, modeled inbound commuter rail ridership is close to observed;
- Modeled outbound commuter rail ridership is overestimated on MNR;
- Modeled commuter rail alightings in Brooklyn and Queens are overestimated;
- Modeled inbound alightings at the two major Manhattan terminals are close to observed;
- Modeled boardings for NJT non-commuter rail services are overestimated; and
- Modeled PATH boardings are close to observed.

Table 6-15. Station Group Transit Assignment Summary

ONS	Model	Observed	Difference	% Difference
Total	2,739,215	2,155,333	583,882	27%
NJT	74,658	91,693	-17,035	-19%
MNR	120,810	103,466	17,344	17%
LIRR	97,726	90,837	6,889	8%
Penn Station/GCT	16,531	28,526	-11,995	-42%
Jamaica/City Terminal	31,473	29,476	1,997	7%
Newark/Secaucus/Hoboken	83,912	50,226	33,686	67%
MTA Subway	2,287,945	1,695,495	592,450	35%
PATH	93,572	107,300	-13,728	-13%
All commuter rail except NYC	293,194	285,996	7,198	3%
OFFS	**Model**	**Observed**	**Difference**	**% Difference**
Total	2,765,783	2,142,358	623,425	29%
NJT	21,313	23,282	-1,969	-8%
MNR	48,181	31,588	16,593	53%
LIRR	14,174	12,049	2,125	18%
Penn Station/GCT	209,530	221,777	-12,247	-6%
Jamaica/City Terminal	47,052	31,615	15,437	49%
Newark/Secaucus/Hoboken	64,165	49,966	14,199	28%
MTA Subway	2,287,946	1,695,495	592,451	35%
PATH	124,938	121,052	3,886	3%
All commuter rail except NYC	83,668	66,919	16,749	25%
TOTAL ONS AND OFFS	**Model**	**Observed**	**Difference**	**% Difference**
Total	5,504,998	4,297,691	1,207,307	28%
NJT	95,971	114,975	-19,004	-17%
MNR	168,991	135,054	33,937	25%
LIRR	111,900	102,886	9,014	9%
Penn Station/GCT	226,061	250,303	-24,242	-10%
Jamaica/City Terminal	78,525	61,091	17,434	29%
Newark/Secaucus/Hoboken	148,077	100,192	47,885	48%
MTA Subway	4,575,891	3,390,990	1,184,901	35%
PATH	218,510	228,352	-9,842	-4%
All commuter rail except NYC	95,971	114,975	-19,004	-17%

Hub-Bound Summary

The hub-bound summary (for a.m. peak boardings to and from the Manhattan CBD) is summarized in Table 6-16. The overall inbound modeled results are within three percent of the observed inbound trips except for local bus. The modeled inbound trips for each of the five corridors entering and leaving the CBD are within 20 percent of observed, except for Queens, which the model overestimates by 23 percent. Modeled trips by mode are within 10 percent of observed, except for bus, which represents about 10 percent of all inbound trips and is significantly overestimated, especially for Brooklyn and Queen. (Ferry trips from New Jersey are underestimated, but they represent only one percent of hub-bound trips.)

Outbound a.m. peak trips from the hub are about one third of inbound trips. The outbound model results (where overall numbers are lower) are 23 percent higher than observed. The highest overestimation occurs across 60th Street and to Queens, where subway is the dominant mode.

Table 6-16. Hub-Bound Transit Summary

Modeled INBOUND Hub-Bound Transit Flows

	Bus	Ferry	Rail	Subway/PATH	Total
60th St	15,751	0	79,877	239,392	**335,020**
Queens	18,926	82	78,615	325,196	**422,819**
Brooklyn	39,789	1,687	0	439,005	**480,481**
Staten Island	0	16,001	0	0	**16,001**
New Jersey	115,676	3,050	51,038	82,027	**251,791**
Total	**190,142**	**20,820**	**209,530**	**1,085,620**	**1,506,112**

Observed INBOUND Hub-Bound Transit Flows

	Bus	Ferry	Rail	Subway/PATH	Total
60th St	12,510	0	79,721	314,619	**406,850**
Queens	9,377	2,058	85,218	246,172	**343,979**
Brooklyn	14,774	1,151	0	401,874	**417,799**
Staten Island	0	14,328	0	0	**14,328**
New Jersey	116,186	14,414	62,451	84,317	**277,368**
Total	**152,847**	**31,951**	**227,390**	**1,046,982**	**1,460,324**

