Uncertainty in Travel Forecasting: Exploratory Modeling and Analysis TMIP-EMAT: A Desk Reference

JULY 2021

U.S. Department of Transportation
Federal Highway Administration

TMIP
Better Methods. Better Outcomes.

Source: FHWA

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HEP-21-032		

4. Title and Subtitle	5. Report Date
Uncertainty in Travel Forecasting: Exploratory Modeling and Analysis TMIP-EMAT: A Desk Reference	July 2021
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
Jason Lemp, Thomas Rossi, Jeffrey Newman, Rachel Copperman	
9. Performing Organization Name And Address	**10. Work Unit No. (TRAIS)**
Cambridge Systematics, Inc. 101 Station Landing, Suite 410 Medford, MA 02155	
	11. Contract or Grant No.
	DTFH61-16-D00016L
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
Federal Highway Administration Office of Planning, Environment, and Realty 1200 New Jersey Avenue, SE Washington, DC 20590	January 2020 to June 2021
	14. Sponsoring Agency Code
	HEPP-30

15. Supplementary Notes

The project was managed by Task Manager for Federal Highway Administration, Sarah Sun, who provided detailed technical directions.

16. Abstract

Traditionally, travel forecasting models have been used to provide single point predictions. That is, a single future scenario is developed, and the model is applied to that scenario. This approach, however, ignores the deep uncertainty that exists in future land use, demographic, and transportation systems inputs, not to mention the uncertainty that exists in the model itself. More importantly, transportation policy decisions made on the basis of such model outputs may be misguided and ineffective. This report demonstrates and motivates the use of travel forecasting models in an exploratory manner that accounts for the inherent uncertainties of the future. Specifically, this report describes the approach that can be supported by a new planning and modeling tool: the Travel Model Improvement Program Exploratory Modeling and Analysis Tool (TMIP-EMAT) that has been developed to facilitate the use of exploratory techniques with travel forecasting models.

17. Key Words	18. Distribution Statement
TMIP-EMAT, Exploratory Modeling and Analysis, Travel Demand Modeling, Regional Models, Corridor Studies, uncertainty analysis, Monte Carlo simulation	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	142	N/A

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

ACS	American Community Survey
API	Application Programming Interface
ASC	Alternative-Specific Constant
ATRI	American Transportation Research Institute
ATUS	American Time Use Survey
AV	Automomous Vehicle
BLS	Bureau of Labor Statistics
BRT	Bus Rapid Transit
CART	Classification and Regression Trees
CAV	Connected and Automated Vehicle
CDF	Cumulative Distribution Function
DTA	Dynamic Traffic Assignment
EMA	Exploratory Modeling and Analysis
EMAT	Exploratory Modeling and Analysis Tool
FAST	Fixing America's Surface Transportation
FHWA	Federal Highway Administration
FTA	Federal Transit Administration
GBNRTC	Greater Buffalo Niagara Regional Transportation Council
GPR	Gaussian Process Regression
GPS	Global Positioning System
HH	Household
ID	Identifcation
IVTT	In-vehicle travel time

LBS	Location-based services
LHS	Latin HyperCube Sampling
LOS	Level of Service
MaaS	Mobility as a Service
MAP-21	Moving Ahead for Progress in the 21st Century
MIT	Massachusetts Institute of Technology
MPO	Metropolitan Planning Organization
NCHRP	National Cooperative Highway Research Program
NHTS	National Highway Traffic Safety Administration
O-D	Origin-Destination
ODOT	Oregon Department of Transportation
ODME	Origin-Destination Matrix Estimation
PM	Performance Metric
PRIM	Patient Rule Induction Method
RSPM	Regional Strategic Planning Model
RTP	Regional Transportation Plan
SPLOM	Scatter Plot Matrix
STOPS	Simplified Trips-on-Project Software
TAZ	Traffic Analysis Zone
TDM	Travel Demand Model
TMIP	Travel Model Improvement Program
TNC	Transportation Network Company
TOD	Transit-Oriented Development
TRB	Transportation Research Board
TSP	Transit Signal Priority

V/C	Volume-to-Capacity
VHT	Vehicle-Hours of Travel
VMT	Vehicle-Miles of Travel
VTPI	Victoria Transport Policy Institute

Executive Summary

Over the past decades, transportation agencies have used predictive planning with a predetermined plan within a specific timeframe. Travel models are designed to provide point-estimate forecasts. This means that the model inputs and other assumption used for developing the forecasts are assumed to be deterministic and unvarying (i.e., no uncertainty exists in the assumptions). However, by its very nature, long range transportation planning is uncertain. In a more general way, the planning community faces problems that involve uncertainty about the future all the time. This uncertainty comes from uncertainty in how systems work, uncertainty in how inputs to a system will change in the future, uncertainty about what are the important features of a system on which to focus, and a number of other uncertainties. It is imperative for planners to recognize this uncertainty and utilize different approaches and tools to support examining and planning for systems with uncertainty.

A new approach, exploratory modeling and analysis (EMA), embraces the examination of uncertainty by explicitly treating computational experiments (i.e., models) as a set of assumptions and hypotheses and aims to explore the impacts of the assumptions on the analysis of interest. Exploratory modeling approaches are preferred when critical information is unavailable. EMA has been used to better understand systems with deep uncertainty by calibrating models that explain the system where some inputs to the system have deep uncertainty associated with them, there are various policies or levers available to a decisionmaker to affect the system, and there are various outputs of the system which are of interest.

The FHWA Exploratory Modeling and Simulation Study focused on exploratory, rather than predictive, modeling of future transportation systems, with particular attention to the impacts of new technology. One outcome of this project was the development of the Travel Model Improvement Program Exploratory Modeling and Analysis Tool (TMIP-EMAT). TMIP-EMAT was developed to help agencies manage uncertainties by illuminating interactions between transportation supply and demand on the urban surface transportation system through exploratory modeling and simulation, provide insights of potential, possible, plausible, probable, or preferred futures, and support robust regional transportation planning decisionmaking incorporating principles of risk management.

TMIP-EMAT can be integrated with existing travel modeling tools to facilitate the application of those models in an exploratory, rather than predictive, manner. It builds upon evolving sensitivity and risk analysis approaches using travel models to forecast uncertainty; it can be used to understand the effects of future mobility impacts on travel patterns, and it incorporates exploratory-type visualizers and optimization search tools to present and analyze the results.

TMIP-EMAT can produce results for performing risk analysis and exploratory analysis. TMIP-EMAT is not a transportation model in and of itself; it is a utility tool that enables an analyst to use the region’s transportation model for exploratory analyses.

TMIP-EMAT has been successfully deployed by planning agencies in the U.S. The Oregon Department of Transportation used TMIP-EMAT to support analysis of future technologies where little or no observed data exist to estimate and validate models, using the Southern Oregon Activity-Based Model. The Greater Buffalo Niagara Regional Transportation Council used TMIP-EMAT to evaluate infrastructure investments along a specific corridor in the region, using their four-step travel model.

Section I

Introduction

1.0 Introduction

Over the past decades, transportation agencies have used predictive planning with a predetermined plan within a specific timeframe. Travel models are designed to provide point-estimate forecasts (Transportation Research Board, 2007). This means that the model inputs and other assumption used for developing the forecasts are assumed to be deterministic and unvarying (i.e., no uncertainty exists in the assumptions). However, by its very nature, long range transportation planning is uncertain. In a more general way, the planning community faces problems that involve uncertainty about the future all the time. This uncertainty comes from uncertainty in how systems work, uncertainty in how inputs to a system will change in the future, uncertainty about what are the important features of a system on which to focus, and a number of other uncertainties. It is imperative for planners to recognize this uncertainty and utilize different approaches and tools to support examining and planning for systems with uncertainty.

Emerging connected and autonomous vehicle (CAV) technology, new mobility services, and changing travel patterns will potentially have significant unpredictable impacts on future surface transportation operations and travel demand. Current practice is to build theoretical extensions of existing models based on stated-preference surveys, technological trends, and expert opinion (Childress et al., 2015; Gucwa, 2014; Vovsha, 2017; Cambridge Systematics, Inc., 2016). However, it is extremely challenging to use travel models to predict what effects these disruptive technologies will have on travel behavior and, in turn, on the surface transportation system, given the lack of substantive data at the age of rapid technological innovations as we are now. Uncertainty associated with significant unpredictable impacts has been termed deep uncertainty (Lempert et al., 2003). Under the condition of deep uncertainty, one of the primary challenge for transportation agencies is to understand the scope of impacts and interactions and the implications on traditional planning strategies. This deep uncertainty associated with potential impacts due to changing travel patterns and emerging technologies calls for a more comprehensive and exploratory approach to planning future mobility. Therefore, a new approach to modeling and simulation is needed. One such approach, exploratory modeling and analysis (EMA), embraces the examination of uncertainty by explicitly treating computational experiments (i.e., models) as a set of assumptions and hypotheses and aims to explore the impacts of the assumptions on the analysis of interest.

> *"Deep uncertainty exists when analysts do not know, or the parties to a decision cannot agree on, (1) the appropriate models to describe the interactions among a system's variables, (2) the probability distributions to represent uncertainty about key variables and parameters in the models, and/or (3) how to value the desirability of alternative outcomes."—Lempert et al., 2003*

Exploratory modeling approaches are preferred when critical information is unavailable (Bankes, 1993). EMA has been used to better understand systems with deep uncertainty by calibrating models that **explain the system**, where:

- Some inputs to the system have deep uncertainty associated with them.
- There are various policies or levers available to a decisionmaker to affect the system.
- There are various outputs of the system which are of interest.

This differs from treating the model as a predictive tool that is an accurate surrogate to the real world (Bankes, 2003).

The Federal Highway Administration (FHWA) Exploratory Modeling and Simulation Study focused on exploratory, rather than predictive, modeling of future transportation systems with particular attention to the impacts of new technology. One outcome of this project was the development of the Travel Model Improvement Program Exploratory Modeling and Analysis Tool (TMIP-EMAT). TMIP-EMAT was developed to help agencies manage uncertainties by illuminating interactions between transportation supply and demand on the urban surface transportation system through exploratory modeling and simulation; provide insights of potential, possible, plausible, probable, or preferred futures; and support robust regional transportation planning decisionmaking incorporating principles of risk management.

1.1 Travel Model Improvement Program Exploratory Modeling and Analysis Tool

A Different Way to Address Transportation Planning Questions

There are many types of analyses where planners use modeling tools to provide quantitative information that can help make planning, investment, and policy decisions, as discussed in detail in section 2.2. This generally involves creating an analysis scenario and comparing results to those reflecting a base condition. The analysis scenario may represent, for example, the execution of a specific investment, such as building or improving transportation infrastructure, the implementation of specific policy or pricing decision, a set of actions related to a long-range transportation plan, or proposed land use changes reflecting a specific development or land use plan.

Typically, travel models are run for a single analysis scenario, created using the best guesses for various assumptions about growth forecasts; changes in factors that influence travel behavior (e.g., fuel costs); and the effects of technological changes that affect travel demand and supply (e.g., shared mobility providers). The core model itself contains a variety of assumptions about how travelers react to changes in the transportation environment, usually expressed as parameters calibrated to reflect observed travel behavior.

So by themselves, travel models provide a snapshot of what is expected to happen under a specific set of assumptions, which provides limited information about the impacts of planning, policy, and investment decisions. Even if multiple scenarios are able to be run (e.g., low-, medium-, and high-growth scenarios), only a few data points are provided, with no context to examine the likelihood of specific performance measure results.

TMIP-EMAT offers a different way to address transportation questions. It can be used to systematically explore uncertainties in input variables and model parameters and the impacts that those uncertainties have on performance metrics. It is useful for examining model forecasts as a **range of model outcomes** rather than a single outcome, and it provides a mechanism for defining uncertainties and visualizing outputs.

TMIP-EMAT offers a different way to address transportation questions. It can be used to systematically explore uncertainties in input variables and model parameters and the impacts that those uncertainties have on performance metrics.

TMIP-EMAT can be integrated with existing travel modeling tools to facilitate the application of those models in an exploratory, rather than predictive, manner. It builds upon the evolving sensitivity and risk analysis approaches using travel models to forecast uncertainty, can be used to understand the effects of future mobility impacts on travel patterns, and incorporates exploratory-type visualizers and optimization search tools to present and analyze the results.

TMIP-EMAT was developed by utilizing and building-upon the EMA Workbench, which is a tool for performing exploratory modeling and analysis (Kwakkel, 2017).The EMA Workbench includes features to define the uncertainty space (range of possible scenarios), define the decision space (range of policy/treatment/project levers), and run the simulations to populate the outcome space (performance measure products for each experiment). A diagram of the uncertainty, decision, and outcome spaces consistent with the EMA Workbench is presented in figure 1. The EMA Workbench also includes features to conduct different types of analyses like scenario discovery and robust search, which are discussed in more detail in section 3.4.

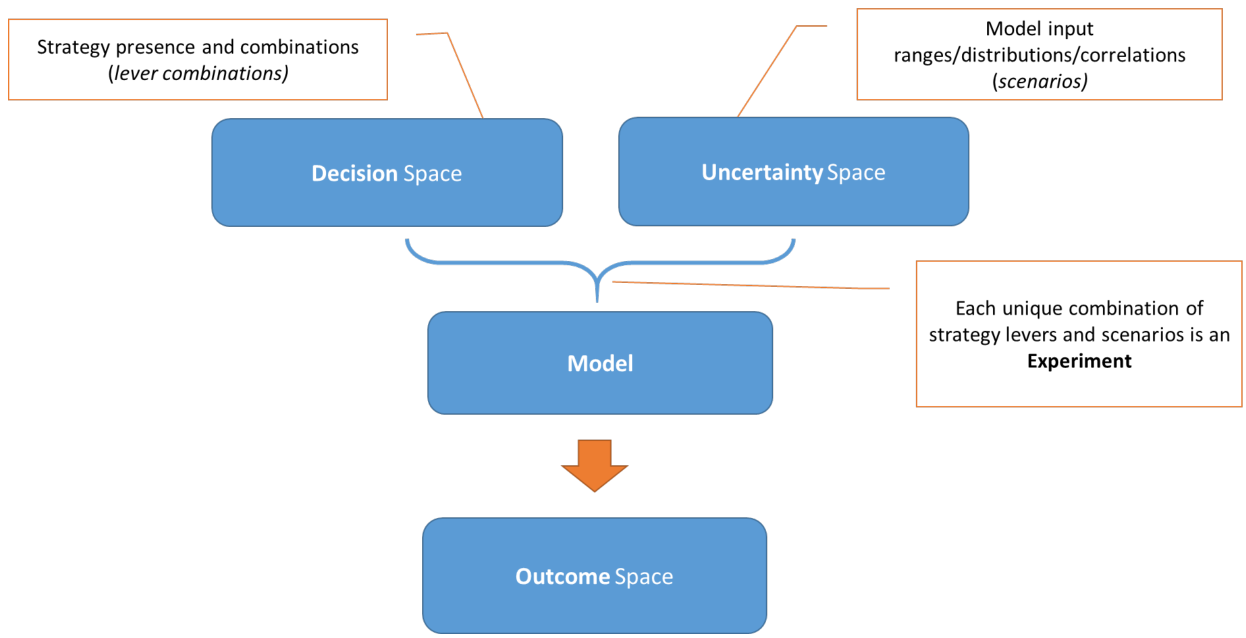


Figure 1. Flow chart. A diagram of uncertainty, decision, and outcome spaces.

(Source: Federal Highway Administration.)

TMIP-EMAT also builds upon the framework developed by RAND (Lempert, 2019) for understanding deep uncertainty, which is based on the following relationship:

$$M = R\,(X, L)$$

Here, *M* are the metrics that we are interested in understanding, or outputs of the model (i.e., performance measures), *X* are the set of uncertainties; *L* are the set of policy levers; and *R* are the relationship between *X*, *L*, and *M*, which in our case is the travel model.

In addition to these types of uncertainties, RAND's terminology (Lempert, 2019) talks about the presence of deep uncertainty, which is when the relationships (R) from the equation above cannot be agreed upon. In our case, the relationships refer to the travel model itself and, in particular, the structural elements of the model. This is of particular importance as new technologies have emerged over the last several years. While the travel modeling community has a lot of experience (collectively) in developing effective ways to forecast traditional modes, such as nonautonomous vehicles and transit, much less is known about forecasting the impacts of e-commerce or CAVs.TMIP-EMAT is designed to integrate with a **core model**, which is any application-/region-specific transportation model. The core model takes a collection of inputs and generates one or more outputs, or "performance measures," of interest. Examples of a core model include, but are not necessarily limited to, the following:

- Regional or statewide travel model.
- Activity-based travel model.
- Trip-based travel model.

- Sketch planning or spreadsheet model.
- Microsimulation model.
- Corridor-level or mode-specific travel model.

The typical user of TMIP-EMAT is envisioned to be an analyst/planner/modeler who is familiar with the capabilities, and limitations, of the core model. The three major steps to working with TMIP-EMAT are defined in figure 2.

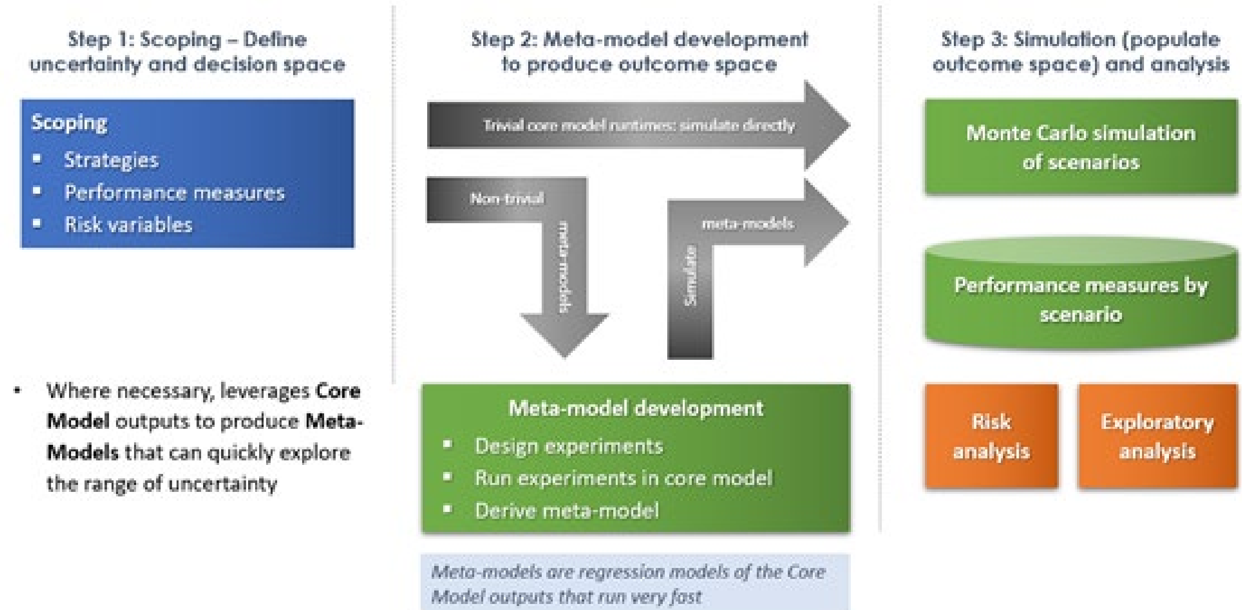


Figure 2. Diagram. Travel Model Improvement Program—Exploratory Modeling and Analysis Tool process flow.

(Source: Federal Highway Administration.)

During **Scoping**, the user identifies the strategies to be analyzed and the measures to evaluate. The user also considers uncertainties that may affect the outcome of the measures and can be represented by parameters of or inputs to the core model. For most types of analysis and cases where the core model run time is nontrivial, TMIP-EMAT utilizes **metamodels**, which are regression models that estimate the core model outputs. The metamodels run very quickly (milliseconds) and, thus, can be used to produce measures comprehensively across the uncertainty space. Where the core model run time is trivial, as in a sketch model, TMIP-EMAT utilizes the core model directly.

To populate the outcome space, TMIP-EMAT utilizes Monte Carlo **simulation** to sample across the uncertainty distributions. For each simulation run, the associated value of each uncertainty is set in the model (metamodel or core model), and the measure estimate is recorded in a database. The user can then examine and analyze the effects of their strategy levers on measures of interest under various uncertainties utilizing various analysis approaches.

A Different Way to Utilize Core Model Outputs

TMIP-EMAT does not diminish the information that is provided by travel models, since the core model is an important part of the TMIP-EMAT process. It provides valuable additional information about the variation associated with the values of performance measures, rather than single point estimates, and about the potential of meeting various planning goals.

TMIP-EMAT can produce results for performing risk analysis and exploratory analysis. The distinction between the two types of analyses is predicated on the user's desire to analyze probabilities and produce performance measures with confidence intervals, as in a risk analysis, or to identify and describe the existence and extremes of best and worst case scenarios as in an exploratory analysis.

Risk Analysis

A risk analysis is dependent on the uncertainty variable distribution and presents the measure results in terms of the probability of each outcome (e.g., using a cumulative distribution function (CDF)), as shown in figure 3. The user can then assess the most likely value and overall range for each measure. In this situation, the user also may conduct a series of statistical significance tests on the performance measures across different strategies. A risk analysis may examine outputs, such as:

- The range and probability of occurrence of performance measure values.
- Cumulative distribution function plots, which provide a visual perspective of the percentiles.
- The relative importance of an uncertainty variable contributing to the variability in the outcome.

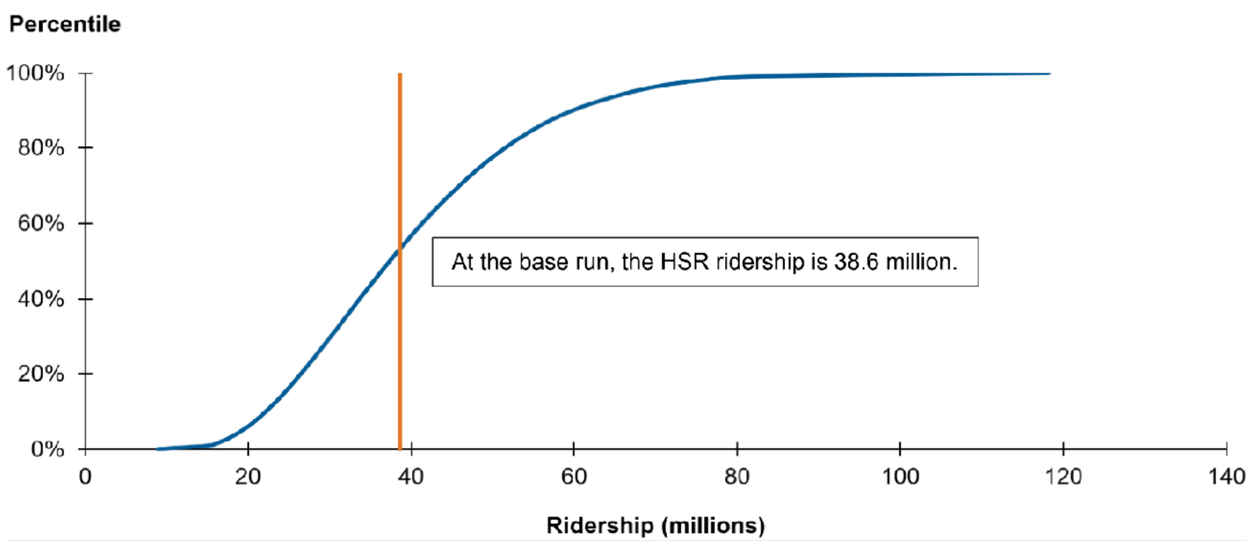


Figure 3. Chart. Example risk analysis outputs.

(Source: California High-Speed Rail 2020 Business Plan Ridership and Revenue Risk Analysis.)

Exploratory Analysis

The purpose of an exploratory analysis is not to describe the probability or likelihood of a certain outcome, but rather to identify scenario conditions and thresholds where different outcomes occur. An exploratory analysis should:

- Identify the implications of uncertainty variables on the exploration space.
- Help the user tell the story of how the strategies may fare under different conditions.
- Identify what can be done to mitigate the worst-case outcomes and encourage the best-case outcomes.
- Provide an analysis of sensitivities and vulnerabilities.

Below are three different approaches to an exploratory analysis, although practice may vary:

1. Understanding the importance or success of strategies given certain conditions of the performance measures.
2. Understanding the significance of the uncertainty variable on the impact of a performance measure (i.e., understanding what the area of concern is).
3. Understanding the importance or success of strategies given certain conditions of the uncertainty variables.

1.2 Organization of the Report

The remainder of this report is organized as follows. Section II focuses on core model selection. Within this section, chapter 2.0 describes the types of modeling and analysis tools available to planners and how TMIP-EMAT can be used to enhance the use of particular tools for specific planning applications. Those readers who are experts in core models, such as travel models, may skip this section if they prefer. Section III presents detail information on TMIP-EMAT and its application. Within section III, the process of executing an exploratory modeling analysis using TMIP-EMAT is laid out in chapter 3.0. Chapter 4.0 summarizes two case studies using TMIP-EMAT that can help guide analysts and planners in using the tool in a variety of environments. Chapter 5.0 provides conclusing remarks.

Section II

Core Model Selection

2.0 Choosing the Right Model

As discussed in section 1.1, TMIP-EMAT takes advantage of an existing core model, often the main travel model used in a planning region. At a time when there is more uncertainty than ever about how people will be traveling both short term and longer term, it is the right time to identify how to make best use of the available tools to ensure that all of the right data needed to best analyze planning issues is made available, and to make sure that the modeling tools themselves are as good as they can be. This chapter discusses the importance of choosing the right model or modeling tool for planning purposes. This is important in ensuring that the TMIP-EMAT tool is based on the best available core model (and is important in any planning application analysis regardless of whether TMIP-EMAT is being used).

2.1 Defining Planning Analysis Goals

The set of tools a planner may choose depends greatly on the values of the community and the conditions in the planning region, which in turn help determine the planning analysis goals. Before designing and developing a new model or adopting a model for a particular project or application, it is important to understand the planning analysis objectives.

The factors and conditions that affect planning goals can vary widely; depending on the region, they may include the following:

- The size of the region (in terms of population and land area).
- Characteristics of nearby areas, such as other metropolitan areas or States, that may exchange substantial amounts of travel with the planning region.
- How quickly the region is growing.
- Economic characteristics (such as dominant industry types, presence of major attractions, universities, etc.).
- Demographic characteristics (for example, age distribution and income levels).
- Density of the highway system and roadway types.
- Level and types of transit service provided.
- Competition among auto, transit, and active transportation modes.
- Parking supply in areas with high levels of travel demand.
- Levels of truck volumes and freight movement, both within and passing through the region.
- Presence of toll roads and managed lanes, and the potential for adding them in the future.
- Penetration of new transportation modes such as transportation network companies (TNC) and shared mobility, which may vary throughout the region.
- Various types of mobility options the community would like to examine or consider.
- Institutional and political factors, such as pricing and taxes on fuel, parking, and public transportation, land use regulations, and school district boundaries.

- Levels of congestion and travel peaking characteristics.
- Whether there are key seasonal and/or day of week considerations with respect to travel patterns.
- The resources available for planning, modeling, and obtaining data.

The factors and conditions listed above are not meant to be exhaustive, but they provide a starting point for thinking about the planning analysis needs of a region. They also provide the backdrop for defining the basic modeling requirements, which may include a set of performance measures that the model needs to be able to forecast, as well as the key policy questions for which the model will be used. Some common performance measures that are often output by models include the following:

- Regionwide vehicle-miles of travel (VMT), vehicle-hours of travel (VHT), delay, and hours of congestion.
- Corridor-level metrics (such as travel times, volumes, delay, reliability).
- Air quality analytics and fuel consumption.
- Transit use (such as ridership, revenue miles, passenger-hours).
- Toll and managed lane volumes and revenues.
- Travel time reliability metrics.
- Levels of active transportation.
- Transportation metrics for different groups of travelers (such as equity groups).
- Truck and freight-related movements.

Smaller regions where some of these factors are not important may prefer simpler models, which could include sketch planning tools or relatively simple travel models. In such cases, the types of policies that the model can examine will be more limited. Larger areas with significant congestion, complex multimodal transportation systems, roadway pricing, and higher concentrations of nonauto travel may require more sophisticated modeling tools. Rapidly growing areas need models that are sensitive to the land use and demographic changes that are occurring and are expected to occur in the future. As a result, it is critical that the model be designed to answer the questions at hand.

2.2 Tailoring the Model to Planning Analysis Goals

As the Transportation Research Board (TRB) Special Report 288 (Transportation Research Board, 2007) noted, there is no one single “correct” approach for all applications or all MPOs. “Travel forecasting tools developed and used by an MPO should be appropriate for the nature of the questions being posed by its constituent jurisdictions and the types of analysis being conducted.” It also is important that any model be designed to represent the transportation system supply and demand conditions in a satisfactory way, and that it be able to address the planning goals that it supports. The report *How-to: Think About Model Design for Your Region*, prepared for FHWA (Bernardin, 2018), discusses many of the practical and theoretical considerations relevant to the choice between activity-based model, trip-based model, or hybrid model, including practical issues like data availability, calibration, cost, and model runtimes, as well as theoretical issues like aggregation and trip chaining. All of these issues are important to deciding among the main travel model types.

The next section discusses model design considerations for answering tradtional planning questions while the subsequent section discusses model design for emerging mobility.

Traditional Questions Posed to Travel Models

The types of questions that a travel model may be asked to address are varied and include everything from traditional highway projects to emerging mobility issues. The model design should reflect these needs. *How-to: Think About Model Design for Your Region* also discusses the model design issues for many typical types of questions asked of travel models in the context of deciding between the main model types (activity-based, trip-based, and hybrid models). This section summarizes from the report some of the nuances involved in addressing specific types of questions within a given model among the main model design issues.

Traditional Highway Projects

These projects generally need to be able to answer the question of how highway capacity impacts congestion on the roadway network. Such projects include not only entirely new roadway facilities, but also capacity expansion to existing facilities. These types of projects are what traditional travel models were originally designed to assess, and generally can be evaluated with any network-based travel model.

Air Quality Conformity Analysis

These analyses are required by the Clean Air Act for large metropolitan regions and require agencies to generate estimates of emissions levels in the region. Emissions estimates are based on VMT estimates by roadway type and speed, which are standard outputs of network-based travel models and, thus, are easily assessed with such models.

Accurate air quality estimates also require a precise estimate of the breakdown of vehicles by type. While most travel models do not directly model vehicle type of passenger travel, models often include truck components, which explicitly estimate travel by trucks that are an important component of emissions estimates.

Equity Analysis

These types of analyses rely on segmentation of key outputs from the model on the basis of demographic or place of residence characteristics. As the report points out, segmentation of outputs by place of residence requires that the model be able to attribute each trip to a resident, which requires a more advanced model, such as a hybrid or activity-based model, though segmentation of home-based trips by place of residence can be done even with a trip-based model.

Likewise, segmentation of model outputs by detailed demographic features requires that the model segment the population by such features, which typically is only possible with activity-based models since such models simulate individual travel decisions of each resident. Because these types of analyses require specialized model features, it is important to consider whether these analytical capabilities are important in the model design.

Traffic Impacts

These types of analyses try to answer the question of how new developments will impact traffic patterns. Like traditional highway projects, traffic impacts can be assessed with any network-based travel model, though finer-grained network and zone resolution is better for these types of analyses. Unlike traditional highway projects, these analyses often require edits to the model to enhance the resolution in and around the new developments. The report also notes that hybrid and activity-based models may be better suited for these types of analyses due to better treatment of internal capture, though in the case of activity-based models, these benefits are largely outweighed by the presence of simulation variation.

Highway Pricing

These analyses deal with pricing (including tolling, parking fees, cordon fees, mileage-based fees, etc.) and how pricing impacts congestion and mode choices and generates revenue. There is a number of key factors that exist (or can exist) related to the types of pricing studies that can be analyzed with a travel model. In particular, the model should be capable of analyzing the key characteristics of the potential market for the priced facility, including those associated with willingness-to-pay (e.g., income and trip purpose); visitor versus resident travel; and trucks. Other model design considerations also are important for highway pricing. For instance, if the model is to be used to assess the impacts of time variable tolls, then the time-of-day component of the travel model must reflect the time period variation of the tolling scheme.

Peak Spreading

Peak spreading travel phenomena can occur wherein travelers shift the timing of their travel to peak shoulders or nonpeak periods to avoid the heaviest congestion levels during the peak hour. These types of analyses usually are only required for models that represent large regions where heavy congestion exists in the peak periods.

Travel Demand Management

These strategies include various types of policies aimed at impacting travel demand in different ways, including, but not limited to, the following:

- Telecommuting.
- Alternative work schedules.
- Transit subsidies like employer-provided transit passes.
- Ride-home guarantees for public transit commuters.

- Carpooling incentives.
- Parking management strategies.

Most travel models struggle to evaluate policies such as these because they typically require a number of assumptions about how travel behavior will be impacted. For instance, the efficacy of a telecommuting incentive in terms of how many people telecommute as a result of a policy cannot be evaluated with a travel model. Assumptions can be made about the number of telecommuters and, given those assumptions, the effect on travel patterns can be evaluated. However, in that case, an assumption is still needed about how telecommuting will affect nonwork traveler for those telecommuters, which may or may not be valid.

Transit Investments

These analyses address the question of how changes in transit services impact transit ridership. The report suggests that there is little empirical evidence suggesting that activity-based or hybrid models perform better than trip-based models in terms of transit investments, and that the Federal Transit Administration (FTA) prefers use of its Simplified Trips-on-Project Software (STOPS) forecasting tool for New Starts project forecasts.

Bicycle/Pedestrian Planning

The key question for bicycle and pedestrian planning analyses deals with how infrastructure and urban design impact bicycle and pedestrian travel. The key factor in addressing these types of analyses in model design deals with the level of spatial aggregation in the model. Because bicycle and especially pedestrian trips tend to be shorter than auto and transit trips, the spatial resolution of the model needs to be higher to accurately predict how policy will impact bicycle and pedestrian travel patterns. The greater spatial resolution means that a greater amount of data and effort is needed as input to the model as it relates to zone structure, land use, and networks.

Land Use Planning

These analyses deal with the impact that land use policy and urban design have on travel. As the report points out, different travel models have differing capabilities to analyze land use policy. Many factors play a role in determining whether these types of analyses may be important to a region, including the role of changing demographics in the region, economic conditions, and the importance of promoting nonauto modes, where micro-level accessibility metrics can be important.

Model Design for Emerging Mobility

The previous subsection deals specifically with traditional questions asked of travel models for which methods have become standardized and/or situations when, typically, robust data is available for model estimation and calibration. However, model design in the context of emerging mobility considerations is different in this respect and applies to a variety of pressing transportation planning questions. Addressing these questions is made more difficult given the relative lack of data available to support the estimation and calibration of model components that specifically address these issues. In this section, we address some of the key model design considerations required to analyze emerging policy and technology questions; and present “use examples” that take a deeper dive into how specifically four topics, currently of great interest to planners, can be and have been addressed in travel demand models. The four topics—CAVs, emerging modes, e-commerce, and work from home—are related to transformational changes in transportation supply and demand that have emerged recently and are still changing, or are expected to emerge soon.

In most of these cases, the data available to assess policy sensitivity and/or quantify the impacts of emerging technology are lacking, either because the issue is too new or does not yet exist. Examples of emerging mobility questions range from the impacts of e-commerce on shopping travel to the adoption of CAVs and their impact on travel patterns. In order to address these issues using a travel model, the model should be designed in such a way that it is sensitive to these issues. This section discusses the emerging mobility trends.

Connected and Autonomous Vehicles

CAVs have been on the radar of transportation professionals for several years. They promise to enhance the efficiency of transportation system while offering improved accessibility. Obviously, no observed data yet exists to assess the impacts of CAVs on travel behavior or transportation supply. Nonetheless, many researchers have begun to anticipate the types of effects CAVs may have and develop approaches for handling CAVs in travel models. One example of such research is the National Cooperative Highway Research Program (NCHRP) Report 896: Updating Regional Transportation Planning and Modeling Tools to Address Impacts of Connected and Automated Vehicles (Zmud et al., 2018).

Transportation Network Companies

The emergence of TNCs has had a big impact on how people travel in many regions. TNCs provide taxi-like service (but at a lower cost), interface with customers using a user-friendly app, and provide more widespread coverage, making them more desirable to a larger portion of the population. Shared mobility options, such as vanpools, offer even lower costs, though with the disbenefit of shared service and lower levels of service.

TNC usage has enjoyed a significant upward trend over the last several years prior to the COVID-19 pandemic. Between 2015 and 2018, the Pew Research Center found that the number of individuals that has used TNC services rose from 15 percent to 36 percent (a 34-percent yearly increase), while the number of individuals that has not heard of such services dropped from 33 percent to 3 percent over the same period (Jiang, 2019). More recently, the number of daily active users of specific TNC services, such as Uber and Lyft, has increased more gradually. Uber use increased from about 1.1 million daily users of U.S. Android phones to 1.2 million from the beginning of 2018 to the beginning of 2020, while Lyft use increased from about 550,000 daily users in 2018 to 700,000 in 2020, representing a total yearly increase of only 7 percent (Thibault, 2020). Neither of these measures is a perfect proxy for actual TNC use at a trip level; they suggest that, while growth in TNC usage continues, it is slowing overall.

Studies have shown that travel patterns of TNC trips are vastly different from those of trips using other modes of travel. In particular, the demographics of TNC users are unique (they tend to be younger, have higher education levels, higher income, and have fewer household vehicles) (Dias et al., 2017); the areas where TNC activity takes place tend to be centrally located in urban areas; and trip purpose, time of day, and day of week characteristics differ from other trip types. Roy et al. (2020) provide an extensive review of the various studies examining TNCs and the data that currently is available.

E-Scooters and Bike Share

Like TNCs, e-scooters and bike share programs have emerged over the last several years, impacting travel patterns in a number of ways. These technologies have provided a new transportation mode option for many trips, especially those in urban areas where e-scooter and bike share availability is highest. Moreover, they are providing new options for the first and last mile of transit journeys in many areas.

E-commerce

E-commerce has been steadily on the rise for a number of years, even prior to the COVID-19 pandemic. It is widely accepted that shopping online has resulted in a reduction in shopping trips (Maat and Konings, 2018); and this has even been shown using survey data. In addition to a drop in shopping trips, e-commerce also results in a corresponding rise in urban freight and parcel delivery activity. And in recent years, parcel delivery vehicles have been trending toward a larger number of smaller vehicles to handle the demand for these services. An enhanced travel model can be used to address these questions.

Telecommuting

While telecommuting has been around for decades, advances in communication technology and data systems have allowed more workers the freedom to work from home in recent years. Telecommuting has long been studied as a transportation demand management concept to reduce congestion levels on roadways.

Based on data from the Bureau of Labor Statistics (BLS), about seven percent of workers in the U.S. had the access to work from home as a benefit in 2019, which was up from about five percent in 2010 (DeSilver, 2020). However, with the COVID-19 pandemic of 2020, a major shift in telecommuting trends began as many workers were forced from their normal places of work. It still is unclear to what extent the trends that emerged during the pandemic will persist long term, or will return to their prepandemic levels in the future.

As noted above, telecommuting levels saw a sharp rise during the COVID-19 pandemic of 2020, and it is yet unclear how those changes will persist in the future. As such, a great deal of uncertainty exists around telecommuting levels in the future. It is likely that as many companies shifted policies and made investments to allow for more telecommuting during the pandemic, the shift toward telecommuting that occurred during the pandemic will not totally disappear.

2.3 Model Design Examples for Emerging Mobility

This section presents "use examples" that take a deeper dive into how specifically four topics, currently of great interest to planners, can be and have been addressed in travel demand models. The four topics—CAV, emerging modes, e-commerce, and work from home—are related to transformational changes in transportation supply and demand that have emerged recently and are still changing, or are expected to emerge soon.

Connected and Automatic Vehicles

As previously described, CAVs represent a technology that has not yet been realized. CAVs have the potential to cause major changes to the transportation system, how people use it, and even how people choose to live. Such changes include impacts on congestion, how far people are willing to drive, how transit is used, where people live and work, automobile ownership, and how young people travel, among others. The objectives of analyzing this technology are to use information we understand about travel behavior to inform sensitivities we incorporate in the travel model and forecast potential future impacts of CAVs.

Key Questions

There are a number of key questions related to CAVs from a policy perspective, including the following:

- *Will CAVs be able to utilize roadway space more efficiently than manually operated vehicles?*

  One of the great promises of CAVs is that they will communicate with one another and be able to form platoons that allow for higher speeds and less congestion. This is very likely for a fully automated and connected fleet of vehicles. However, a great deal of uncertainty remains about roadway utilization when the vehicle fleet is mixed with some CAVs and other manually operated vehicles, as well as the extent to which CAVs will be able to be more efficient users of roadway space.

- *How will CAVs impact how individuals travel?*

  This question is critical and has a number of subquestions around vehicle ownership, travelers' values of time, travel time budgets, auto occupancies, and parking policy, among many others.