Modeled OUTBOUND Hub-Bound Transit Flows

	Bus	Ferry	Rail	Subway/PATH	Total
60th St	15,371	0	6,211	240,031	**261,613**
Queens	11,826	4	3,896	66,538	**82,264**
Brooklyn	1319	504	0	82,311	**84,134**
Staten Island	0	2,213	0	0	**2,213**
New Jersey	18,329	114	6,424	9,284	**34,151**
Total	**46,845**	**2,835**	**16,531**	**398,164**	**464,375**

Observed OUTBOUND Hub-Bound Transit Flows

	Bus	Ferry	Rail	Subway/PATH	Total
60th St	5,257	0	6,685	161,053	**172,995**
Queens	209	298	5,618	47,480	**53,909**
Brooklyn	278	200	0	86,563	**87,041**
Staten Island	0	2,835	0	0	**2,835**
New Jersey	32,178	1,036	10,233	16,198	**59,645**
Total	**37,922**	**4,369**	**22,536**	**311,294**	**376,425**

6.3 LBS Data Comparisons

As described in Section 3.4.3, the LOCUS LBS data set was used in model validation. This section describes some of the comparisons performed using LOCUS data. The comparisons relate to number of tours and trips, tour and trip origins and destinations, and time of day and are therefore related mainly to

CEMDAP components (except for the CEMSELTS components for workplace and school location choices—these comparisons are described in Section 6.1.1).

6.3.1 Number and Orientation of Tours

The number of tours produced by purpose are derived from the outputs of the following CEMDAP components:

- GA1 – Child's decision to go to school
- GA4 – Adult's decision to go to work
- GA6 – Adult's decision to go to school
- GA15/16 – Independent and joint activity participation
- GA17 – Decision of adult to undertake other serve-passenger activities
- WSCH2 – Number of before-work tours
- WSCH3 – Number of work-based tours
- WSCH4 – Number of after-work tours
- NWSCH1 – Non-worker number of independent tours
- NWSCH2 – Non-worker decision to undertake independent tour before pick-up/joint discretionary tour
- NWSCH3 – Non-worker decision to undertake an independent tour after pick-up/joint discretionary tour

The modeled tours were aggregated to the three segments available in the LOCUS data set (regular, non-regular, and non-home based). Note that modeled non-home based tours are limited to work-based subtours.

Table 6-17 shows the number of modeled regular tours by county (with Manhattan segmented by CBD and other) and from the LOCUS data set. Table 6-18 shows the same comparison for non-regular tours, and Table 6-19 shows the comparison for non-home based tours.

As shown in Table 6-17, the number of modeled regular tours is more than 40 percent higher than the number of regular tours in the LOCUS data set. This difference is fairly consistent across the region. A major reason for the difference is that many children, especially younger children, do not carry active mobile devices. Therefore, many school tours—which are included in the regular tours—would not be included in the LOCUS data.

Table 6-17. Modeled Number of Regular Tours Compared to Number of Regular Tours in the LOCUS Data Set

County	NYBPM	LOCUS
Manhattan CBD	361,243	305,534
Manhattan Other	551,981	389,743
Queens	1,293,098	863,580
Bronx	814,592	464,304
Kings	1,465,112	939,341
Richmond	271,539	163,824
Nassau	738,841	559,529
Suffolk	815,928	597,886
Westchester	529,431	389,354
Rockland	179,197	125,152
Putnam	54,800	38,990
Orange	216,317	140,454
Dutchess	162,745	106,211
Bergen	507,540	403,106
Passaic	289,685	207,300
Hudson	384,583	288,403
Essex	451,127	293,449
Union	309,700	228,972
Morris	281,387	212,159
Somerset	187,915	138,976
Middlesex	463,971	328,057
Monmouth	342,961	247,065
Ocean	303,969	201,563
Hunterdon	69,550	46,860
Warren	59,228	39,286
Sussex	80,660	56,354
Mercer	214,357	139,284
Fairfield	529,670	374,132
New Haven	465,229	298,202
Total	**12,396,356**	**8,587,070**

Table 6-18. Modeled Number of Non-Regular Tours Compared to Number of Non-Regular Tours in the LOCUS Data Set

County	NYBPM	LOCUS
Manhattan CBD	412,720	487,027
Manhattan Other	615,181	639,285
Queens	1,281,734	1,634,249
Bronx	732,795	945,789
Kings	1,419,070	1,775,002
Richmond	277,972	329,058
Nassau	851,053	991,413
Suffolk	907,459	1,082,437
Westchester	576,476	695,517
Rockland	192,463	229,268
Putnam	56,319	70,568
Orange	204,967	262,110
Dutchess	165,987	210,488
Bergen	589,624	674,529
Passaic	291,745	364,401
Hudson	376,318	477,282
Essex	432,140	556,555
Union	317,364	389,795
Morris	299,695	350,056
Somerset	186,231	229,689
Middlesex	494,022	602,023
Monmouth	411,236	456,017
Ocean	376,011	426,477
Hunterdon	74,123	89,982
Warren	60,188	75,832
Sussex	79,054	101,875
Mercer	225,885	275,680
Fairfield	544,922	665,912
New Haven	496,853	609,802
Total	**12,949,607**	**15,698,118**

Table 6-19. Modeled Number of Work-Based Subtours Compared to Number of Non-Home Based Tours in the LOCUS Data Set

County	NYBPM	LOCUS
Manhattan CBD	225,265	329,551
Manhattan Other	77,581	87,016
Queens	61,387	113,901
Bronx	34,637	73,387
Kings	81,538	144,421
Richmond	7,242	18,338
Nassau	51,765	90,436
Suffolk	52,076	103,411
Westchester	38,395	72,736
Rockland	9,295	21,812
Putnam	1,980	3,884
Orange	10,940	24,718
Dutchess	8,294	22,697
Bergen	32,182	67,931
Passaic	18,493	37,161
Hudson	25,630	49,432
Essex	36,307	64,649
Union	18,938	40,504
Morris	20,729	39,261
Somerset	11,927	23,978
Middlesex	33,976	72,222
Monmouth	22,187	36,848
Ocean	15,777	35,106
Hunterdon	3,236	9,273
Warren	3,099	4,755
Sussex	3,745	6,711
Mercer	16,390	38,449
Fairfield	35,642	70,567
New Haven	31,672	52,734
Total	**990,325**	**1,755,889**

As shown in Table 6-18, the number of modeled non-regular tours is about 20 percent lower than the number of regular tours in the LOCUS data set. Again, the difference is fairly consistent across the region. Table 6-19 indicates that the NYBPM simulates more than 40 percent fewer non-home based tours than there are in the LOCUS data set. This may be partly due to the non-home based tours simulated in the NYBPM being limited to work-based subtours. Another possibility is that LOCUS is capturing subtours that are not reported in the RHTS, from which the NYBPM components were estimated. These tours consist of non-home based trips, which are typically underreported in household travel surveys.

The percentages of tours by purpose in the LOCUS data are fairly consistent with those in the RHTS:

- NYMTC survey – 41% mandatory, 56% non-mandatory, 3% subtours
- NYMTC LOCUS – 37% mandatory, 56% non-mandatory, 8% subtours

The percentage of tours that are mandatory (regular) is a bit lower in the LOCUS data, and the percentage of tours that are subtours is higher in the LOCUS data, consistent with the reasons noted above.

County-to-county (home to regular location) patterns for regular tours were also compared, including comparisons to the expanded RHTS data set. The following observations are noteworthy:

- For regular tours, even when considering the lower number of regular tours in the LOCUS data set, the original model results were lower for many intra-county movements, notably for Queens, New York, Bergen, Hudson, Passaic, Richmond, Rockland, Somerset, and Union Counties. A few other county-to-county flows were clearly different, including most of the intra-New York City flows (excluding Manhattan) being higher in the model results and most of the New Jersey-Manhattan results being higher in the model results. Compared to the LOCUS data, the model produced fewer regular tours to and from Manhattan and more regular tours from the other New York City boroughs.