- *How will CAVs be adopted over time?*

  Truly autonomous vehicles have not yet been realized. When CAVs first become available on the marketplace, they will likely be relatively expensive and, thus, cater to higher income consumers and others willing to pay the higher prices. As CAV technology matures, prices will drop, and market penetration will grow. The process by which CAV technology is adopted is highly uncertain and speculative at this point, but is critical for making travel forecasts.

  In addition to the natural market forces that will play into the adoption of CAV technology, there exists further potential for key policies to impact this technology adoption process. As a historical example, seat belt laws forced automakers to start including this technology on all new vehicles. Similar types of laws could have impacts on how CAV technology is adopted.

- *When will CAVs be available?*

  A related question is when this technology will be available for public consumption. This impacts how we think about both short- and long-range forecasting.

Uncertainty Associated Connected and Automatic Vehicles

Researchers have identified a number of ways in which CAVs may impact how people travel. The following provides a list of these potential changes and the reasonings for each:

- **Roadway Capacities.** Higher roadway capacities will be possible, especially on limited access highways due to the ability of CAVs to coordinate movements. This effect, however, may be muted or even reversed in mixed fleet traffic if CAVs drive more cautiously and at slower speeds.
- **Improved Accessibility.** For individuals who are unable to drive an automobile on their own (e.g., disabled or older children), CAVs may improve accessibility. Such mobility impaired individuals can be chauffeured by an autonomous vehicle.
- **Value of Time.** Travelers' sensitivities to spending time in automobiles may reduce as a result of being able to focus attention on nondriving activities (e.g., reading or work).
- **Parking.** Parking needs and costs may reduce as CAVs can drop off a passenger and find parking elsewhere. Likewise, terminal times may be reduced.
- **Auto Operating Costs.** Auto operating costs may reduce as the driving behaviors of CAVs may be more fuel efficient than human operations. On the other hand, in mixed fleet traffic, CAVs may drive more cautiously and at slower speeds, which may negate efficiencies in acceleration and deceleration patterns.
- **Auto Occupancy.** Auto occupancy levels may reduce as people can more freely travel independently from one another, due to nondriver's traveling independently or due to easier vehicle sharing. Vehicle sharing may increase between household members as a result of unoccupied CAVs (i.e., zero-occupancy CAVs) able to travel between different household members.
- **Vehicle Ownership.** People may choose to own fewer vehicles due to the added opportunities for vehicle sharing.
- **Nonpassenger Trips.** Trips made to transport an unoccupied CAVs from one place to another will emerge, such as to travel to/from parking locations and/or cater to the needs of multiple household members.
- **Induced Travel**. Additional trips may be induced due to lower total costs of driving.

Modeling Connected and Autonomous Vehicles

Modeling CAVs comes with some serious challenges, largely because the CAV future has not yet begun. No observed data exists to inform how people will use this new technology, what specific features it will have, or how quickly it will be adopted. As a result, modeling the effects of CAVs is highly speculative.

Model Development

There are multiple ways in which the travel demand modeling community has envisioned that CAVs will impact both travel behavior and congestion. These include a number of variables described below. For each, we explain how these effects can be addressed in a trip-based travel demand model system.

- **Market Penetration.** Implicitly, all existing models have an assumption about the market penetration of CAVs, which is that it equals zero. Many studies examining CAVs have taken a simple approach of assuming full CAV market penetration, meaning all vehicles on the roadway network are CAVs. While a full market penetration scenario may one day occur, it is likely far into the distant future (at least 30 to 50 years). Considering a mixed fleet is a much more relevant scenario for most planning purposes.

  A full market penetration scenario makes the analysis simple because it does not require making assumptions about how different road users behave or about how mixed fleet traffic behaves. It also avoids the discussion of what assumption to make about CAV market penetration in any future year.

  However, for CAV mixed fleet analysis, market penetration must be considered as an explicit input to the model. Considering market penetration as an explicit input to the model is only relevant if the input has an effect on other components of the model. The input to the model may be user defined (e.g., the user specifies the market penetration as part of defining the scenario); or defined based upon a set of assumptions, such as market penetration projections. A user defined input to the model is probably more suitable for most applications since there is so much uncertainty associated with what market penetration levels will be in any future year. And also, given these high levels of uncertainty, market penetration is a key variable that could be used in an exploratory analysis, as described in more detail later.

  In the remainder of this case study, we consider market penetration as an explicit input and describe how that impacts other elements of the analysis.

- **Market Segmentation.** Segmentation is only important if market penetration is assumed to be between 0 and 100 percent. Treating market penetration as an input necessarily requires that market segmentation be addressed. The following are two basic approaches for dealing with market segmentation in the context of CAVs:

  - First, CAV users can be treated separately from non-CAV users, each group with their own behavioral characteristics.
  - Second, market segmentation could be avoided by creating behavioral characteristics that represent a weighted average of CAV and non-CAV users. If market penetration

equals 20 percent, then non-CAV users would be weighted more heavily in the average (at 80 percent) and CAV users less heavily (at 20 percent).

Averaging sensitivities or behaviors always comes with the risk of poorly representing all travelers' true behaviors because each group actually behaves differently from the average. Thus, from a behavioral perspective, the first approach of considering CAV users and non-CAV users as distinct travel markets in the model is the preferred approach. In the remainder of this case study, we assume CAVs and non-CAVs are treated as distinct traveler groups in the model.

- **Highway Capacity Adjustments.** One of the primary impacts touted of CAVs is that they will reduce congestion because CAVs will be better able to more efficiently use the space on the roads via shorter headways. Most experts believe that the primary benefits of CAVs from a capacity standpoint will be achieved on freeways and expressways, whereas arterials and collectors may see very limited improvements in effective capacity. In this case study, we assume that capacity improvements are limited to freeways with no impact on arterials and other roadway classes.

  Market penetration plays an important role as well since a future where all vehicles on the road are CAVs will have very different impacts on congestion than a mixed fleet scenario. Because market penetration plays a role, it is often useful to develop a relationship between market penetration and freeway capacity effects, such as the one shown in figure 4. In this particular example, capacity increases monotonically with market penetration, though other relationships may be suitable. For instance, a mixed fleet scenario could result in less efficient use of highway capacity (particularly at low levels of market penetration), depending on how CAVs interact with non-CAVs, which would lead to a relationship where the lowest levels of capacity are somewhere between a market penetration of 0 and 1.

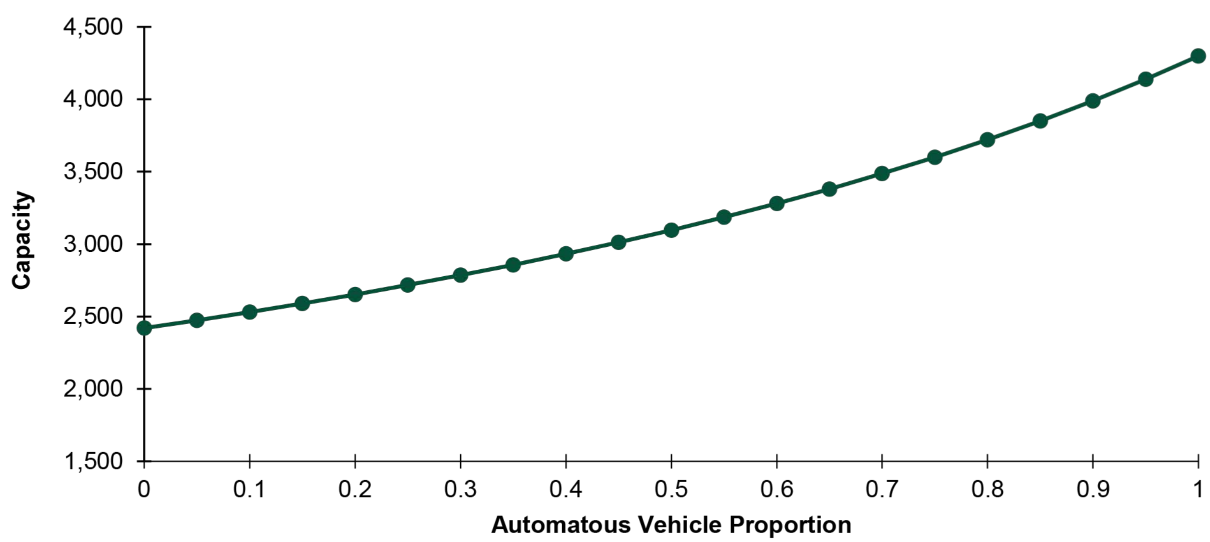


Figure 4. Chart. Formulaic highway capacity as a function of autonomous vehicle market penetration.

(Source: Friedrich, 2016.)

Highway capacity adjustments typically can be made fairly easily by factoring the roadway capacities for all links in the network with certain functional classifications.

Because most travel demand models assign traffic using aggregate traffic assignment algorithms, the market segmentation approach does not offer any benefit in more accurately reflecting changes to highway capacity resulting from CAVs in a mixed fleet scenario.

- **Auto Value of Time Adjustments.** Another way in which CAVs may impact behavior is in how travelers perceive the time they spend traveling, which is controlled in the modeled by the auto value of time. The argument here is that if travelers do not need to focus their attention on the task of driving, then they can spend that time engaged in other activities (e.g., reading, working, or even sleeping). Typically, assumptions about the reduction in auto value of time range from a reduction of 10 to 50 percent (see, for example, Litman, 2017; Childress et al., 2015; Kroger et al., 2016; LaMondia et al., 2016; Zhou and Kockelman, 2017; and Kohli and Willumsen, 2016; National Academies of Sciences, Engineering, and Medicine, 2018). This generally is consistent with the reductions in value of time often used for fixed guideway transit services, such as commuter rail, compared to bus transportation (typically 10 to 25 percent reduction; see FTA, 2008).

  While most travel models have explicit inputs for the value of time used in mode choice models, these inputs are not always easily modifiable for the auto mode in isolation. Often values of time are assumed to be fixed across modes or specific factors may be used for certain fixed guideway transit modes. As a result, changes to model code may be required to implement auto value of time adjustments.

  To implement this feature with the market segmentation approach, separate assumptions must apply to CAV travelers and non-CAV travelers. In particular, CAVs will enjoy the auto value of time adjustments, while non-CAV travelers will retain the original auto value of time of the model.

- **Parking Needs and Costs.** Parking needs and costs may be reduced as CAVs can drop off a passenger and find parking elsewhere. Likewise, terminal times may be reduced. Reductions in parking costs may be assumed to be anywhere from 0 to 100 percent. It is important to note that Government legislation may impact the extent to which such activities are allowed (e.g., zero-occupant parking journeys may not be allowed).

  This feature can be implemented in different ways depending on how parking costs are input to the model. For instance, the mode choice model coefficient specifically related to parking costs (if one exists) could be reduced, parking cost attributes at the zonal level (if they exist) could be reduced, or parking cost constants by land use category (if they exist) could be reduced.

  When market segmentation of CAVs and non-CAVs is used, implementation comes with additional challenges and may require custom changes to model code.

- **Fuel Costs.** Auto operating costs may be reduced as the driving behaviors of CAVs may be more fuel efficient than human operations. On the other hand, in mixed fleet traffic, CAVs may drive more cautiously and at slower speeds, which may negate efficiencies in acceleration and deceleration patterns. These potential benefits are often ignored in the development of CAV scenarios for travel demand models. As a result, there are few examples of the assumptions that have been made in practice to adjust fuel costs.

Typically, auto operating cost is a direct input to travel models and, thus, can be adjusted directly. A market segmentation approach may require custom coding that allows the auto operating cost input to be scaled by market segment (e.g., non-CAVs would be scaled by a factor of one, while CAVs would be scaled by the adjustment factor for CAVs).

- **Nonpassenger Trips.** Nonpassenger trips are those generated by CAVs where there are zero occupants in the vehicle. These arise as a result of CAVs dropping off or pickup up passengers in locations with limited parking where the CAV parks offsite. They also can arise if CAVs return home to serve the disparate needs of persons from the same household.

  Nonpassenger trips require special procedures to generate these trips and add into assignment procedures. This typically is done by examining the number and location of CAV trips generated by the model. For instance, trips with attraction ends downtown may be prime candidates for parking somewhere different than the actual trip end. These trip attractions would then serve as trip productions of nonpassenger trips. The locations of trip attractors for nonpassenger trips would need to be assumed, based on available parking space, parking costs, or some other measures.

  A market segmentation approach requires that the nonpassenger trips apply only to demand generated by CAVs, while non-CAVs generate no nonpassenger trips. This is straightforward in a market segmentation approach as the CAV demand is generated in separate trip tables from non-CAV demand.

Model Validation

Given the aforementioned lack of observed data, traditional model validation of CAV impacts is not possible. However, sensitivity testing is critical to understand how the modeled features of CAVs impact the results. Sensitivity testing should focus on ensuring that individual variables that are updated or changed as part of the CAV analysis provide reasonable sensitivities as intended.

Forecasting

Given that CAVs have not yet hit the marketplace for consumers and observed data does not yet exist, a great deal of uncertainty exists around how CAVs will affect travel. For instance, the adoption of CAVs is expected to follow an S-curve development pattern historically exhibited by new vehicle and other technologies. Figure 5 illustrates an example of this pattern over a 50-year horizon. Younger and higher income households are more likely to be among the first adopters of the new technology.

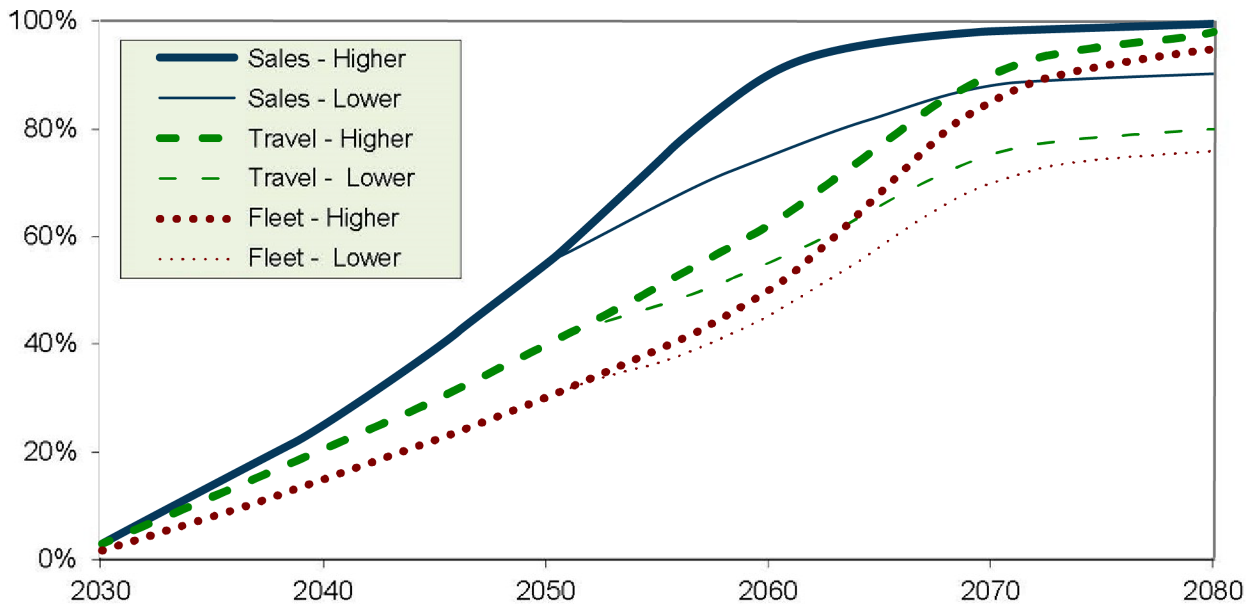


Figure 5. Chart. Autonomous vehicle sales, fleet, and travel projections.

(Source: Litman (2020).)

CAV market penetration rates are important assumptions for evaluating the CAV impacts on travel, but are not the only source of uncertainty. For example, a great deal of uncertainty also exists in how CAVs will impact roadway capacities, travel time sensitivities, and parking costs at different levels of CAV market penetration. Either additional assumptions need to be made to establish the relationship between these other CAV impacts and CAV market penetration, or each of these CAV impacts can be examined independently through exploratory modeling to explore the range of potential effects on the transportation system or specific policy analysis.

Emerging Modes

The last several years have seen a proliferation of new modes, including TNCs, scooters, and e-bikes. The characteristics of these modes and their use have changed dramatically in a short period of time. Increasingly, transportation planners are asked to evaluate how these modes are impacting regional travel patterns now and in the future.

Key Questions

There are a number of key questions related to emerging modes from a policy perspective, including the following:

- *What is the congestion impact of TNCs on the transportation system?*

  TNCs provide a taxi-like service that is accessible to many more individuals across larger geographies than traditional taxi services. Meanwhile, TNCs can effectively consume much more of the roadway space per passenger trip served due to idling and repositioning of the vehicle between passenger trips. How do these elements impact congestion in a region?

- *How will growing TNC usage impact the transportation system in the future?*

  Growing TNC usage allows people to have better, door-to-door accessibility options that travelers have traditionally only been able to obtain from the auto mode. This enhanced accessibility could potentially allow travelers to own fewer cars in the future and rely more heavily on shared mobility options, including traditional transit, TNCs, and other emerging shared mobility modes. What impacts will these potential changes have on the transportation system of the future?

- *Are TNCs competing more with transit or auto mode?*

  This is a critical question as traditional transit mode shares were on the decline across the country for several years prior to the COVID-19 pandemic. There is at least anecdotal evidence that TNCs are competing with transit more so than serving as first-/last-mile complements to transit. From this perspective, will traditional transit services, which many disadvantaged populations rely on, continue to be viable as TNC mode share increases?

- *What are the safety impacts of shared mobility modes?*

  Shared mobility modes typically are mostly available in the central urban areas of large metropolitan regions. In many such places, particularly where traditionally the auto and transit modes have reigned, travelers are having to learn how to share the highway network space with these new modes. Anecdotal evidence suggests that safety concerns have emerged as a result.

- *How do we manage curb and sidewalk space?*

  In some locations, e-scooters are prohibited from sidewalk use. However, most e-scooter and bike share programs utilize public/sidewalk space for device storage, which creates less space for pedestrians, as well as driveway access concerns. Meanwhile, TNCs have the potential to consume large amounts of curb space when not in transit. In some locations, dedicated TNC drop-off and pick-up zones have been created, but in many other locations, such zones do not exist. Consumption of curb space can create parking and delivery vehicle problems that need to be addressed.

Uncertainty Associated with Emerging Modes

Emerging modes may have a number of related impacts on how people travel. Some of these impacts are listed below.

- **Improved Accessibility.** Because TNCs have become ubiquitous in large metropolitan areas, they provide a transportation option for traveling longer distances when it otherwise may be much more difficult. For instance, someone who takes transit to work may be able to utilize a TNC to travel to a farther location to go for lunch or a work meeting. Likewise, shared mobility options may provide a faster transportation option than walking if vehicles are readily accessible such as in a downtown location.
- **Vehicle Ownership.** Because of improved accessibility offered by these emerging modes, private vehicle ownership may become less important for mobility. This likely will impact vehicle ownership levels over the long term.

- **Transit Use.** Likewise, because the TNC mode, in particular, provides a more personal transportation option, it may compete directly with transit for certain types of trips. This may result in a long-term downward trend in transit ridership. On the other hand, there are likely cases, especially longer-distance commuting trips where TNC may serve as a first-/last-mile option that is complementary to transit.

Data Availability

In most regions, usage of these modes remains relatively low. The National Household Travel Survey (NHTS) was conducted in 2017 and contains a robust sample of TNC users nationwide, with about 0.5 percent combined mode share of taxi and rideshare. As shown in table 1, only 8 percent of NHTS respondents reported using TNCs in the previous month. Scooters and e-bikes are even newer modes, and so even less data is available, but usage of these modes has been increasing over the last few years. Furthermore, there are certain areas, particularly in the more urban and centralized areas, where usage of these modes is much higher than national averages; in these areas, understanding trends is even more critical.

Table 1. National Household Travel Survey rideshare app use in past 30 days.

	Nationwide			
Rideshare App Use (Times per Month)	**Unweighted Persons**	**Unweighted Percentage**	**Weighted Persons**	**Weighted Percentage**
0	246,551	93%	276,271,841	92%
1	4,942	2%	6,245,758	2%
2	4,413	2%	6,128,432	2%
3	1,779	1%	2,544,664	1%
4	1,612	1%	2,485,143	1%
5	1,456	1%	2,458,909	1%
6	769	0%	977,803	0%
7+	2,376	1%	4,017,215	1%
Total	**263,898**	–	**301,129,765**	–
Overall Average per Person			**0.01126**	

(Source: National Household Travel Survey (2017).)

Transportation Network Companies

The information that currently is available on TNCs comes from several sources. First, recent household travel surveys have been recording TNC activity and have been used to examine the unique characteristics of TNC travel. To the extent that many characteristics of TNC trips are similar across different regions, these surveys are useful for understanding who and in what circumstances travelers are using these services. Table 1 provides a summary of rideshare use among NHTS respondents.

Using local household travel survey data would be better for estimating localized demand for TNCs, although many regional survey datasets may suffer from small sample sizes of trips made by the mode, making estimation and calibration of models more challenging. Nonetheless, these challenges are likely to be mitigated as usage of the mode increases and newer survey datasets are collected.

Several intercept surveys specifically of TNC users also have been conducted in cities, such as San Francisco (see Rayle et al., 2016) and Boston (see Gehrke et al., 2018). Targeted surveys such as these can provide detailed information on local use patterns.

While survey data is useful for understanding certain characteristics of TNC trips (such as demographics, travel purpose, time of day, and day of week), due to small sample sizes it may be less useful for understanding other TNC trip characteristics that also are important (such as overall trip rates, key drop off and pick up locations, and average wait times). Furthermore, survey data typically only reflects travel by *residents* in a region, while a large share of TNC travel is made by *visitors* to the region. As a result, transportation professionals have made attempts to obtain TNC operator data. Though these efforts have not been successful on a widespread basis, TNC operator data has been made available either for specific projects or publicly in a few locations, including San Francisco, Austin, Chicago, and New York. Efforts to examine these data have demonstrated key characteristics of TNCs and contrasted those against other travel, including taxis (Castiglione et al., 2016; Komanduri et al., 2018; Roy et al., 2020). However, the TNC market may varyfrom one region to another, and these data may not be useful for model calibration in regions other than the specific region in which data was collected.

TNC operator data also may contain information related to the preponderance of repositioning trips made by TNC drivers in between drop-off of one customer and pick-up of the next. Caution should be executed in transfering these data to regions other than the region in which data was collected, but the data may be useful in assessing general trends in terms of land use or key location types (e.g., airports and hotels) where TNC pick-ups typically occur.

E-Scooters and Bike Share

As with the TNC mode, the data for these technologies is limited, since these modes have only emerged in the past several years, and these modes only make up a small fraction of overall mode shares in most regions. They also serve a trip market that is less well studied and understood, tending to serve very short trips often in the urban core. These typically are nonhome-based trips and have a high prevalence of use by visitors to a region. While recent household travel survey datasets might be useful to support identification of trip types using these modes, sample size issues are likely to be an issue in many regions until these modes gain more traction with higher mode shares.

Another option for obtaining data related to these modes lies with the e-scooter and bike share operators. While there has been little success in agencies obtaining data from TNC operators without legislation that requires certain datasets be made public, there is much less experience in working with micromobility operators in a similar fashion. Efforts to engage these operators may be another way to obtain more and useful data.

Data related to the availability of these modes in different areas of a region is likely more readily available. If docking stations exist at predefined locations, it is possible to code the accessibility of these modes based upon the proximity of such stations. In some regions, dockless equipment is available; in which case, assumptions may be needed on availability that vary based upon land use or area type. If count data is available, that information can be used to help calibrate models.

Modeling Emerging Modes

In this section, we describe a hypothetical case study for introducing emerging modes in the model development process of a trip-based model. (It should be noted that many travel models do not have an explicit visitor model component, and visitors to the model region likely make up a disproportionate share of TNC passengers in a region, meaning that many TNC trips may not be directly considered.)

There are three key considerations that are addressed here for modeling these modes:

1. **Mode Choice Model Development.** Mode choice model design is a critical component to incorporate these new modes into the travel modeling framework.
2. **TNC Repositioning Trips.** These are nonpassenger trips made by TNC drivers to reposition themselves from the drop-off location of one passenger trip to the pick-up location of the next.
3. **Spatial Resolution.** Due to the shorter distance trips typically served by e-scooters and bike share modes, the model's spatial resolution is more critical.

These elements are considered more carefully below.

Data Preparation

To add these modes to the mode choice model, procedures for generating the modal attributes of these modes must be created. While it may be possible to create entirely new procedures for generating these attributes, in this case study, we consider approaches that allow for reusing old methods already in place in many models:

- **Travel Times.** TNC zone-to-zone travel times can be borrowed from other auto modes. Similarly, shared mobility modes' speed characteristics are similar to those of a standard bike, and, thus, travel times can be borrowed from those used for the bike mode. While it is important to consider that motorized shared vehicles may be prohibited from certain facility types like bike trails, most trip-based models do not have that level of detail for the bike network and rely on the highway network for generating bike travel times.
- **TNC Fares.** TNC fares are important because, unlike traditional taxis that have fixed fare structures related to trip distance or zone definitions, TNC fares vary widely depending on

demand and supply levels (e.g., surge pricing) and by type of service (e.g., shared services). While it may not be feasible to account for these fare structures directly in the model, an understanding of the fare structures is needed in order to develop an appropriate approximation that can be used in the model. For example, the Uber website (https://www.uber.com/us/en/price-estimate/) publishes fares that can be validated using randomly sampled fares from the app. The fare structure usually has a flat fixed minimum fare plus a distance-based rate adjustment. When multiple operators exist in the same area, a composite fare structure that approximates the average users' experience can be used.

Another element of TNC fares is differential pricing for different types of service. In particular, both Uber and Lyft offer shared ride services where a user agrees to ride with other passengers along their journey, potentially increasing total travel time, but at a reduced fare. To accommodate these different fare structures, as well as the difference in attractiveness and travel times of the different modes, wholly separate alternatives could be used in the mode choice model. However, the prevalence of shared TNCs may not be high enough to warrant modeling as a separate mode.

- **Shared Mobility Fares.** Shared mobility modes also include a fare component. These modes typically charge a flat fare to use a scooter or bike and add an additional fare that varies according to the amount of time the vehicle is in use. However, characteristics of the fare structure within the region of analysis should be examined carefully before making assumptions related to fares for bike share and scooters because fare structures can vary considerably between regions. When multiple operators exist in the same area, a composite fare structure that approximates the average user experience can be used. For the purposes of this case study, we assume that shared mobility mode fares can be generated as a function of a constant and travel time.
- **Wait Times.** Wait time is another key attribute of TNCs, and different options exist for handling it. One simple approach is to assume wait times vary by area type, since more TNCs operate in the urban centers of regions than in rural areas. In the example above where fares were randomly sampled using the app, wait times also were collected and could be summarized by area type. However, it is important to consider that individuals using TNCs are often aware of typical wait times and may be able to minimize wait by scheduling their trip in advance.
- **Accessibility Measures.** Accessibility measures also are an important attribute that could be included in the mode choice model for both TNC and shared mobility modes. This is particularly true for shared mobility since some of the accessibility characteristics associated with TNCs are the result of differential wait times, which can be accounted for directly as discussed above.

  Shared bikes and scooters, however, are often only available in the urban core of a region. If docking stations are required, then availability depends on the locations of those stations. Docked service availability can be approximated by estimated (walk) access times to the docking stations from each zone. If dockless equipment is used, then an approach similar to the wait times for TNCs can be used, where access times vary based upon area type. Furthermore, during peak periods (typically midday and PM peaks), the availability of these vehicles may be more limited due to high usage levels of deployed vehicles. This temporal variation in availability can be handled through shadow prices, though that level of detail in modeling probably is not needed for most applications. Time-of-day factors that represent differences in demand by time of day are likely sufficient.

In addition to the attributes of these new modes, the geographic resolution is another element that deserves consideration in model design. While not critical for TNC modes, one of the key challenges with shared mobility modes, similar to walk and bike modes, is that these modes typically are used for short-distance travel. As a result, the geographic resolution of the model can be very important for accurately forecasting these modes. Microzones or parcels, which replace traffic analysis zones (TAZ) as the spatial unit of measurement for mode choice in some activity-based models, are more suitable for modeling these modes.

Model Estimation

Mode choice model design is a critical component to incorporating these new modes into the travel modeling framework. At a minimum, a TNC mode or a shared mobility mode could be added to the mode choice model. A variety of different correlation structures (e.g., in a nested logit model formulation) could be envisioned for including these modes.

In addition to acting as primary mode options, both TNC and shared mobility modes have the potential to provide access or egress options for using transit (so-called first and last-mile options). Because the bike mode offers similar level of service characteristics to shared bike and scooter modes with respect to travel times and given that bike access/egress is often not even included explicitly as a mode option in these models, explicit treatment of shared bike and scooter modes is not necessary. TNCs have distinctly different characteristics than other auto and walk access/egress mode options, including the following:

- Unlike park-and-ride, TNC mode could be used as either an access or egress mode (or even both).
- Cost structures and wait times for TNC are different from other auto modes.

Therefore, depending on the structure of the transit modes in the mode choice model, TNC access and egress to transit could be explicitly included as a mode within the mode choice model.

More recent survey data (collected in the last few years) likely contains trips taken by the TNC mode, and potentially shared mobility options. Recent transit onboard survey data also may contain transit trips that use TNC access and/or egress. However, even in these best-case scenarios, the data may be thin due to the modest overall share of these trips. Given this scarcity of data, especially shared mobility modes, it may not be possible to estimate parameters of the model directly.

In cases where model parameters cannot be estimated using local data, asserting model parameters may be acceptable. Travel time and cost sensitivities are often constrained to be identical across all modes in mode choice models, and estimates of these sensitivities for standard modes could be applied to these emerging modes. Mode constants could then be calibrated. If the mode choice models are estimated as nested logit models with specific correlation structures, asserting the correlation structures for new emerging modes in the model is less straightforward, but can be done. Some potential guidelines include:

- Models with a motorized nest (that includes auto and transit) also can include TNC as a main mode.
- Nonmotorized nests could include shared mobility modes (note that even though e-scooters may technically be motorized, they share more in common with nonmotorized modes).
- TNC is likely distinct enough from auto or transit to fall into nests specific to either of those mode clusters, or as its own nest.

It is important to recognize that these are not hard and fast rules. Different approaches may be suitable in different circumstances.

Repositioning Trips

Repositioning trips are specific to the TNC mode. No standard approach for incorporating TNC repositioning trips in travel models currently exists, and, thus, ad hoc methods have been used. In an activity-based model, it is possible to explicitly link personal travel TNC trip ends to create a trip table of repositioning trips, and similar, less refined approaches may work for trip-based models.

One option is to model reposition trips by balancing the TNC demand trip table. The trips generated by the passenger travel model are balanced by adding additional vehicle trips to the trip table. These additional trips were assumed to be the zero-occupant vehicle trips between drop-off locations and pick-up locations.

This method requires a number of assumptions. For instance, if we observe one TAZ with 100 drop-off trips and 50 pick-up trips, a simple solution would be to add 50 zero-occupant trips with a starting point in this TAZ and destination in some other TAZ. However, this method assumes that the other 50 drop-off trips have a pick-up trip in the same TAZ, which is not necessarily the case. It could be that all 100 drop-offs have a subsequent pick-up in a different TAZ, and all 50 pick-ups have a previous drop-off in a different TAZ.

TNC operator data can help to estimate the relationships between repositioning trip ends. Region-specific data is best, but if such data does not exist, TNC operator data from other regions provides a baseline from which to inform assumptions specific to another region. The type of model that can be used here is a trip distribution model (either a gravity model or a destination choice model). In either case, the model matches the drop-off locations to subsequent pick-up locations, similar to the trip distribution methods already in use by travel models. Once a trip table for these trips is generated, these vehicle trips should be assigned to the network during traffic assignment step.

However, even these methods leave out two important features of TNCs. First, it may not be possible to account for destination-less repositioning, for instance, if the TNC driver drops off a passenger and immediately begins driving without having another passenger locked in to be picked up. Second, since many travel models do not have an explicit visitor model component, and visitors likely make up a disproportionate share of TNC passengers in a region, the overall levels of repositioning trips may be vastly underestimated. Explicit representation of visitor travel in the travel model would improve the realism of both the overall model system, as well as the TNC repositioning component.

Model Validation

There are two key ways that models should be validated when incorporating emerging modes. First, mode choice model constants must be well-calibrated to current levels of mode usage. While the survey data available for estimating mode choice models with these new modes may be thin, in most regions, there is more data available to estimate current levels of use for these modes. Creating calibration targets for each new mode will allow for the calibration of modal constants to ensure the model generates a reasonable number of trips in the base year model.

Secondly, sensitivity testing provides information about the reasonability of the model results. Sensitivity tests may include adjusting sensitivities to attributes like travel times and costs or adjusting assumptions about modal attributes like fare structures and/or waiting times. Other sensitivity tests may be less specific to these modes, but could provide more information about the model's response to larger regional questions. For instance, a common type of sensitivity test is to change land use assumptions. Increasing the size of the urban center (by increasing densities) as a sensitivity test might be expected to result in an increase in usage of emerging modes.

Forecasting

Emerging modes have been gaining market share for the past several years due to their added convenience and improvements to accessibility. Forecasting of emerging modes comes with a number of associated uncertainties; all of which are potential candidates for exploratory modeling, including the following:

- Overall use levels continue to rise. Exploring scenarios where the overall usage rates grow due to uncontrolled attributes could provide value (e.g., by manipulation of the modal constants).
- In the future, improvements in modal level-of-service attributes may be possible. For instance, in the case of TNCs, CAV technology may make manually operated TNCs obsolete. Given that the driver of the TNC is the greatest cost for providing TNC service, this could drastically reduce the rates charged by TNC operators.
- As described above, asserting model parameters for these modes may be necessary due to lack of data currently. These asserted model parameters are great options for exploring uncertainty in the model specifications.

E-commerce

The steady rise in e-commerce activity over the past several years has had important impacts on transportation. Accounting for these effects within our travel modeling tools can be quite complex. As discussed earlier, not only has e-commerce resulted in a reduction in shopping trips (which can be measured using survey data), but it also has caused the proliferation of business and home parcel delivery trips at rates never before realized. As a result, there are two key elements to forecasting the effects of e-commerce using a travel model: personal and household shopping effects and freight and delivery vehicle impacts. In this case study, we describe an approach to explicitly consider the levels of e-commerce in a travel model.

Key Questions

There are several key questions related to work-from-home behavior from a policy perspective, including the following:

- *Will the rise of e-commerce continue to affect personal shopping travel patterns?*

  Conventional wisdom suggests that people are traveling less for shopping purposes than in the past due to the availability of online shopping. In fact, data over the past couple of decades supports this idea. However, it is difficult to parse out the cause and effect of increased online shopping and decreased shopping related travel. Presumably, increased online shopping is having some effect, but other ancillary drivers may also be playing a role, like consolidation of retail businesses and larger box stores, which can cater to multiple shopping needs (rather than specific niche shopping needs).

- *How has e-commerce impacted parcel delivery and how has it impacted other freight goods movement?*

  With the increase in e-commerce, there also has been a growth in parcel delivery travel activity over recent years. There is likely a direct correlation between the levels of e-commerce and volume of parcels, but understanding this relationship is important to crafting policy.

  Further, while in the past, the parcel delivery vehicle fleet largely consisted of single-unit trucks (e.g., UPS and FedEx), more recently many parcel delivery vehicles are smaller light-duty vehicles. These changes have an impact on the delivery patterns. Larger vehicles obviously can carry more volume but may be less efficient both in terms of fuel consumption and time.

- *What are the levels of parcel delivery? What types of travel patterns do parcel deliver vehicles have (e.g., routing of delivery vehicles)?*

  Simply measuring the levels of parcel deliveries can be quite challenging since data is not readily made available to the public. Nonetheless, measuring these levels is critical to determining a baseline of parcel delivery activity and relating this to e-commerce activity.

  The number of deliveries also may be impacted by volume. For instance, a customer that receives a lot of parcels may have many parcels delivered in a single delivery. Conversely, each parcel may be delivered separately. These two options have very different implications for the amount of travel generated by each parcel. Speedy delivery guarantees from shippers also may have an impact on whether parcels are delivered in one shipment or many (i.e., if delivery has been guaranteed in one or two days, then it is less likely that parcels can be packaged together).

  Another important consideration is the scheduling of parcel delivery stops, which is typically handled by a shipper. While the algorithms for scheduling stops are proprietary, it may be possible to replicate these processes using standard approaches (e.g., the traveling salesman problem).

- *Does an increase in online shopping and parcel delivery increase or lessen congestion?*

  There are differing schools of thought with regard to whether e-commerce, as a whole, has resulted in more or less efficient distribution of goods. E-commerce has essentially replaced some individual personal shopping trips (e.g., between home and shopping locations) with delivery stops made as part of a longer parcel deliver vehicle tour. While it is likely true that each individual delivery trip is shorter distance and more efficient than a roundtrip shopping trip, a person making a shopping trip may be more likely to purchase multiple goods with a single trip; whereas parcels may more often be of a single good, requiring multiple parcel delivery stops. These tradeoffs currently are not well understood.

  Furthermore, some research has suggested that people tend to have relatively fixed travel time budgets (Mokhtarian and Chen, 2004). That is, individuals are willing to travel a specific amount of time per day (e.g., one hour), and they tend to fill this time budget. In that case, a reduction in personal shopping travel may be replaced with a commensurate increase in travel for other purposes (e.g., willingness to travel farther for work). The result would be an increase in parcel delivery travel with no change in the amount of total personal travel.

Uncertainty Associated with E-commerce

E-commerce may have a number of different impacts on travel. As noted above, e-commerce reduces the need for personal shopping travel while increasing parcel delivery travel. These can have a myriad of impacts on how people travel:

- **Travel budgets.** The reduced personal shopping travel may offer the opportunity for more personal travel for other purposes. Research has shown that total travel budgets seem to be fixed over time, even when travel becomes more efficient.
- **Activity patterns.** Even if travel budgets are not fixed, the reduced need to shop in-person frees up time in people's schedules for other activities. Understanding what these activities are and how they are substituted is necessary to forecast the impacts of a changing e-commerce landscape.
- **Parcel deliveries.** As noted above, e-commerce likely has a direct correlation with the number of parcel deliveries in a region.
- **Timing of travel.** Personal shopping travel is often associated with nonpeak travel patterns (though this is not always the case). Reducing such travel has the potential to adjust time-of-day distributions in a region. Furthermore, the timing of parcel delivery travel has unique characteristics as well, and understanding those is important for forecasting the impacts of e-commerce.

Data Availability

National economic data can provide big-picture trends in the growth of e-commerce over time. Figure 6 shows national commerce data from the U.S. Census for year-by-year e-commerce retail trade as a percent of total retail trade. E-commerce, as a share of total retail, has grown from less than 1 percent in 1998 to about 10 percent in 2018. While this trend will inevitably level off at some point, if recent trends continue in the short term, e-commerce as a share of total retail could eclipse 20 percent by 2025. In theory, these trends should continue to reduce the number of personal shopping trips while increasing the number of parcel delivery trips. One key question deals with how to predict e-commerce trade into the future.

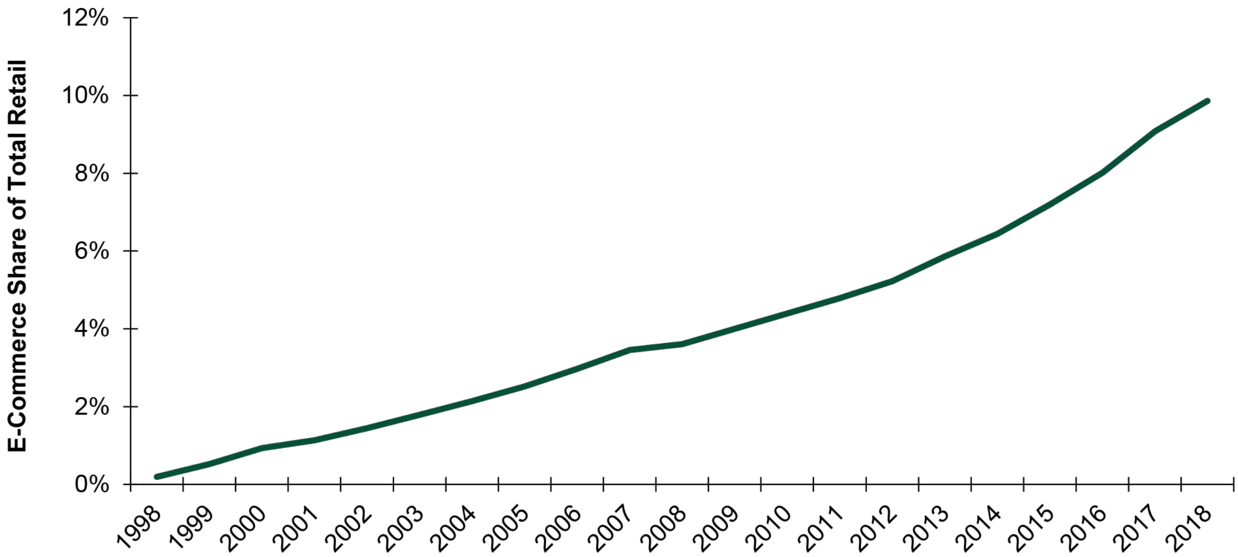


Figure 6. Chart. E-commerce retail trade sales in U.S. as a percentage of total retail sales.