  This information was used (along with ACS data) in validation of the regular workplace location model. After calibration adjustments, the modeled distributions better matched both the LOCUS and ACS data, and other validation measures also improved, including screenline summaries and transit assignment summaries (presented in Section 6.2.1). Some notable differences between the model results and LOCUS data remained, including fewer modeled regular tour locations in the Manhattan CBD and more in the rest of Manhattan, more modeled regular tour locations for most inter-borough movements in New York City and fewer inter-county modeled regular tour locations for most New Jersey counties close to New York City (Bergen, Hudson, and Union—but not Essex).

- For non-regular tours, the distributions by home county are similar between LOCUS and the NYBPM.

- For non-home based tours, the distributions by home county are similar between LOCUS and the NYBPM. The percentage of subtours in Manhattan is higher in the NYBPM; this may reflect that the model better captures the relatively high number of work-based subtours in Manhattan.

6.3.2 Number and Orientation of Trips

The number of trips produced by CEMDAP by purpose are derived from the outputs of the following CEMDAP components:

- WSCH8a – Worker number of stops on commute tour
- WSCH8b – Worker number of stops on before work/after work/at-work tour
- NWSCH5 – Non-worker number of stops in a tour
- NWSCH6 – Non-worker number of stops following pick-up/drop-off

Table 6-20 shows the number of trips by county (with Manhattan segmented by CBD and other) from the model results and from the LOCUS data set. The number of modeled trips is 13 percent higher than the number of trips in the LOCUS data set. This difference is smaller than the percentage difference in total tours, which is about 20 percent. LOCUS probably includes many stops on tours that are unreported in the RHTS. Again, this seems to reflect the underreporting of non-home based trips in the survey—and perhaps some home-based trips for non-mandatory purposes. It is notable that the difference is greater in New York and Connecticut and smaller in New Jersey.

Table 6-20. Modeled Number of Trips Compared to Number of Trips in the LOCUS Data Set

County	NYBPM	LOCUS
Manhattan CBD	6,221,597	5,298,799
Manhattan Other	3,336,389	2,880,394
Queens	6,712,649	5,729,447
Bronx	3,581,192	3,250,468
Kings	7,189,175	6,299,910
Richmond	1,336,610	1,163,647
Nassau	4,741,363	4,053,604
Suffolk	4,937,284	4,493,894
Westchester	3,475,481	2,914,611
Rockland	836,445	926,114
Putnam	274,060	241,935
Orange	1,127,954	1,042,197
Dutchess	900,736	781,638
Bergen	2,901,614	2,870,629
Passaic	1,480,797	1,416,375
Hudson	1,923,153	1,874,294
Essex	2,780,287	2,263,848
Union	1,666,188	1,620,662
Morris	1,642,658	1,553,276
Somerset	1,168,393	973,898
Middlesex	2,855,380	2,431,049
Monmouth	2,046,677	1,959,177
Ocean	1,699,560	1,677,974
Hunterdon	325,337	341,400
Warren	263,464	255,666
Sussex	377,358	351,682
Mercer	1,252,709	1,099,264
Fairfield	3,370,467	2,811,951
New Haven	2,741,088	2,316,142
Total	**73,166,065**	**64,893,945**

County-to-county (origin-destination) patterns for trips simulated in the NYBPM were compared to the LOCUS data set. The following observations are noteworthy (even considering the differences in the number of trips between the NYBPM output and LOCUS):

- Modeled flows are noticeably higher in LOCUS for intra-county movements, especially for Kings, Queens, Nassau, Suffolk, Bergen, and Hudson Counties. This finding is consistent with the notion that short trips, which are mostly intra-county, may be underreported in the RHTS in spite of a GPS subsample and related adjustment to compensate for this underreporting.

- The model has noticeably fewer trips to and from Rockland and Hunterdon Counties than LOCUS.

- At a more aggregate district level, the origin-destination patterns compare more favorably (see Table 6-21) though LOCUS shows higher incidences of intra-district trips.