(Source: U.S. Census E-commerce Statistics, available at: https://www.census.gov/programs-surveys/e-stats/data/tables.html.)

The drop in shopping trips over the last 20 years can be measured using historical and more recent household travel surveys. National data on trip rates by purpose from the NHTS datasets for 1990, 1995, 2001, 2009, and 2017 are shown in figure 7. The figure clearly shows that person trip rates overall have been dropping since 1995, and that this drop has been fueled by reductions in shopping trips by about -33 percent from 1995 to 2017.

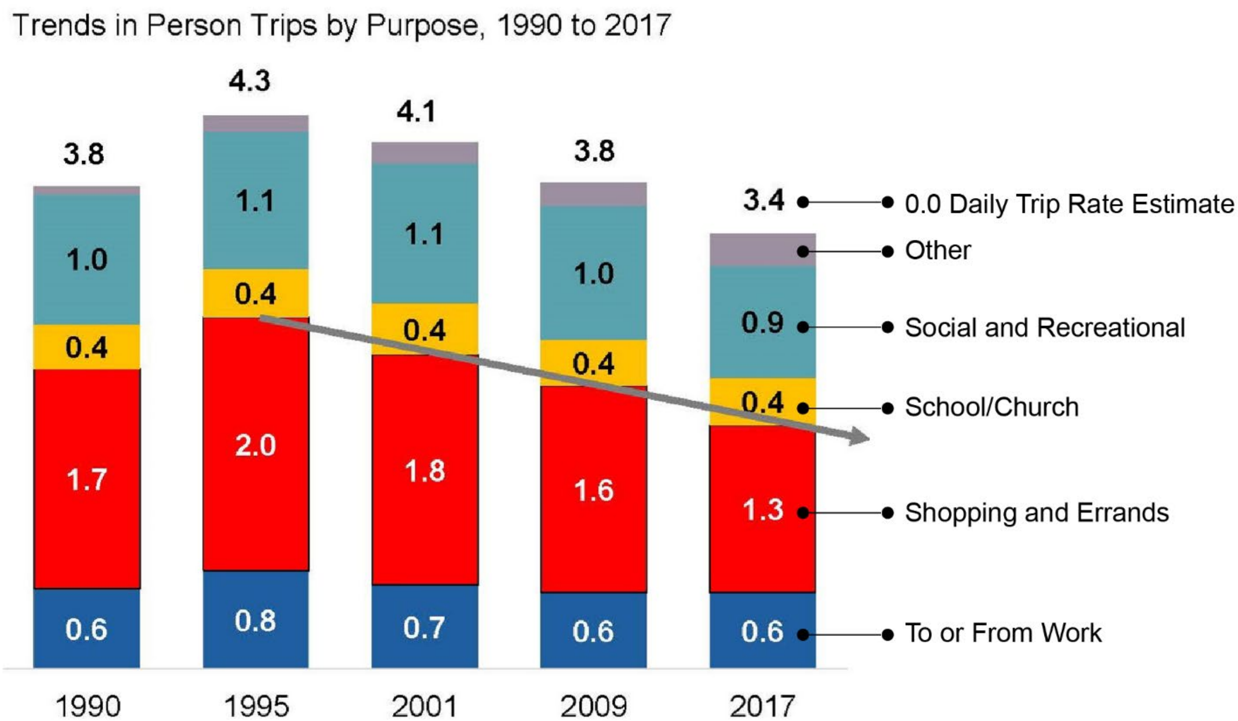


Figure 7. Bar chart. Trends in trip rates by purpose over years from national survey datasets.

(Source: McGuckin and Fucci, 2018.)

Further analysis of data may look into how these trends in shopping travel have occurred across different segments of the population. It could be that higher income households with greater access to the Internet and online shopping have seen larger drops in shopping travel overall than lower income households. Vehicle availability also may be related as households with insufficient vehicles may have more incentive to use online shopping to fulfill these needs.

Trends in delivery vehicles are not easily measured. Data from parcel delivery companies is not readily available as it is proprietary. While some parcel delivery information may be available from some Global Positioning System (GPS) truck databases (e.g., INRIX or StreetLight truck data), these data vendors will not release the specific truck companies that form their samples. In the future, the transportation planning community should push for companies like Amazon, UPS, and FedEx to release parcel delivery and truck information that can be used for planning purposes. In the meantime, however, other datasets may be useful, such as Rakuten data (https://www.rakutenintelligence.com/), which can provide information on where parcels are being delivered and by which company, but only for the sample of users of their service.

Modeling Shopping Activity

In this section, we outline our case study illustrating the steps that can be used to address the impacts of e-commerce on shopping activity in a trip-based model. Note, this example focuses only on household travel and not modeling of truck movements (i.e., parcel delivery trips), which is discussed in the next section.

Data Preparation

The data preparation process for model development looks similar to typical data preparation, but with some added layers. In a trip-based model, trip rates by trip purpose must be estimated, typically using recent household travel survey data collected in the region. This process does not change when trying to address e-commerce, since recent shopping trip rates still must be input to the model.

However, because e-commerce affects shopping travel, in particular, it will be useful to isolate shopping travel from other trip purposes. For home-based travel, many models already treat shopping travel as a distinct trip purpose, as we will do here. For nonhome-based travel, we also will need to summarize travel with a shopping purpose from other trip purposes. For nonhome-based travel, we need to only know the share of nonhome-based trips (work versus other) that have at least one trip end at a shopping activity.

Model Estimation

Similarly, model estimation proceeds as is typically done using a recently collected household travel survey to estimate trip rates for different segments of households. There need not be any change in the standard approach to model estimation.

However, in order to recognize the reality that e-commerce continues to change year by year and also has an impact on shopping travel, the model must be designed to be responsive to the levels of e-commerce. For this case study, we assume that e-commerce trade as a share of total retail is a metric that can be forecast in the future.

In order to make use of the historical e-commerce data, we also will use the historical shopping trip rates estimated from NHTS sample shown above. It is possible to measure the drop in shopping trips and relate that drop to levels of e-commerce described from the U.S. Census. Figure 8 shows an example of this using an exponential curve.[1] This curve suggests that when the e-commerce share of retail reaches 20 percent, shopping trips per person will reduce to only 0.8 trips per day, which is a reduction of about 40 percent from 2017 shopping trip levels.

1 Note that we included a point for the 1995 shopping trip rate from the NHTS of 2.0, even though U.S. Census data only goes back to 1998. We assumed e-commerce share of 0.1 percent in 1995.

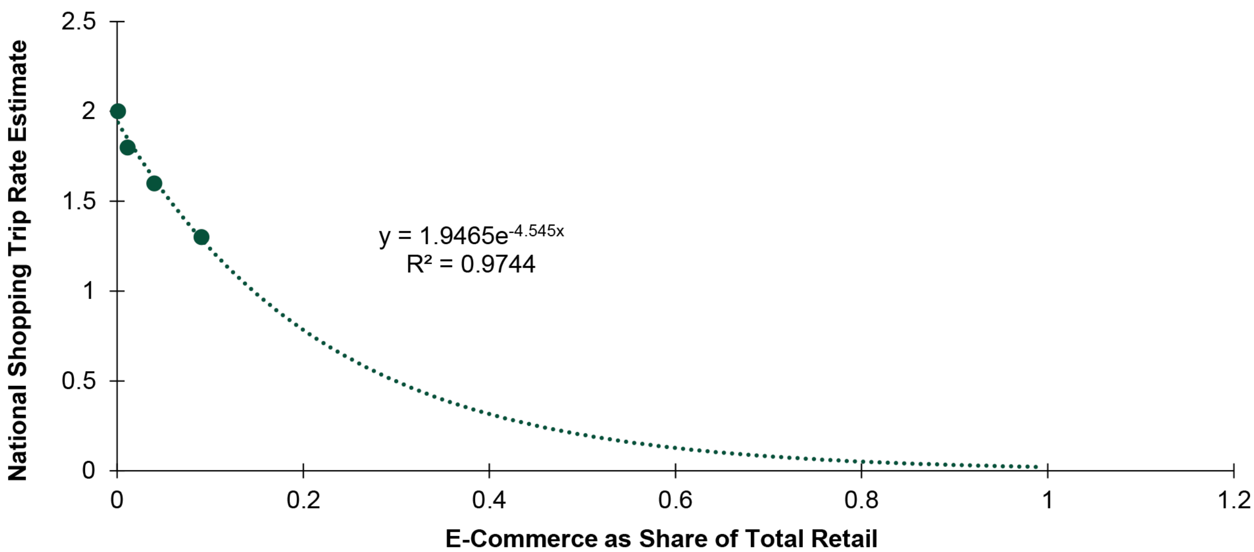


Figure 8. Chart. Example of exponential curve-fitting to e-commerce and shopping trip rate data.

(Source: U.S. Census Bureau.)

In this case study, we use this relationship to automatically adjust the shopping trip rate used by the trip generation model by the appropriate amount based upon the model's forecast year. This is done using the following steps:

1. Using the fitted curve, calculate the shopping trip rate in the model's base year. Here we use 2017 as the base year with a shopping trip rate of 1.3.
2. If the forecast year is input as 2020, the calculated trip rate is 0.8. In this case, this corresponds to a reduction in shopping trips of 40 percent.
3. The calculated reduction in shopping trips can be directly applied to the home-based shopping trip rate to generate the forecast year shopping trip rate:

$$TR\_adj = TR\_orig * (1 - \%\ reduction)$$

4. For nonhome-based travel, we must make an adjustment to this reduction based on the percent of nonhome-based travel that has one activity end that is for the purpose of shopping. This can be done using the following formula:

$$TR\_adj = TR\_orig * (1 - \%\ shopping) + TR\_orig * (\%\ shopping) * (1 - \%\ reduction)$$

It is important to note that other factors may be impacting shopping travel behavior. Furthermore, the relationship in figure 8 was fit using only four data points. Different assumptions about the curve and fitting the curve would lead to vastly different conclusions. For instance, we may not expect shopping trips to completely disappear at high levels of e-commerce. Figure 9 shows an example of this with a curve that fits the data almost, as well as the one above. In this case, shopping trips approach 0.5 per person per day as e-commerce increases, rather than approaching zero trips.

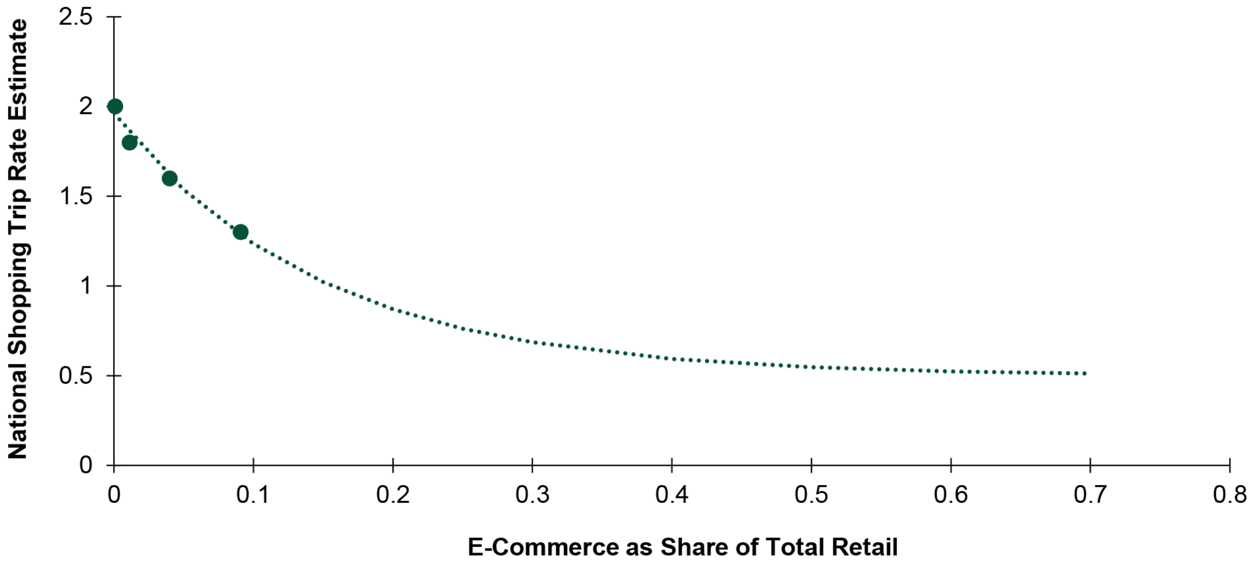


Figure 9. Chart. Alternative example of exponential curve-fitting to e–commerce and shopping trip rate data.

(Source: Federal Highway Administration.)

Model Validation

The traditional approach to model validation can still be used to validate the new model. All of the traditional measures for model fit and appropriateness apply, but there is one new model component that relates shopping trip rates to e-commerce levels, which is calibrated to the existing data, but can not validated except in regard to sensitivity testing to ensure model is working as expected.

If a historical household travel survey dataset for the region is available, the change in shopping travel trip rates can be directly measured between the recent and historical survey data and compared against the fitted curve. This could give some confidence that the relationships are valid.

Another tool that could be utilized, and in fact, is recommended by the Model Validation and Reasonableness Checking Manual (Cambridge Systematics, Inc., 2010), is a full backcast. A backcast is when a model is applied and forecast for a year that occurred in the past, using historical input data for things like transportation networks, population and employment, and land use. The backcast model results can then be compared against historical data like traffic counts, transit boardings, and other traditional metrics.

Other model sensitivity tests also would be valuable. A number of questions may be valuable in devising sensitivity tests, such as the following:

- How does the model respond to an increase in e-commerce activity? Are the drop-in trip rates that results from growing e-commerce activity reasonable? What impact does this have on other key metrics, like trip lengths and travel times?

- What are some of the limitations of the model sensitivities? For instance, how much confidence is there in the model's implied response to very high e-commerce levels? Because we have not yet observed e-commerce levels above about 10 to 15 percent of total retail, any forecast beyond those levels represents an extrapolation of existing data. How confident are we in the model's forecasts at e-commerce levels that are much higher? Developing an appreciation of our confidence in model results is as important as the results themselves.

Forecasting

Any forecast using the model must recognize the added input of e-commerce levels. This could be treated as a user-specified input value or a simple trend forecast could be made that provides an estimate of e-commerce levels for any specified forecast year. The latter could be informed based upon historical data.

Building the model as described in this section lends itself nicely to the exploratory modeling and analysis process. Since uncertainty exists about the magnitude of e-commerce in the future, the e-commerce levels could be treated as an uncertainty in any exploratory modeling and analysis work. Uncertainty also exists about the impact that e-commerce has and will have on levels of shopping travel. There may be wide ranges for the distribution of the effect of e-commerce levels on shopping travel, which would appear as uncertainty in the fitted curves shown in figure 8 and figure 9. As such, additional levers could be built into the model that account for this uncertainty in how e-commerce levels impact the trip rates. This would require a more complex model development process, but could allow tests of these uncertainties as well.

Modeling Freight and Parcel Delivery

In this section we describe the steps that can be used to address the freight and parcel delivery impacts of e-commerce in a trip-based model.

Data Preparation

Traditionally, regional truck models do a poor job of estimating parcel delivery trips within a region. This is largely because the data to estimate these trips simply do not exist. The natural entities to acquire these data from are the shippers themselves (e.g., UPS, FedEx, and Amazon). However, shipping companies are not necessarily interested in sharing their data with the planning community due to it containing proprietary information of which these companies are highly protective.

The emergence of new data sources, such as the parcel delivery data noted above, could allow for estimating parcel delivery activity. In particular, these data could be used to estimate differences in the amount of parcel delivery activity across different segments of the population. However, these data are incomplete representations needed for modeling since they only contain information about where deliveries are being made. They do not contain information about how trips get linked between these locations nor is there any clear way of expanding these data.

Other national-level data estimates and anecdotal information could be used as guides to make assumptions about how to size the volumes of parcel deliveries. Some key pieces of information include the following:

- In the U.S. in 2019, Amazon shipped 2.5 billion packages, FedEx shipped 3 billion packages, and UPS shipped 4.7 billion packages (see https://www.forbes.com/sites/andriacheng/2019/12/12/how-serious-is-amazons-threat-to-ups-fedex-study-finds-it-could-soon-beat-them-in-us-package-delivery-volume/?sh=4d1ae53368f4).
- Amazon drivers deliver between 80 to 250 parcels per day (see https://www.businessinsider.com/amazon-delivery-drivers-reveal-claims-of-disturbing-work-conditions-2018-8#:~:text=A%20couple%20of%20years%20ago,250%20parcels%20daily%2C%20drivers%20said).
- UPS drivers deliver about 120 parcels per day (see https://www.wired.com/2013/06/ups-astronomical-math/#:~:text=At%20UPS%2C%20the%20average%20driver,giant's%20director%20of%20process%20management).

These figures can provide some rough approximations of the overall levels of parcel delivery activity at a national level. Making some assumptions about how national trends scale to a region, it may be possible to reasonably determine the overall size of the parcel deliver activity at a regional level.

Model Estimation

While it may be tempting to think about how to model parcel delivery activity within the existing framework of a regional truck model, the travel patterns of parcel delivery vehicles are likely quite distinct from those of other trucks. As a result, it may be warranted to develop parcel delivery vehicle models that are separate from other truck travel. In the case study of a trip-based model, we take the latter approach using the following steps:

- Using Rakuten or similar data, estimate the distribution of parcel delivery activity to each TAZ. In this case, we will use a linear regression model. The independent variables of this model may include the following:
  - Employment by type at the TAZ level.
  - Households by market segment at the TAZ level.
  - Accessibility or area type variables at the TAZ level.

  These variables are already readily available since they are used as inputs to the standard trip-based model.

- Using national statistics, estimate the number of parcels delivered in the region per weekday. This requires a number of assumptions.
  - First, we assume that the regionwide number of parcels is proportional to the population of the region. Nationally, UPS, FedEx, and Amazon delivered 10.2 billion packages in 2019 (see above). In our case study, we assume the region has a population of 5 million people, which equates to 160 million parcels per year in the region.
  - Second, we convert the yearly figure to daily weekday parcels. Because fewer parcels are delivered on weekends than weekdays, we use an assumption of 6 weekday equivalents per week on which deliveries are made. This results in an estimate of about 500,000 parcels delivered each weekday.
  - Last, we use the estimates of average parcels delivered per day to derive an average number of vehicles used to deliver the parcels. If we assume 150 parcels per driver per day (range above was 80 to 250), we get an estimate of about 3,400 drivers, each making 150 deliveries per day.
- The next step in the process is to link trips. This step requires some big assumptions, since we have no data to support building the model. However, we can reasonably guess that one objective of each driver is to minimize the time spent traveling, and we can reasonably guess that shippers want to maximize the number of deliveries a driver can make by ensuring deliveries are close to one another. While it is unlikely that shippers could cluster all deliveries for any given driver in a single TAZ, it is likely that each driver may only visit a small number of TAZs that are close in proximity (in addition to the trip to/from a warehouse). By assigning each parcel delivery to a driver and minimizing the travel for each driver, the trips can be chained, and a trip table developed.
- The last step is to relate parcel delivery activity to e-commerce activity. Since no data exists here either, one simple assumption is to consider a one-to-one relationship between e-commerce and number of parcels that need to be delivered. The e-commerce activity is already being used as an input to adjust shopping trip rates. For any forecast year, the level of e-commerce activity is compared against the base year, and the number of total regionwide parcels is factored by the ratio of the e-commerce level in the forecast year and the level in the base year.

Model Validation

Unfortunately, model validation data does not really exist to test the validity of the parcel delivery model developed for the case study. However, it would still be possible to conduct a collection of sensitivity tests and check reasonableness of model forecasts. Sensitivity tests that could be conducted include the following:

- Test the reasonability of parcel deliveries per day per driver. This could be done by computing the total travel time based upon the set of trips generated for a driver and comparing against what is feasible within the time constraints of a day.
- Test the impacts of raising or lowering the e-commerce levels for a given forecast year. Given the concurrent reduction in shopping trip rates, this sensitivity test would shed light on

how the model is forecasting e-commerce to impact levels of congestion and other travel metrics in the region.

Forecasting

A great deal of uncertainty exists with e-commerce, both because there are gaps in the transportation data available to study it, and because it is a rapidly changing technology. As a result, the methods described here lend themselves well to exploratory modeling, especially on the parcel delivery side where the existing data is particularly thin. Large uncertainties are associated with the modeling components developed to forecast parcel delivery, as discussed.

Work from Home

While working from home has been around for decades and has long been studied as a transportation demand management concept to reduce congestion levels on roadways, advances in communication technology and data systems have allowed more workers freedom to work from home in recent years. Based on data from the BLS, about seven percent of workers in the U.S. had the access to work from home as a benefit in 2019, which was up from about five percent in 2010 (DeSilver, 2020). These trends were exacerbated during the COVID-19 pandemic of 2020, where many workers were forced from their normal places of work to work-from-home arrangements. While it is unclear whether the trends that emerged during the pandemic will persist long term, the recent acceleration of work-from-home activity has encouraged agencies to have a renewed focus on examining work-from-home behaviors and policies supporting these behaviors.

Key Questions

There is a number of key questions related to work-from-home behavior from a policy perspective, including the following:

- *What are the characteristics of the telecommuters?*

  Many professions cannot work from home due to the characteristics of their jobs. Examples of such professions include many retail and food service jobs, industrial and warehousing jobs, and medical professions. Many of these jobs have lower wage levels and, thus, there is a correlation between household income and ability to work from home, with lower income workers working from home less.

  Types of jobs that can more easily cater to working from home include jobs in many office settings, such as financial sector, professional services, and public administration.

  Regular and infrequent telecommuters also may live in different kinds of locations than other workers and have differing transportation needs. For instance, they may be less likely to rely on transit for commuting needs and more likely to own automobiles.

- *Are the features of regular telecommuters different from infrequent or flex workers?*

  There is a growing segment of workers whose usual workplace is a nonhome location, but who routinely telecommute one or more days per week. The characteristics of these workers may be quite different from workers whose workplace location is their home.

Among regular telecommuters, there also may be key differences. For instance, some workers with home as their regular workplace may have professions that require regular out-of-home work travel, such as real estate agents or certain types of businesses like contracting. The characteristics of these workers likely differ from the characteristics of telecommuters who primarily work from their home.

- *What days of week will flex workers actually come into work and what days will they work from home?*

  A company that staggers its office workforce across different days has lower office space demands, which could potentially save money. However, some professions may require workers to collaborate and be at the worksite on the same days. Understanding the differing needs of different types of jobs where work-from-home behavior is prevalent is important for understanding the behaviors of these workers.

- *Will workers continue working from home after the COVID-19 pandemic ends?*

  As noted above, many employers were forced to have their workers start working from home during the pandemic. For many companies, this meant making new investments in equipment and other services to support a remote workforce. Part of the reluctance by some companies to allow working from home in the past may have been around the overhead of making their systems work for a remote workforce. Now that those investments have been made, there may be less reluctance to allow continued work-from-home arrangements, particularly given the potential of cost savings for things like office space.

  Moving past the pandemic, it will be important to monitor whether some of the shift to work-from-home persists or if work-from-home patterns return to a prepandemic state. While there will likely be a downturn in working from home following the pandemic, levels of working from home are unlikely to return to prepandemic levels.

Uncertainty Associated with Working from Home Behaviors

Working from home has a number of impacts on travel behaviors. Obviously, working from home eliminates commute trips, but work-from-home patterns also have other distinct features that set them apart from other individuals. For instance, telecommuters are likely to block off time to adhere to work schedules, but also are more likely to be available for running household errands and providing rides to other household members than they would be if they had to report to a fixed workplace on a given day. This translates to some key travel pattern differences between telecommuters and other workers:

- **Types and purposes of travel.** Telecommuters may have no work travel, but may engage in other out of home activities that onsite workers do not, like shopping or escorting activities. Other telecommuters may have regular work-related travel.
- **Time of day.** Telecommuters have more flexibility during the travel day than onsite workers with fixed schedules.
- **Childcare needs.** While many telecommuters likely utilize childcare services for young children that may be in the household, others may forego childcare services to take advantage of cost savings.

- **Vehicle ownership.** In many locations, telecommuters may have different needs in terms of vehicle ownership due to differences in modal options available for their daily travel needs.

Data Availability

A fair amount of data is available for understanding and modeling telecommuting trends, although most of the data is from before the COVID-19 pandemic. Levels of telecommuting are highest among higher income earners. Based on the American Time Use Survey (ATUS) (Bureau of Labor Statistics, 2019), conducted by the BLS, a number of characteristics helps explain who works from home, including the following:[2]

- Women worked at home more (26 percent) than men (22 percent).
- Those with an advanced degree worked at home most (42 percent), followed by those with a bachelor degree (34 percent), those with some college (19 percent), high school graduates (16 percent), and those with less than high school (10 percent).
- Those with the highest incomes (above $1,620 per week) worked at home most (34 percent), followed by those making $1,001 to $1,620 per week (21 percent), those making $651 to $1,000 per week (12 percent), and those making $650 or less per week (10 percent).
- Occupations with the highest levels of working from home include management, business, and financial operations (37 percent); professional and related services (33 percent); and sales (24 percent).

Rates of telecommuting increased substantially during the COVID-19 pandemic of 2020. Nearly one-half of all workers were working from home during the pandemic, with the highest rates of work from home by those from the highest income group (Guyot and Sawhill, 2020).

Most recent household travel surveys already ask questions about a worker's regular workplace and ability to work from home on a part-time basis. These questions are critical to allow that estimates of a baseline work-from-home population can be calculated.

On the other hand, household travel surveys typically do not ask about the purpose of in-home activities. This is partly because it is often difficult for respondents to say specifically that they were working when they also may have been engaging in other nonwork activities during at-home periods. As a result, it is typically not possible to determine whether a worker worked from home on their travel day or did not work at all. Given the growing interest in work-from-home patterns, in the future, household travel surveys should be designed to capture at-home work activities explicitly, even if the duration of at-home work activities cannot be established accurately.

Mobile location data offers a new data source for transportation analyses. While these data are good for precisely identifying activity locations of a person over long periods of time, they do not provide contextual information. Home locations can typically be identified easily based upon

[2] Note that figures represent the percentage of workers that worked for any amount of time at home on a given day that the worker worked.

diurnal patterns of a given person and out-of-home work locations also can often be identified. However, a worker that telecommutes presents challenges to these data. For example, during the COVID-19 pandemic, when many more workers began working from home on a regular basis, it can be difficult to differentiate those working from home from the many workers that lost their jobs during this time. However, these data do hold promise in understanding the travel patterns of individuals that work part-time from home and part-time onsite. Moreover, these data could be used alongside survey data to help establish longer-term patterns among different groups of travelers.

Modeling Working from Home

In this section, we describe the steps that can be used to incorporate telecommuting in a trip-based model.

Data Preparation

In order to model work-from-home behaviors in a trip-based model, at a minimum, workers need to be split into three distinct categories:

1. Workers that work onsite on the travel day.
2. Workers that work from home on the travel day.
3. Workers that do not work on the travel day.

This may not be so straightforward since household travel surveys may not ask the reason a worker did not travel to work on the travel day. In that case, assumptions may be needed to infer which workers that did not work fall into the work from home category, and which fall into the did not work category.

In addition to these requirements on the household survey data, having detailed employment classification data also can be valuable. This is because certain sectors of employment are more likely going to be jobs where employees can work from home versus other sectors of employment. Therefore, the more detailed the employment classification that is possible, the better it will be for model development.

Model Estimation

In a trip-based model, the main modeling elements that could be adjusted are trip productions and attractions, where work from home behavior would reduce home-based work trip rates (and potentially increase home-based other trip rates). While the effects of work-from-home patterns could be estimated from survey data (if the survey samples are sufficiently robust) and tied back to the level of telecommuting activity, one key challenge is that the ability to work from home is not uniform across a region, and certain jobs are more likely to allow workers to telework. To account for these effects, one option for modeling work-from-home behavior in a trip-based model is as follows:

- In addition to any other household segmentation, households are segmented by whether any worker from the household works from home on the travel day.

- Trip generation rates are then estimated separately for these two categories of households. Note that telecommuting household segmentation across other variables may need to be collapsed to ensure reasonable sample sizes. While home-based work trip productions will be low for telecommuting households, home-based other trip productions may very well be higher.
- The production of nonhome-based trips for telecommuting households may deserve special attention since the geographic distribution of these trips may be different for telecommuting and nontelecommuting households based upon the locations of telework-eligible jobs in the region. Having a robust employment type classification scheme could be helpful in differentiating between households for nonhome-based trip productions.
- The new input to the model would be the level of telecommuting activity.

The segmentation of households by telecommuting categorization also may be carried through to the trip distribution and mode choice model components. Telecommuting households may have different distributional patterns for trips even after accounting for the differences in trip purposes resulting after trip generation. However, the trip generation elements of the model are likely where the biggest effects are likely.

Model Validation

This model can be validated in similar ways to the traditional approach. For the model base year, the telecommuting share of households that is input to the model need only be based upon the telecommuting levels observed in the base year. The survey can be further used to validate the types of locations that telecommuting and nontelecommuting households travel to.

If a historical household travel survey dataset for the region is available and the survey collected information about telecommuting that can be used to compute telecommuting rates, a full backcast of the model may be appropriate, using historical input data for transportation networks, employment, and population.

Sensitivity testing also is critical and can help add confidence to the telecommuting components of the model. The types of questions that can be useful to ask as part of sensitivity testing include the following:

- Are the differences in trip generation (and other model components) between telecommuting households and nontelecommuting households reasonable?
- What times of day do telecommuting households travel in? How far do they travel and for how long? How does this compare against nontelecommuting households?
- What are some of the limitations of the model sensitivities? How confident are we in the forecasts when telecommuting is very high? It is likely that, at some point, the forecasts break down. For instance, the timing of travel for telecommuters likely is somewhat related to congestion on the network. If telecommuting becomes quite high, this may reduce peak congestion so much that the time-of-day shifts forecast by the model are unreasonable.

Forecasting

Since the beginning of the COVID-19 pandemic in 2020, there has been a renewed interest in telecommuting behavior, largely because the pandemic forced millions of workers to being working from home on a regular or semi-regular basis. With such great uncertainty about the extent to which people will continue working from home following the pandemic, exploratory modeling provides an approach for systematically investigating how telecommuting levels may impact the transportation system.

In addition to the future levels of telecommuting behavior being highly uncertain, the mechanics of how telecommuters travel are not well understood either, even with the existence of survey data that reflects telecommuting patterns. Some of these features could be examined using exploratory modeling as well, for instance, adjusting the trip rates of telecommuting households up or down.

2.4 Getting the Right Data (and Using It Correctly)

As noted above, some of the biggest challenges associated with developing models to analyze both conventional travel and emerging mobility options have to do with getting the right data. It is difficult to get data on new types of travel, for which there may be only a few years of usage to examine (or nothing, in the case of modes such as CAVs, which are not yet being used). And much of the data on new travel mode usage is not publicly available. At the same time, however, there are newer data sources that had not been available until recently, and both the amount and accuracy of the data are beyond what planners have had access to in the past.

Model Input Data

The main inputs of travel demand models are socioeconomic and land use data and transportation networks. The best sources for model input data are well established and may be legislatively mandated in some jurisdictions (for example, the need to use U.S. Census data or State-generated employment data). The main concern for analysts is understanding and accounting for the errors associated with these data sources, especially forecasts, as described in section 2.5.

Survey Data

There is a wealth of information available on the collection and validation of travel survey data; perhaps, the most comprehensive source is the *Travel Survey Manual*, first published by FHWA in the mid-1990s, and later updated by the TRB Standing Committee on Travel Survey Methods (2010). Travel surveys are a unique source of some of the information that is valuable for developing and validating models, although household travel surveys can cost millions of dollars and can take years to plan for, collect data, and process the data. For many analyses, household surveys are by themselves insufficient to provide the necessary data; for example, if transit demand is part of the analysis, transit rider surveys are a critical resource in identifying how transit is used and who is using it. Additionally, as discussed in section 2.5, there are a number of sources of error in survey data that must be considered when using such data.

Data on the Amount of Travel on Transportation Facilities

The data available on the amount of travel on transportation facilities are used in model validation to help confirm that the model produces accurate estimates of travel for the base year scenario, where observed data are available. These data consist mainly of vehicle traffic count data (often segmented by vehicle type and time of day) and, in situations where transit is being considered, ridership counts at the route level and station boarding counts. The data available in a particular region is highly dependent on the data collection programs undertaken by the States and other agencies involved. As with other data sources, it is important to consider and account for the errors associated with these data sources, as discussed in section 2.5.

Big Data

"Big data" for measuring travel demand have revolutionized how we think about travel forecasting, because these data can be collected more regularly and with larger sample sizes than is possible within the construct of traditional survey data collection (which is the typical means for collecting data on travel demand). Significantly, most big data sources are collected passively, meaning that they do not rely on responses from the travelers and, therefore, do not have associated response bias. There is a number of ways to effectively use big data to support travel modeling in terms of model development and validation. There also are several sources of big data that can be used in different ways, including the following:

- **Mobile Location Data.** Location-based services (LBS) data fall into this category because these data are collected from mobile devices that individuals carry with them. These data are typically marketed for use in generating origin-destination (O-D) trip tables, but can be useful for any application that requires identification of trip ends. These data are typically not robust enough to decipher the trajectory of movement from origin to destination, at least not on a widespread basis.
- **Connected Car and GPS Tracking Data.** These data typically come from transponders inside personal automobiles and can be used to track the trajectories of trips from origin to destination. These data are often used to develop speed estimates on major roadways and infer typical routes used between O-D pairs.
- **Truck GPS Data.** These data also come from transponders inside vehicles, but in this case, those vehicles are trucks. These data can be provided as O-D trip table form or can be provided as disaggregate vehicle traces.

These data types are described in more detail below.

Mobile Location Data

The most common type of mobile location data currently available comes from LBS. These data are collected by applications running on mobile devices using the embedded GPS technology. Some of the properties of these data include the following:

- **Sample Size:**
  - The raw data are massive in size relative to more traditional sources of travel behavior data, such as surveys, with penetration rates in the tens of millions (if not hundreds of millions) of devices nationwide.
  - Once these datasets are processed and filtered, device sample sizes are typically smaller since some percentage of the data collected by LBS are not usable for inferring travel. While algorithms to process the data vary from vendor to vendor, processed LBS data typically include a set of devices representing at least 5 to 10 percent of the population, which is about an order of magnitude larger than most surveys (where household travel surveys typically recruit a sample of less than 1 percent). LBS datasets often include many days of information for each device, making samples even larger relative to household travel surveys, which typically only collect a few days of travel information from respondents.
- **Passive Data Collection:**
  - Unlike surveys, LBS data are collected passively from mobile devices. As a result, travel episodes are not directly observed and must be inferred from the geographic and temporal data that are directly collected.
  - Contextual data about travel are not directly observed (e.g., mode, purpose, travel party, demographics). Some of this information can be inferred or estimated based upon repeated observations of the same device, or based on fusion with other data sources.
  - Ban et al. (2018) note that these datasets are not generated via rigorously designed processes (as surveys are); and as a result, the representativeness of the generated sample cannot be guaranteed.
- **Persistence of Device Identifiers:**
  - Device identifiers are persistent for weeks or months, which allows for reasonably accurate and precise inference regarding home and work locations.
- **Data Quality:**
  - LBS data are collected at irregular intervals that depend on the usage patterns of the device user. As such, the quality of the data collected varies from one device to another.
  - Low-quality data are filtered using methods that vary from vendor to vendor.

As noted above, specific methods are needed to infer travel episodes from these data since travel is not observed directly. Data processing algorithms are needed to identify trip ends and filter erroneous or incomplete records (A number of such algorithms have been published in the literature (e.g., Alexander et al., 2015; Widhalm et al., 2015; Wang and Chen, 2018; Cambridge Systematics, Inc. and Massachusetts Institute of Tehcnology, 2018; and Lemp et al., 2019.).

LBS data are typically used to generate O-D trip tables. Since data are collected from mobile devices, they are generally considered to represent an estimate of total trips in a region, including trips by all modes and by residents and visitors. Some of the key uses of these data are as follows:

- **O-D Trip Patterns.** While traditional data sources, like travel surveys, can provide reasonably accurate estimates of average trip lengths and trip length distributions, these data sources have sample sizes that are too small to provide robust information about O-D trip patterns in most cases. Because of the much larger sample sizes obtained from LBS data, they provide much better spatial and temporal coverage than survey data (Adler et al., 2017). The enhanced resolution likely means that the O-D trip pattern estimates derived from the data are more robust than those developed from surveys, and these O-D estimates can be used for model calibration or other purposes.
- **Visitor Travel.** Traditionally, visitor travel is either ignored in regional travel models or modeled in very limited ways, possibly based on inadequate survey data. Visitors to a region can be easily identified in LBS datasets based upon having a home location outside the region. As a result, LBS data can be used to estimate the relative levels of trips generated by zone, as well as the distribution patterns of visitor trips.
- **Seasonality and Day of Week.** LBS datasets can provide estimates of travel patterns for different days of week and seasons of the year, which are useful information for certain applications. Measuring differences in travel patterns for different seasons or across days of week can provide a benchmark from which to estimate differences in key metrics like VMT and VHT. These data also can be used to support specialized models in areas with heavy seasonal tourism and weekend visitors. Such information is not available from surveys with specific (usually short) timeframes for collection of data from specific travelers.
- **Travel Variability.** LBS datasets can be used to understand how people’s travel patterns vary from day to day. Since individual devices can be tracked over weeks and even months, a clearer picture of travel pattern variability at a disaggregate level can be gleaned. Note that sometimes only aggregate trip-level information is available.
- **Model Update Frequency.** In comparison to traditional survey data, LBS data have a lower cost, are collected continuously, and are readily available for consumption. Household travel surveys can cost millions of dollars and can take years to plan for, collect data, and process the data. The long lead time often means that the data are several years old before they are ever used to update the regional model. LBS data, on the other hand, can be acquired at a fraction of the cost and with a very short lead time of a few months or less. This allows for more frequent updating of the data used to support modeling and using more recent data. While LBS data may not offer the full gamut of contextual information offered by surveys, other key metrics of travel can be updated, including the following:
  - Trip rates by geography.
  - Temporal distributions.
  - Trip lengths.
  - Mix and share of trip purpose.

It also is worth noting that these data need not be used in isolation. Techniques that use these data in tandem with other more traditional datasets, like Census data products, traffic counts, and household travel surveys, could provide even more value.

While these data hold promise for a variety of transportation planning and modeling applications, they are not without limitations. Users of these data need to understand these limitations to ensure these data are not used in ways that are not appropriate. Some key challenges of these data are as follows:

- **Data Expansion.** As noted by Ban et al. (2018), LBS data may not be representative of the population. Moreover, these data still represent a sample of overall population. Therefore, data expansion is a critical component of using LBS data for any purpose. Expansion processes adjust the LBS sample of device home locations (and potentially work locations) to match the true population. This is important because biases may exist from zone to zone and across area types. Data expansion is typically accomplished using Census demographic information.
- **Demographics.** Even when the data are expanded, there are potential issues related to the representativeness of the sample. Certain demographic biases may exist with the data, which cannot be controlled for in the expansion process since demographic information is not available. Key groups that may be underrepresented include the following:
  - Children less than 10 years of age who typically do not carry mobile phones.
  - Older individuals who, on average, likely have lower mobile phone usage and lower rates of app usage.
  - Lower income individuals, who may not have access to mobile phones.
  - Individuals concerned about personal data privacy, who may use their mobile phones in different ways.
- **Other (potential) biases.** Vendors of raw LBS data typically will not share information about the apps from which data are collected. While most of these data vendors can ensure that data are coming from many tens or hundreds of different apps, the uncertainty associated with the collection of apps offers potential for bias which cannot be checked (Adler et al., 2017). In addition, it is unclear whether general app usage has any correlation with travel patterns. For instance, potential biases could emerge if some people that are particularly concerned with privacy (who might do things to mask the ability of apps to track their data) travel differently than others.
- **Validation.** Direct validation of these data requires an apples-to-apples comparison of trips inferred from raw LBS data against actual trips performed by individuals in a region. Adler et al. (2017) used GPS-collected survey data and matched respondents to devices in an LBS dataset. They found that LBS data underrepresents certain trips, especially shorter duration ones, though they also note that improvements to the trip inference algorithm may be able to mitigate these issues in the future.

Besides direct comparisons, some attributes of these data can be easily compared against other more traditional data sources (such as surveys) to validate the data (Adler et al. 2017;

Cambridge Systematics, Inc. and Massachusetts Institute of Technology, 2018; Ban et al., 2018). It is worth noting that errors in the inferred data may be masked when comparing to more aggregate statistics (Ban et al., 2018).