Table 6-21. Modeled Origin-Destination Trip Flows Compared to LOCUS

		LOCUS						
		1	2	3	4-5	6-7	8-10	11
1	Manhattan	69%	21%	2%	2%	4%	1%	1%
2	Other NYC	11%	83%	3%	2%	1%	0%	0%
3	Long Island	2%	6%	92%	0%	0%	0%	0%
4-5	Other NYS	3%	5%	0%	88%	2%	0%	2%
6-7	NJ near*	3%	1%	0%	1%	87%	7%	0%
8-10	Other NJ	1%	1%	0%	0%	7%	91%	0%
11	Connecticut	1%	1%	0%	2%	0%	0%	96%
	Region	13%	25%	13%	9%	15%	16%	8%

		Model Results						
		1	2	3	4-5	6-7	8-10	11
1	Manhattan	66%	20%	1%	4%	5%	1%	2%
2	Other NYC	10%	76%	9%	3%	1%	0%	1%
3	Long Island	2%	18%	80%	0%	0%	0%	0%
4-5	Other NYS	6%	7%	0%	77%	3%	2%	5%
6-7	NJ near*	5%	2%	0%	2%	80%	12%	0%
8-10	Other NJ	1%	1%	0%	1%	11%	86%	0%
11	Connecticut	3%	3%	0%	5%	0%	1%	88%
	Total	13%	26%	13%	9%	15%	16%	8%

6.3.3 Time of Day

The following comparisons were made by time period:

- Regular tours for LOCUS data, RHTS data, and model results (Table 6-22)
- Trips for LOCUS data and model results, and observed VMT from traffic counts and modeled VMT from the highway assignment (Table 6-23)

Table 6-22. Regular Tour Start Time Comparison by Time of Day

	Time Period			
	AM	MD	PM	NT
LOCUS	74%	12%	4%	10%
RHTS	81%	12%	4%	3%
Model	77%	14%	5%	5%

Table 6-23. Trip Time Comparison by Time of Day

	Time Period			
	AM	MD	PM	NT
LOCUS	22%	27%	30%	21%
Model Trips	25%	30%	25%	21%
Traffic Count VMT*	21%	31%	27%	21%
Model VMT*	23%	28%	28%	21%

* - On links with time of day counts

The following observations regarding these comparisons can be made:

- LOCUS shows a higher share of regular tours beginning in the nighttime period than the model results. The model results, however, show a higher share of nighttime regular tours than the RHTS data from which the time of day models were estimated. It is possible that the RHTS underreports nighttime regular tours. Note that as reported in Section 6.2.1, the model underestimates the magnitude of nighttime VMT by about five percent. The nighttime percentage of modeled daily VMT, however, matches the VMT share from traffic counts, as well as the nighttime share of trips in the LOCUS data set.

- The model shows a slightly higher percentage of regular tours beginning in the AM peak period than LOCUS. The model also shows a higher share of daily VMT in the AM peak than the traffic counts. However, as shown in Section 6.2.1, the magnitude of AM peak VMT is very close to the AM peak count VMT.

- As reported in Section 6.2.1, the magnitude of modeled mid-day VMT is about 13 percent lower than mid-day count VMT. However, Table 6-23 indicates that the mid-day percentage of modeled trips is higher than the mid-day percentage of trips in the LOCUS data set.

6.4 Sensitivity Testing

Sensitivity tests are designed to examine how vulnerable the model's forecasts are for certain plausible scenarios and how sensitive the model is to transportation level-of-service parameters and other factors that affect travel demand. While the sensitivity tests may indicate that the model is more or less sensitive than expected, the tests are not designed to say whether the travel model is "correct"; rather, they provide information about the overall behavior of the model.

As noted in the model validation plan, sensitivity testing involves revising key factors and observing the effects on forecasted travel. NYMTC defined the following sensitivity tests to be performed for the 2019 NYBPM:

- **Vehicle operating cost changes** – Increase vehicle operating costs by 50 percent
- **Transit network change** – Eliminate service for a section of a subway line

In addition to the 2019 tests, a 2050 forecast year scenario was run. A forecast year scenario can serve as a sensitivity test across a number of changes to the model's inputs. Specifically, the population and employment are increased (not uniformly across the region), and transportation system changes and improvements are assumed.

The sensitivity tests are summarized in a technical memorandum (Cambridge Systematics, Inc. and NYMTC, 2023). Overall, the model reacted reasonably to the changes involved in these tests. Auto travel decreased when vehicle operating costs (VOC) were increased, and some travel shifted from driving alone to carpooling. Many riders shifted to a nearby, parallel subway line when a section of the Lexington Avenue line was removed, and a small share shifted to other modes. The 2050 scenario showed an unusual and as yet unexplained lack of growth in travel between Long Island and Manhattan that will be further investigated as part of a model enhancement. . However, the NYBPM is useful for planning analyses at the regional level.