On the other hand, many of the ways in which these data provide value is by providing cheaper and more robust estimates for key travel pattern characteristics for which we often have very limited information. For instance, often the only other source of O-D trip pattern data is a household travel survey, which, as described previously, suffers from small sample sizes, making O-D trip tables at levels important for modeling (i.e., transportation analysis zones) very sparse. Visitor travel also typically has limited data available from which to validate because data on visitors often is not collected. It is important to appreciate that these data cannot be validated along every dimension.

- **Contextual information.** Contextual information such as demographics and key travel characteristics are not available from the data. Travel surveys will remain an important source of data for connecting context to travel patterns.
- **Differences among data vendors.** Each data vendor uses its own unique set of algorithms (which are generally not made available to data users) to process and compile trip pattern information. Since each one is unique, differences can emerge that may mean certain limitations exist for one dataset that do not for another. Furthermore, given these data are still quite new, algorithms continue to change to improve inference processes and add new features. In this continually evolving environment, it becomes quite challenging to make any definite judgments about these data more generally. The best advice is to evaluate the dataset along the dimensions that are important for its use case to ensure the data provide value.

Connected Car and Global Positioning System Tracking Data

GPS tracking data are fundamentally different from LBS data because they are collected at frequent and regular intervals (e.g., every few seconds), which makes these data appropriate for tracking the route used for a trip. However, these data are generally collected by devices located in autos and trucks, and so these data cannot measure travel by other modes, such as transit or bicycle. Some features of these data are as follows:

- Each device identification (ID) in the dataset represents a vehicle, rather than a person (as in the case of LBS data). Device IDs tend to be updated much more frequently than is the case for LBS data, meaning a vehicle cannot be tracked over days or weeks. In fact, in some GPS datasets, the vehicle IDs are updated after each trip.
- The GPS precision of these data is more precise than LBS data, allowing for individual data points to be snapped to roads and even lanes on the road.
- Data collection is much more frequent and occurs at regular intervals. Often data are collected at intervals of 60 seconds or less. This allows for tracking the path used by a vehicle in addition to the trip origin and destination.
- The sample of devices included in any given dataset often comes from a small sample of vehicles. These vehicles either have onboard navigation systems or manufacturer-installed

connected car systems. These vehicles tend to be newer and biased toward more luxury vehicles, which results in a sample that is not representative of the set of vehicles actually on the roads. As a result, expansion is really not possible with these data, and it is important to understand the sample effects for any application using these data.

Chen et al. (2017) provide more detailed information about the key properties of these types of datasets, including spatial and temporal properties of raw GPS data, as well as key properties of trips inferred from these data.

These data are often used to support specific applications rather than calibrating or updating a full regional model. For instance, these data can be valuable for providing O-D information for specific links on the network (i.e., select link analysis). They also are used for various types of corridor and subarea analyses.

The main way in which these data can be used to support model development is through travel time and speed information from the data that can be aggregated at the link level. Speed and travel times at the link level that are output from the model can then be compared against observed speeds and travel times by time of day. These data also may be used to calibrate volume-delay relationships and identify free flow speeds for individual facilities.

Truck Global Positioning System Data

Truck GPS data are similar to GPS tracking data, except that they are specific to commercial vehicles. These data have GPS spatial precision and are collected in frequent and regular intervals, which allows for both identification of trip ends and routing. Unlike LBS and connected car data, however, truck GPS data are often available at a disaggregate level, allowing the user to have access to individual ping location data.

It is important to note that different sources of truck GPS data provide travel patterns for different types of trucks. For instance, American Transportation Research Institute (ATRI) data predominantly come from heavy trucks (FHWA classes 8 to 13 or combination units), while INRIX data offerings include medium trucks (FHWA classes 5 to 7) and light trucks.

These data can be used to identify truck trip ends and build O-D truck trip tables. However, expansion is often a challenge as there are typically not good bases for the expansion process. Expansion is often performed by a factor method or Origin-Destination Matrix Estimation (ODME) process; both of which aim to align the O-D patterns more closely with truck counts.

These data also are often used for model development. By utilizing land use and employment data along with travel impedances, trip generation and distribution models of truck trips can be related to key variables. Expansion is again important, particularly for calibration of trip generation rates.

2.5 Validating and Testing the Models

Model validation has been recognized as a topic of great importance for several decades. FHWA first published the Model Validation and Reasonableness Checking Manual in 1997 and published an update in 2010 (Cambridge Systematics, Inc., 2010). There have been a multitude of research efforts on model validation focusing on specific modeling topics over the years.

An important focus of model validation is to ensure that the model produces useful results for the planning analyses that it is used for. Certainly, this includes checks of how well the model can represent existing conditions by comparing base year scenario results to observed data (though there has recently been greater recognition of the error associated with the observed data). The importance of checking not only the final model results, but interim results of individual components to ensure that there are not offsetting errors in different components, has long been understood, as discussed in both versions of the FHWA model validation manual.

The importance of model sensitivity testing also has been recognized since at least the 1990s. While models have long been used to estimate the impacts of new ways of traveling in a region (e.g., new types of transit, toll roads in regions where tolls had not previously existed), the recent emergence of new types of travel technology and behavior have greatly increased the importance of ensuring that models are properly sensitive to changing conditions. This increased focus on model sensitivity needs to consider the greater uncertainty in modeling things that we have little information about.

Model sensitivity testing originally consisted of changing values of individual input data items or model parameters, and examining how much model results changed. For example, the sensitivity of mode choice to transit fare could be tested by increasing fares for some or all of the transit system by a fixed amount or percentage and seeing how transit demand changed. Later, analysts designed more sophisticated sensitivity tests, especially for more complex model types, such as activity-based models. For example, asserting a traveler age distribution that skews older than the existing population could be used to assess how well the model is able to capture the effects of an aging population on a variety of travel demand results, including activity purposes, time of day, and mode choices. Section 3.3 discusss how TMIP-EMAT can be utilized to perform a systematic and comprehensive sensitivity analysis.

The emergence of new travel modes and travel behaviors requires that the use of more sophisticated model sensitivity testing procedures be increased to properly assess the effects of these issues. The greater uncertainty associated with these issues means that analysts need a greater understanding of the sensitivity to the assumptions made to model mobility options on which there is little or no information. The greater level of sensitivity testing should extend beyond changes in model inputs, such as travel times, to include model parameters themselves. For example, testing the sensitivity of the greater freedom to perform different activities in a CAV requires testing the sensitivity of the model to the in-vehicle travel time parameters.

Analysts need a greater understanding of the sensitivity to the assumptions made to model mobility options on which there is little or no information.

2.6 Recognizing Uncertainty in Travel Models

As can be seen from previous section, there is a lot of uncertainty associated with emerging mobility options. This section discusses how to recognize uncertainty in travel models, and how modelers can incorporate the uncertainties into the planning process.

Uncertainty associated with travel models can be categorized into four types of uncertainty, as illustrated in figure 10:

1. Measurement uncertainty associated with model inputs.
2. Measurement uncertainty associated with model assumptions.
3. Forecast uncertainty associated with model inputs.
4. Forecast uncertainty associated with model assumptions.

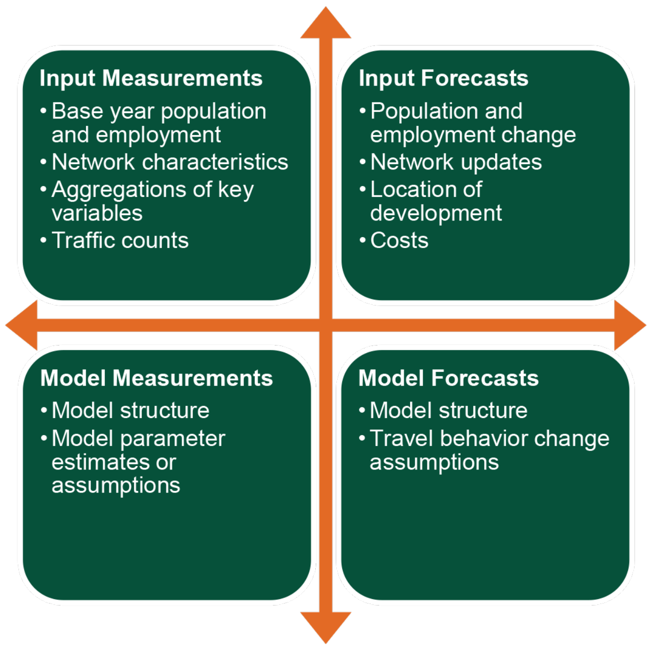


Figure 10. Diagram. Dimensions of uncertainties in models.

(Source: Federal Highway Administration.)

Measurement Uncertainty Associated with Model Inputs

Measurements associated with model inputs include information about base year inputs to the model that can be directly measured from what existed at the time of the model base year, including parameters of the model estimated from observed data. While the ability to measure these inputs typically means there is a lower level of uncertainty associated with measurements than forecasts, measurement error still exists. Uncertainty can be in the form of error in direct measurements, but also can take the form of aggregation error, where a group is assumed to be homogenous when it is not. Sources of uncertainty that fall into this category include the following:

- **Base year population and employment.** Measurements of population and employment can suffer from error for a number of reasons. Population data often comes from the Census, which is considered the gold standard, but the American Community Survey (ACS) samples only a small percentage of the population, which can lead to deviations actual versus measured population totals. Employment estimates often come from economic databases that also are based on survey data.

- **Network characteristics.** While agencies often try to maintain detailed and accurate representations of the highway and transit infrastructure, this is not always possible. Characteristics of the network typically are defined on the basis of models of traffic flow, which are themselves abstractions of reality; network links assigned common classifications are assigned common attributes; and low-level highway links (e.g., local roads) typically are absent from the network and replaced with centroid connectors meant to provide representative travel characteristics of trips accessing the network that is coded. Also, networks do not consider every characteristic of the systems facilities they represent; for example, highway network links typically include the "number of lanes," but usually do not explicitly include the presence and length of turning lanes.
- **Aggregations of key variables.** A number of inputs to travel models could be represented as having more disaggregate characteristics than what is actually represented. Model inputs such as employment are one example, where travel models typically do not distinguish between different types of employees (e.g., full-time, part-time, etc.). The zone system typically used by the model is another example where the exact origins and destinations of modeled trips are lost, and instead represented by an average centroid location for each zone.
- **Traffic counts.** While traffic counting *technology* may not suffer from much uncertainty, many average day counts are based on single or two-day counts that are factored to represent an average day at that location. These factors will vary from location to location by different and unmeasured amounts, resulting in noise in the underlying data. Furthermore, any particular day may have unusual traffic as a result of factors that are uncontrolled. While counts are not a direct input to the model, they are used to calibrate and validate the model in the base year.

Measurement Uncertainty Associated with Model Assumptions

Measurement uncertainty associated with model assumptions deals with the structure and parameters of the model itself. Measurements of the model typically come from estimating or calibrating model components using observed data, such as travel surveys or counts. In such cases, uncertainty exists about the true value of parameters in relation to statistical estimates. In addition, assumptions are made about model structure and parameters, and these assumptions also serve as sources of uncertainty in the model.

As discussed in section 2.2, there are several types of travel models that typically are used for different purposes, including sketch models, trip-based models, and activity-based models. These models use different assumptions about the mechanics of travel behavior. For instance, activity-based models typically are designed as collections of choice models that estimate travel behavior for individual travelers, and are built on assumptions of random utility theory. Therefore, there is measurement uncertainty surrounding the embedded assumptions defined in the type of travel model utilized.

Forecast Uncertainty Associated with Model Inputs

This category includes the forecasts of inputs to the model when it is being used to analyze future scenarios. Most of the uncertainties typically considered when developing travel forecasts fall into this category. In some cases, these can be considered policy variables (e.g., when different networks assumptions are tested). Forecast uncertainties typically considered include, but are not limited to, the following:

- **Location and intensity of population and employment.** Population and employment are key inputs to travel models, as noted above. Forecasts of population and employment can be developed on the basis of land use models, or on the basis of community plans or visions. Either way, a great deal of uncertainty is associated with future land use, which can have impacts on travel patterns.
- **Future networks.** Roadway and transit networks also are key inputs to travel models. Since future infrastructure can be more carefully planned and set forth based upon policy, these inputs are often viewed more as policy inputs, rather than sources of uncertainty. However, uncertainty also may exist depending on the specifications of the analysis. For instance, toll roads operated by private companies may have the ability to set future tolls in different ways; and from a public agency perspective, this could be considered a source of uncertainty.
- **Costs.** Forecast year auto operating cost, parking costs, and transit fares may vary from base year input values.

Forecast Uncertainty Associated with Model Assumptions

As discussed in section 2.2, different types of travel models are more or less suited to answer particular planning questions, and even within each type of travel model different assumptions within the model structure and model parameters affect the uncertainty of the forecasts. For instance, the components that are considered in a mode choice model may include a collection of different attributes (such as time and cost), but may ignore other attributes that are actually important to the choice context (e.g., reliability, safety, etc.). These also are structural decisions about what attributes are important and what role they play in travelers' decisions. Travel behavior uncertainties deal with the assumptions about how travel behavior will change (or remain the same) in the future. Simple examples of these uncertainties, as propogated in travel models, include assumptions regarding price elasticities, values of time, and trip rates.

It is important to keep in mind the inherent uncertainty in travel models when using them to support transportation analysis. The discussions on travel models and travel modeling-related data are not intended to provide indepth knowledge, but to highlight some of the important considerations to encourage good modeling practices. More information on travel model development, some of the recent research on emerging data and its limitations can be found on at https://www.fhwa.dot.gov/planning/tmip/publications/other_reports/, and the references in this report. However, it is important to keep in mind that emerging data are rapidly evolving.

Section III

Conducting Exploratory Modeling and Analysis Using TMIP-EMAT

3.0 Exploratory Analysis Using Travel Model Improvement Program-Exploratory Modeling and Analysis Tool

TMIP-EMAT is not a transportation model in and of itself. It is a utility tool that enables an analyst to use the region's transportation model for exploratory analyses. EMA has been used by planners to better understand systems with deep uncertainty by calibrating models that explain the system, where some inputs to the system have uncertainty associated with them, there are various policies or levers available to a decisionmaker to affect the system, and there are various outputs of the system which are of interest. An EMA research methodology explicitly treats the computational experiment (i.e., model) as a set of assumptions and hypotheses and aims to explore the impacts. This differs from treating the model as a predictive tool that is an accurate surrogate to the real world (Bankes, 1993).

TMIP-EMAT is designed as a tool to engage stakeholders and policymakers in discussions around developing effective policies and facilitating discussions throughout an iterative and continuous planning process. With TMIP-EMAT, analysts, stakehoders and policymakers can explore key relationships between model inputs and outputs using interactive tools, study the range in outcomes to highlight that uncertainties exist in these relationships, and use the results to inform a robust decisionmaking approach. You can find the program code and the technical documentation at https://tmip-emat.github.io.

In this chapter, the steps required to carry out an analysis using TMIP-EMAT are described. The results of a TMIP-EMAT analysis are presented, including the interpretation of results and how those results can be used to inform policy.

3.1 Scoping for Exploratory Modeling Analysis

The first step in an exploratory modeling analysis is the scoping step, which defines the goals and objectives of the analysis, as well as more specifically how those goals and objectives will be explored using the model. Several key components are necessary in the scoping step:

1. **Develop High-Level Scope Goals.** The first step in the scoping process is to define the goals and objectives for doing the analysis and translate those into to policies that could support those goals and identify the uncertainties that could affect meeting the policies' ability to meet the goals. These goals need not be overly specific, but should define a set of high-level objectives to explore using the model.
2. **Identify Model Functionality.** In this step, the scoping goals are matched against the inputs and sensitivities of the core model along with potential performance measures. Inputs and sensitivities include both exogenous inputs, or uncertainty variables, that policymakers have little or no control over (e.g., fuel price) and endogenous policies, referred to as policy levers (e.g., tolls or transit fares). Performance measures are outputs of the model that can be used to assess how well goals and objectives are met based upon the set of policy options that are explored during the exploratory modeling analysis.

3. **Finalize Scope.** The set of uncertainties, policy-levers, and performance measures is assessed on the basis of priorities, as well as the functionality of the model and then finalized. There is a number of limiting factors that may prevent the full set of desired uncertainties, policy levers, and performance measures to be included within the final scope.

High-Level Scoping

The first step of the scoping process is a high-level scoping exercise, which is intended to get agency planners and modelers thinking about the goals and objectives of the analysis and what sets of policies, strategies, and uncertainties would be of interest in relation to those goals and objectives. Rather than limiting the exercise based upon the capabilities of a specific travel model, this step involves thinking more broadly about the goals and objectives.

In the context of TMIP-EMAT, a goal is what the community is trying to accomplish. Some examples of common goals used in the TMIP-EMAT process include the following:

- Increase transit ridership.
- Reduce congestion and improve reliability.
- Understand telecommuting effects on sprawl.
- Improve economic opportunities and access in economically disadvantaged neighborhoods.
- Reduce auto usage in the suburbs.
- Understand future impacts of automated vehicles.
- Improve financial stability of transportation system through tolls/managed lanes.
- Improve freight movements.

The specific goals of the analysis help to define other parameters of the analysis as well. For instance, in some cases the goals dictate the examination of a set of policies in conjunction with uncertainties, while in other cases, the analysis may be purely exploratory, examining effects of uncertainty only.

In some cases, the goals dictate the examination of a set of policies in conjunction with uncertainties, while in other cases, the analysis may be purely exploratory, examining effects of uncertainty only.

The choice of goals impacts this decision. For instance, if a goal is to reduce congestion, it would make sense to formulate policy levers that have a chance of impacting future levels of congestion (e.g., pricing, highway expansion project, improvements to transit service, or other congestion management policies). Often multiple policy levers are considered that may each have different impacts on the set of goals. Conversely, if a goal is to better understand the impacts of automated vehicles, then policy levers may not be of interest since there is no clear objectives for impacting the system performance. In this case, an uncertainty driven analysis may be warranted.

Identify Model Functionality

Once the high-level scoping exercise is complete, it is necessary to identify the specific set of uncertainties and policy levers that will be tested and define how the uncertainties and policy levers are defined within the travel model. We define uncertainties and policy levers as follows:

- **Uncertainties.** Uncertainties are factors outside the decisionmaker's control that may have an effect on performance measures and may help or hinder the ability of policy levers to reach stated goals. Some examples of uncertainties are as follows:
  - Fuel prices.
  - Values of time.
  - Land use and demographics changes.
  - Impacts of automated vehicles.
  - Behavior-related sensitivities of the model.
  - Telecommuting levels.
- **Policy Levers.** Policy levers are factors within the decisionmaker's control that are implemented to help reach a defined goal or goals. Examples of policy levers include the following:
  - Highway capacity expansion.
  - Transit service changes.
  - VMT charge, tolls, and managed lanes.
  - Parking pricing.
  - Travel demand management strategies.
  - User-specific travel cost changes (e.g., transit subsidies for low-income population).
  - Land use policy (e.g., infill development).

As part of identifying the uncertainties and policy levers to include in the analysis is translating the policy levers and uncertainties from high-level concepts (e.g., impacts of autonomous vehicles) to adjustable model inputs or parameters (e.g., allowable vehicle spacing or capacity on a highway). The representation of each uncertainty or policy lever may include a combination of input variables and parameters to appropriately represent the lever or factor within the model.

An important component of this exercise is analyzing the model functionality. For instance, if the existing model does not have a mode choice component, testing policies around the impacts of added transit service would require large changes to the underlying core model to be tested effectively. Conversely, a model that includes auto operating costs as an input can easily test the impacts of changes to fuel prices. Examples of matching specific goals with policy levers and uncertainties are given later in this section.

In addition to the selection of uncertainties and policy levers, the scoping process also must define the range of each input to the model. Since these represent the inputs to the model that will be varied in the analysis, careful attention to these ranges is important as they will drive the results of the analysis. Some policy levers are simple binary variables—either the policy is enacted or not. Other policy levers and uncertainty variables are continuous variables that require careful consideration and review of other sources to arrive at a reasonable set of ranges.

Some policy levers are simple binary variables—either the policy is enacted or not. Other policy levers and uncertainty variables are continuous variables that require careful consideration and review of other sources to arrive at a reasonable set of ranges as they will drive the results of the analysis.

The scoping exercise also must identify key performance measures that can be used to evaluate the efficacy of policies and the extent to which goals of the analysis are achieved. Again, performance measures should be tied to the specific goals of the analysis, and the performance measures should be outputs from the core model (or metrics that can be derived from the results of the core model). Examples of performance measures that might be used in an exploratory modeling analysis include the following:

- VMT (by vehicle class, speed, area, etc.).
- Transit ridership.
- Mode share.
- Travel times for specific roadways or corridors.
- Congestion and reliability measures.
- Economically disadvantaged population measures, such as travel times or accessibility.
- Revenue generated.