Key findings are summarized below.

6.4.1 Vehicle Operating Cost Changes

In this test, total vehicle operating costs were increased by 50 percent for all vehicle classes: passenger cars (including non-commercial pickups, minivans, and SUVs), light trucks/commercial vehicles, medium trucks, and heavy trucks. Costs remained unchanged for other modes, including transit, taxi, and TNC, as operating costs for these modes are not readily available.

The test was performed using the 2019 base year scenario as the baseline. Only costs that vary by tour/trip are included in vehicle operating costs (VOC). This is because the model uses the tradeoffs between the costs of individual tours/trips – such as the relative costs of using an auto or a transit mode – and so fixed costs that are incurred by vehicle users/owners regardless of whether they use auto for a particular trip (e.g., cost of auto insurance or drivers license), are not factors in these travel choices. Therefore, the VOCs used in the NYBPM and related sensitivity tests consider only fuel, maintenance, and tire costs associated with the vehicle operation.

The results of the VOC increase test are summarized below. However, more details are provided in the attached excel spreadsheet Highway Templates for 2019 test-R.xlsx

Regional VMT and VHT – In the higher VOC scenario, total daily VMT decreased by 1.3 percent. The VMT changes varied by vehicle type:

- Drive Alone -2.1%
- HOV 2 +0.7%
- HOV 3+ +2.7%
- All Auto -1.2%
- Taxi +0.6%
- Medium Truck -0.3%
- Heavy Truck -0.4%
- Other Comm. -0.2%
- All Truck -0.3%

The changes in VMT seem logical. Auto travel decreased due to the higher operating costs, and some travel shifted from driving alone to carpooling. Taxi fares were not changed in the higher VOC scenario, and so some auto travel shifted to taxi. The percentage VMT changes for trucks were lower than for autos since mode shifts to taxi or transit were not possible. The small changes reflect low decreases in truck trip length.

In general, the sensitivity of the model to changes in vehicle operating costs appears to be reasonable. Traffic volumes decreased by a modest amount, reflecting some decreases in trip lengths and shifts to modes where costs did not increase. The relatively small percentage changes reflect that people still need to travel for various purposes and that for some trips such as work commutes, it is necessary for travelers to absorb the cost increases. The average increase in cost for an auto trip in the higher VOC scenario is less than a dollar.

6.4.2 Transit Network Change Test

This test involved eliminating service on the Lexington Avenue subway line (4/5/6) between 59th Street and 86th Street. In this test, the NYBPM transit network was edited to remove service for the 4/5/6 train (Lexington Avenue Line) between 59th Street and 86th Street. Specifically, segments on the 4/5/6 line designated by the highlighted green stations/stops in Figure 1 were deleted in the route system file through the Delete Section of Route in TransCAD's Route System Toolbox. The AM peak period was the focus of the transit sensitivity test, as transit assignment is validated primarily for the AM peak (6:00 AM - 10:00 AM) due to availability of observed data.

The results of the transit network change test are summarized below.

- Subway boardings north of 60th Street in Manhattan significantly increased, most likely due to increased transfers.
- In the base scenario, total AM peak period ridership on the 4/5/6 is about 51,000 northbound and about 127,000 southbound between Grand Central-42nd Street and 59th Street & Lexington Avenue. Between 59th Street & Lexington Avenue and 86th Street & Lexington Avenue (the deleted section in the test scenario), the ridership is similar: 52,000 northbound and 124,000 southbound.
- When the segment is removed in the test scenario, ridership on the 4/5/6 south of 59th Street drops substantially, by about 75% northbound (about 38,000 riders) and by over 95% southbound (about 123,000 riders). That makes sense given that most of the riders south of 59th Street are also on the deleted segment.
- Roughly 20% of the riders who no longer take the 4/5/6 in the test scenario shift to the Q line. The northbound increase on the Q line (between 57th Street and 63rd Street is about 7,000; the southbound increase is about 27,000. Other riders probably shift to other north-south subway lines in Manhattan.
- About 2,000 trips apparently switch from transit in the base scenario to an auto mode, including taxi/TNC, in the test scenario. Others likely shift to non-motorized modes.

6.4.3 2050 Scenario

The main inputs to the 2050 scenario are the socioeconomic data, which is used to create the 2050 synthetic population using PopGen2, and the 2050 transportation network, including the highway and transit components. NYMTC's official 2050 forecast was used to develop the TAZ-level population, household, and employment estimates; about a 12 percent increase regionwide is indicated. The Series X highway project files were used to revise the network to reflect projects and changes expected to occur by 2050. Forecast transit routes were identified as routes that did not exist prior to 2019 but are expected to exist after 2019.

The complete model was run with feedback, using the 2050 zone data and PopGen2 results, 2050 highway network, and 2050 transit route system. Results were compared to the base year scenario results and checked for reasonableness. Selected comparisons are summarized below.

<u>Daily auto vehicle trips and VMT</u> – The 2050 scenario showed an increase of 12 percent regionwide over 2019. Increases of 10 to 13 percent were indicated for each subregion except within Manhattan, which showed about an eight percent increase. Manhattan's population is expected to grow at a lower rate than the rest of the region.

<u>Daily transit trips</u> – The 2050 scenario showed an increase of 17 percent regionwide over 2019. Increases of 10 to 20 percent were indicated for each subregion except Long Island, which showed only a one percent increase, the NJTPA core, which showed a 58 percent increase, and Connecticut, which showed a four percent increase.

It is notable that transit travel is estimated to grow much more rapidly in the NJTPA Core area than in the region as a whole. This appears to be related to specific transit improvements in this subregion. Transit demand from Long Island, however, is projected to grow very little by 2050. A deeper examination of this issue indicates that the NYBPM projects a decrease in work commuting between Long Island and Manhattan, and a corresponding increase in commuting within Nassau and Suffolk Counties. Since the Long Island Railroad, the dominant transit service on Long Island, primarily serves the Long Island to Manhattan commuting market, the modeled transit ridership shows little growth from 2019 to 2050.

We have not yet been able to find an explanation for this result, which is inconsistent with estimated growth in travel for other markets (for example, between the mid-Hudson subregion and Manhattan). The model was thoroughly checked, and no bugs were found that could cause this result. The socioeconomic data were also checked, as were the transit and highway networks on Long Island and the connections to Manhattan. This issue will be examined further by NYMTC.

7.0 References

Arizona State University and Cambridge Systematics (2020). *Synthetic Population Generation: An Application of PopGen*. Submitted to New York Metropolitan Transportation Council.

Cambridge Systematics, Inc. (2021). "NYBPM 2012 Update – Task 7 Summary." Memorandum to NYMTC.

Cambridge Systematics (2022). "LOCUS data summaries." Memorandum to NYMTC.

Cambridge Systematics, Inc. (2023a). "Model Re-Estimation Work Summary." Memorandum to NYMTC.

Cambridge Systematics, Inc. (2023b). *New York Best Practice Model 2019 Update: Users' Guide*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc., Arizona State University, University of Texas, Austin, Gallop Corporation, EA Harper Consulting, CDM Smith, Watchung Transportation, Florida International University, and David Rubin (2021). *New York Best Practice Model 2012 Update: Summary Report*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc. and EA Harper Consulting (2021a). *New York Best Practice Model 2019 Update: Model Validation Plan*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc. and EA Harper Consulting (2021b). *New York Best Practice Model 2019 Update: Transit Validation Database*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc. and EA Harper Consulting (2021c). *New York Best Practice Model 2019 Update: Highway Validation Database*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc., Gallop Corporation, and EA Harper Consulting (2023). *New York Best Practice Model 2019 Update: Model Validation*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc. and NYMTC (2023). "NYBPM 2019 Sensitivity Test Results." Memorandum to 2019 NYBPM project files.

Cambridge Systematics, Inc., University of Texas, Austin, and Arizona State University (2017). *New York Best Practice Model – Model Design Plan*. Prepared for the New York Metropolitan Transportation Council.

Cambridge Systematics, Inc., University of Texas, Austin, and Arizona State University (2018). *New York Best Practice Model – Model Estimation Report*. Prepared for the New York Metropolitan Transportation Council.

Dean, B. (2021). "Uber Statistics 2021: How Many People Ride with Uber?" Backlinko, https://backlinko.com/uber-users, Retrieved 11/18/21.