In practice, one of the main limiting factors in the scoping process is the functionality of the existing core model. Policymakers and planners desire to evaluate uncertainties and policies that the travel model cannot appropriately model. This suggests a need to adapt or update the existing travel model to better handle these uncertainties or policies. These updates can either happen during the exploratory modeling process or as part of a larger modeling development effort.

Often, the budget and schedule mandate model enhancements and sensitivity changes are done in a piecemeal manner that do not comprehensively account for specific uncertainties and/or policy levers. So, while it is not uncommon that changes to the model are performed during exploratory modeling, limiting the extent of these changes is desirable.

It makes sense to consider the exploratory modeling process during a model update effort so that the types of variables that are anticipated for exploratory modeling can be included in the model in the most realistic possible ways and done so efficiently. In some cases, this may mean adding functionalities that are not used at all in the base year model (e.g., CAVs). It also may mean adding sensitivities that are not fully calibrated, but allow for testing of specific policies or uncertainties.

Finalize Scope

An important component to the process is performing sensitivity tests to isolate how individual policy levers and uncertainties impact the key performance measures being considered.

An important component to the process is performing sensitivity tests to isolate how individual policy levers and uncertainties impact the key performance measures being considered. Performing such tests is important to do before finalizing the scope for the analysis. If an uncertainty or policy lever, as it is coded in the model, has little impact on the performance measures of interest, it may not be worthwhile including in the analysis since there is an opportunity cost associated with each policy lever and uncertainty. For each policy lever and uncertainty that are included, the experimental design will typically increase by about 10 full model runs. By determining when a policy lever or uncertainty is unimportant, the analysis can either replace that policy lever or uncertainty with another more impactful one, or remove it and reduce the number of full model runs required.

The final number and set of policy levers and uncertainties that can be included in the application is dependent on the chosen core model(s) run times, the computer resources available, and schedule constraints for project analysis. As a rule of thumb, for meta-model development, 10 core model runs need to be run for each uncertainty and policy lever. Therefore, if there are 4 uncertainties and 4 policy levers, then the core model will need to be run 80 times. In addition, other model set-up constraints, such as number of individual highway networks that need to be coded, may dictate the number and set of policy levers that can be feasibly developed.

Scoping Examples

The following provides some specific examples of how the scoping process works. Two examples are provided to illustrate the approach.

Example 1

As a first example, consider the following goal for a hypothetical exploratory modeling exercise:

- Examine the long-term vitality of transit in the region to serve various segments of the population under different transit-focused policy initiatives.

In this case, policy levers should be included in an exploratory analysis approach. Policy levers that might be considered include the following:

- Transit expansion to improve accessibility by transit to more of the region’s population.
- Transit fare reduction for the entire population or subsidies offered to lower income populations to encourage transit ridership.
- Increase service frequencies under the existing transit system to encourage transit ridership.

- Parking price increases, especially in downtown areas, to encourage people to shift away from auto modes and to transit.
- Elimination or reduction of parking costs at park-and-ride stations to encourage transit ridership.

In addition to policies selected for evaluation, it is necessary to consider the uncertainties that may affect the "vitality of transit" and the uncertainties that may affect the ability of the selected policies to help achieve long-term transit vitality. In this case, the impacts of automated vehicle adoption could play a role in how competitive transit is in the future, and thus, may be an important set of uncertainties to consider. Other potential uncertainties in this case may include fuel price, which directly impacts the competitiveness of auto modes; telecommuting levels since transit services often cater more to work travel than nonwork travel; and sensitivities of the model to costs like parking prices and transit fares.

The performance metrics that are included should help measure the "vitality of transit." A number of transit-related performance measures may be relevant, such as ridership by transit mode; travel times by transit, especially to/from downtown and/or other major activity centers; transit accessibility for economically disadvantaged populations and the population as a whole; mode shares; and transit revenues. Other performance metrics also could be considered (like VMT or VHT), but may be less relevant to the specific goal associated with this example.

It should be noted that the goal as stated, "Examine the long-term vitality of transit in the region to serve various segments of the population under different transit-focused policy initiatives," suggests an exploratory analysis approach. However, if the goal was to quantitatively determine which policy lever best met the goal of increasing transit ridership and revenue, then a risk analysis approach may be more appropriate. A risk analysis allows for a probability distribution, and thus, a confidence value to be placed on the performance measure outputs, allowing for an analyst to state, with statistical confidency, that one policy lever produces higher transit ridership or revenue than another policy lever.

Example 2

As a second example, consider a case where the goals are as follows:

- Reduce congestion.
- Maintain current costs of travel.
- Consider the impacts of automated vehicles.

In this case, policy levers would clearly need to be included. This example lends itself nicely to an exploratory analysis rather than a risk analysis, since automated vehicles can be considered a "deep uncertainty" for which it is difficult to put an a priori probability distribution around the uncertainties associated with automated vehicles. Policies should be focused on those that can reduce congestion while maintaining the current costs of travel. Transit-related policies may be one avenue here—for instance, expanding transit and/or reducing transit fares. Other policies may include travel demand management policies like encouraging telecommuting behaviors, or more traditional policies such as highway capacity expansion.

In this case, there is a clear directive to consider the impacts of automated vehicles. Uncertainties related to automated vehicles would include the ways in which at different levels of market penetration it is anticipated these vehicles will impact travel behavior, including changes to highway capacity, value of time sensitivities, and changes to parking behaviors that may impact the amount people pay to park, among others (see earlier discussion of CAVs for more details).

Key performance measures of interest in this case would include metrics related to congestion like VMT, VHT, delay, or corridor travel times, and those related to travel costs.

3.2 Interfacing between Travel Model Improvement Program-Exploratory Modeling and Analysis Tool and the Core Model

In order to perform the model runs of the core model efficiently and effectively, it is important that the core model be interfaced with TMIP-EMAT. This is done by developing an Application Programming Interface (API) that links to TMIP-EMAT, while also providing programmatic control of the core model, as shown in figure 11.

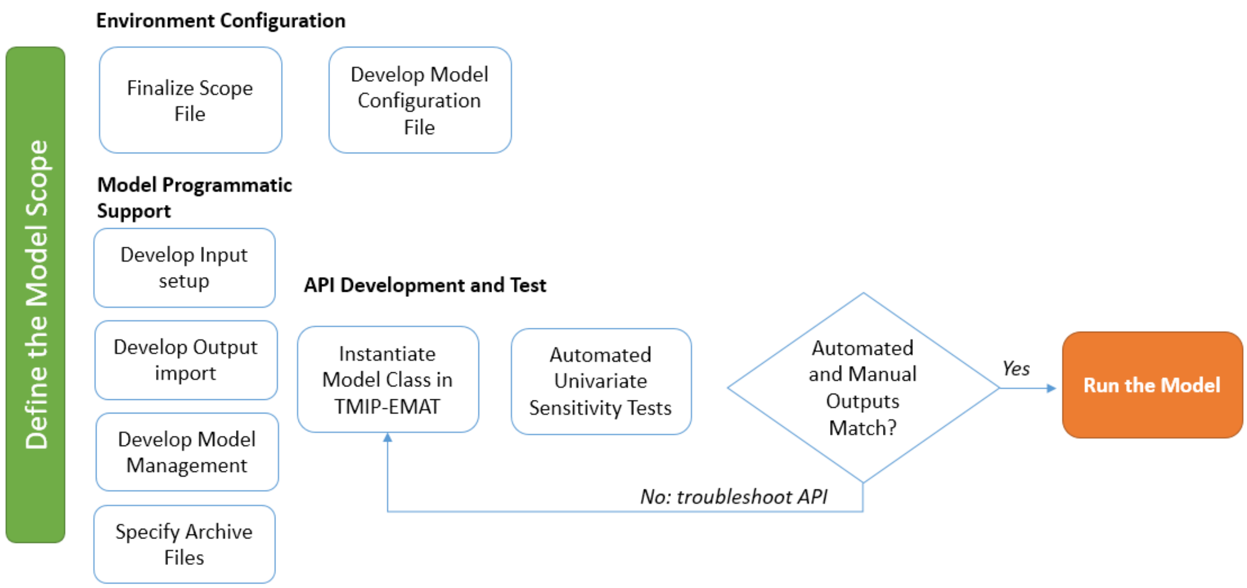


Figure 11. Diagram. Travel Model Improvement Program-Exploratory Modeling and Analysis Tool deployment steps.

(Source: Milkovits et al., 2019.)

The TMIP-EMAT interfacing and API development process is described in detail in section 2.4.2 of the Beta Test Report (Milkovits et al., 2019). Some of the key elements include the following:

- **Environment Configuration**, where the Anaconda environment is set up, and the scope file is finalized for use in the API development.

- **Model Programmatic Support** involves developing code that can initialize a core model run, set the levels of input variables, launch and run the core model, and generate metrics from the model that can be parsed by TMIP-EMAT.
- **Python Development** consists of developing scripts within TMIP-EMAT API that connect to the core model and results outputs.

3.3 Running Experiments

Exploratory Modeling versus Scenario Analysis

Both exploratory modeling and its more limited cousin, scenario analysis, involve running a model a number of times with different input values and evaluating the results. The different sets of input values used comprise a design of experiments. From the perspective of actually running the underlying model, the main differentiators between traditional scenario analysis and the exploratory modeling process are: 1) the breadth of the design of experiments, and 2) how those experimental designs are constructed.

Scenario analysis involves the development of different types of scenarios. They may be point estimates of a specific policy alternative that is being tested, they may involve a "representative" set of model inputs that reflect different model settings and policies, or they may involve extreme scenarios that serve to provide an estimate of lower or upper bounds on a specific forecast. More often, however, scenario analyses vary only one or two model inputs at a time. It also should be noted that sensitivity analysis is a form of scenario analysis. Thus, TMIP-EMAT can be set-up to do a systematic set of sensitivity analyses as part of model validation. In general, the set of inputs that is varied and the levels used for those inputs are selected explicitly by a domain expert (either an analyst or another stakeholder) based on that person's judgment about the relative importance of each of the manipulated inputs with respect to that particular scenario. For example, an agency may choose to evaluate a "boundless growth" scenario that pairs rigorous economic growth across all sectors of the economy with high population growth and relaxed policies on carbon emissions, a "directed growth" scenario that funnels new development into the urban core and invests heavily in green transit technology, and a "default" scenario that presumes no change in current policies. In this example, each scenario assumes that two to three distinct uncertainties or policies change in conjunction with each other and have a very set level with regard to how it will change (i.e., a set population total for the high growth scenario).

An **exploratory modeling and analysis** approach would address this modeling task differently. There remains an import role for the domain expert in establishing the scope of the analysis: deciding what sets or ranges of policy levers and/or exogenous uncertainties should be considered, given the modeling tools available and policy questions that need to be addressed. However, given these scoped inputs, the actual design of experiments can be developed using an efficient mathematical algorithm such as a *Latin Hypercube*, developed to focus on identifying the independent and correlated effects of each of the policy levers and exogenous uncertainties. In the context of the hypothetical agency described above, instead of examining three scenarios, an exploratory analysis would include establishing a set of policy levers (e.g., local carbon emission policy, development regulations, and green transit technology investment) and exogenous uncertainties (e.g., economic growth, population growth, nationally imposed carbon taxes, each with a range of varying levels), allowing for no a priori assumptions to be made regarding the likelihood of each factor occurring in conjuction with other factors. Then an experimental design can be created by TMIP-EMAT, potentially defining as many as 60 different experiments (defined by the number of uncertainties and policy levers) to run with various combinations of all of these factors and levels of the factors.

While the overall number of model runs is larger for the exploratory modeling approach (a rule of thumb is 10 model runs per uncertainty variable and policy lever) when compared to more limited scenario analysis, the tools offered by TMIP-EMAT can make both executing these model runs and analyzing the results a manageable effort. Adopting the exploratory modeling approach also offers a number of concrete benefits:

- It allows for **a direct representation of the uncertainty** resulting from the set of uncertainty variables included in the analysis. The efficient experimental design results in a range of outcomes, not a point forecast. The range of outcomes can be used to study the uncertainty impacts over the domain of those variables. By explicitly accounting for uncertainty, risk can be assessed if that is important (e.g., in toll or revenue forecasts). This can be done by explicitly assigning distributions to the various uncertainties and examining the resulting performance measures and their respective distributions.
- The collection of experiments allows the analyst to build a **direct understanding of the relationships** between policy levers, uncertainties, and performance measures. The only way this is possible is by having many core model runs so that interactions of different inputs and their joint impact on model outputs can be measured.
- It is possible to develop **metamodels** for individual performance measures that are capable of predicting the resulting performance measures for a wider range of potential model inputs than are actually run using the core model. Metamodels are described in more detail later in this section.

As noted above, one of the main benefits of using an appropriate experimental design is analyzing how the combination of inputs impacts the resulting performance measures. This is of particular value when an analyst has a collection of policies and wants to test these policies systematically. The exploratory modeling approach allows the analyst to test how different combinations of the potential policies might interact with one another to produce different results.

The exploratory modeling approach allows the analyst to test how different combinations of the potential policies might interact with one another to produce different results.

These sorts of exploratory modeling tasks, using a large number of model runs, can relatively easily be undertaken using a strategic planning model, with limited detail and relatively quick runtimes. However, the tools in TMIP-EMAT make it possible to use the same exploratory modeling approach while taking advantage of much larger and more detailed travel demand models (TDM). By using a full-scale TDM for exploratory analysis, it becomes possible to drill deeper into the model results, as the set of outputs generated by the TDM are quite detailed and nuanced and include local area factors and regional factors alike. Any detailed model output from the TDM can be analyzed using TMIP-EMAT, not just high-level aggregate measures that also might be available from a strategic planning model.

Using a Metamodel to Accelerate Experiments

Within an exploratory modeling effort, an analyst may wish to conduct multiple different sets of experiments using different designs for different purposes. Evaluating models to discover interesting clusters of scenarios and uncover important patterns in the data requires a different set of experiments from developing an optimized set of policy levers, and both require different experiments from a probabilistic risk analysis. To support all of these different exploratory analyses, it is useful to have a model that can be evaluated quickly, ideally many times per second. This level of speed is not what we typically expect from TDMs, which generally require hours or days to complete a single experiment.

To speed up the process of running the model many times, TMIP-EMAT includes a facility to automatically create a *metamodel* (Cambridge Systematics, Inc., 2018). By default, metamodels derived through TMIP-EMAT include two stages: a linear regression model to capture overall trends and simple linear relationships, and a gaussian process regression (GPR) model that can capture a wide variety of nonlinear effects. The GPR is a nonparametric model and has been found to generally perform well and improves the prediction over using only a linear regression model. The implemented combination of these two model structures provides the best of both worlds, giving a generally good model fit in many cases without the need for careful parametric tuning.

Constructing Efficient Experimental Designs for Metamodel Development

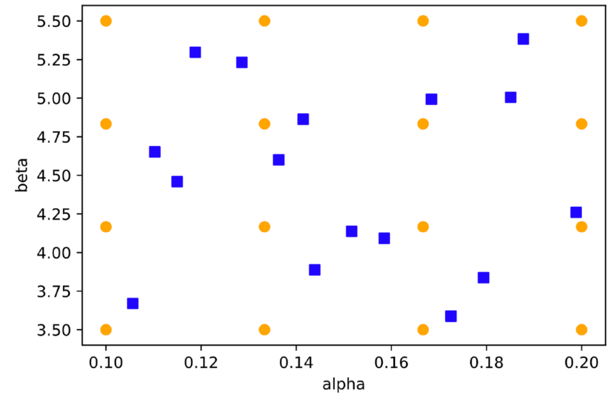
TMIP-EMAT provides algorithms to build efficient designs of experiments based on a defined modeling scope that usually are sufficient for developing metamodels. The creation of an experimental design is fully automated, although the analyst may wish to review the identified experiments to ensure that the scope parameters were set correctly. Specifically, this review entails checking to see that the scoped range is covered by the experiments, and confirming that the sampled experiments are sensitive to the chosen distributions of the uncertainty variables (if anything beyond a default uniform distribution is specified) with more likely outcomes sampled at a higher frequency. The use of nonuniform distributions can be beneficial for well-characterized uncertainty, if the analyst wants to ensure better quality (i.e., denser) results in the heart of the uncertainty space, and specifically for conducting risk analysis.

Review that the scoped range is covered by the experiments and confirming that the sampled experiments are sensitive to the chosen distributions of the uncertainty variables (if anything beyond a default uniform distribution is specified) with more likely outcomes sampled at a higher frequency.

By default, TMIP-EMAT is implemented with a **Latin Hypercube** sampling experiment design approach, although this design algorithm can be overridden by experienced users, if desired. This default design was selected because it is generally an efficient design to support the development of a metamodel. Metamodels for deterministic simulation experiments, such as most travel demand models, are best supported by a "space filling" design of experiments, such as Latin Hypercube draws (Sacks et al., 1989). A Latin Hypercube sample for one dimension is constructed by subdividing the distribution of each input factor into *N* equally probable ranges, and drawing one random sample within each range.

The Latin Hypercube design of experiments is advantageous over a factorial or grid-based design, as every experimental observation can provide useful information, even when some input factors are potentially unimportant or spurious. Figure 12 provides an illustration of this, because every design experiment is unique in each input dimension, and all of the observations provide unique and useful information even if one dimension is spurious. By contrast, the factorial grid of observations collapses on itself with some data points simply replicating others, severely limiting the amount of useful information from the nine experiments.

A representation of a factorial design (orange circles) which covers mostly the edges of the solution space, and a Latin hypercube (blue squares) with the same number of runs, with better coverage of the middle of the space.



If the beta dimension isn't important, the factorial design collapses to only 4 data points (wi), while the Latin hyper-cube still has 16.

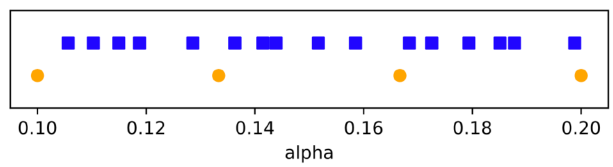


Figure 12. Diagram. Contrasting experimental designs.

(Source: FHWA, 2018.)

One additional advantage of using a Latin hypercube design of experiments, as is done in TMIP-EMAT, is that the required number of experimental runs is not directly dependent on the dimensionality of the input variables and, importantly, does not grow exponentially with the number of dimensions. With a factorial or grid-based design, the number of experiments required does expand exponentially with the number of input dimensions. Practical experience across multiple domains has led to a "rule of thumb" that good results for prediction can be obtained from 10 experimental data points per input variable dimension (Loeppky et al., 2009).

One important consideration in the design of experiments for exploratory analysis is the treatment of differences between *policy levers* and *exogenous uncertainties*. For many analytical analyses, the design of experiments will want to preserve this distinction and maintain a full factorial cross section between these two groupings. That is, for any complete set of policy levers being analyzed, the design should include experiments that match those policies for all of the various set of uncertainties being modeled and vice verse. This partial factorial design will ensure that reasonable results can be obtained, and subsets of experiments that manipulate policy levers or uncertainties always represent the direct effects of these manipulations, and are never merely a fluke of how policy levers and uncertainties interact, especially since the treatments of policy levers and exogenous uncertainties are manifestly different from each other in application. This option may be desirable if metamodels are not needed for the analysis.

However, metamodel development has a singular goal that is quite different from policy analysis: to make the mathematical operation of the metamodel as close as possible to that of the original core model. To this end, any relationship between the core model and any "real world" systems or decisionmaking paradigm is incidental. From a purely mathematical perspective, the metamodel should only be improved by maximizing variability across all input dimensions simultaneously for all the reasons outlined in the discussion of the Latin hypercube above. The TMIP-EMAT design_experiments tool is able to switch easily between metamodel development-focused designs and exploratory analysis designs simply with a change in settings.

The TMIP-EMAT design_experiments tool is able to easily switch between metamodel development-focused designs and exploratory analysis designs.

3.4 Exploring the Results

A key element of exploratory modeling is the interpretation of results, since the results of this type of analysis are different than typical scenario analyses. Instead of producing point forecasts, the exploratory modeling approach produces a range of potential outcomes for a given policy or set of policy options. This allows for more indepth analyses of the overall set of model inputs and outputs. TMIP-EMAT offers a number of different tools to support these analyses, including the following:

- **Visualization** tools to display multiple model outputs together.
- **Scenario Discovery** tools that are focused on finding clusters of scenarios that are interesting to the user.
- **Directed Search** tools to find robust policies that will work well across many scenarios.

Each of these sets of tools is potentially relevant across a broad range of applications, but certain tools are more useful in certain applications. For example, when integrating TMIP-EMAT with a strategic planning model, such as VisionEval's Regional Strategic Planning Model (RSPM), analysts may want to focus on the scenario discovery tools, which can be used to highlight interesting clusters of inputs and outputs for further analysis. These clusters can form the basis for deeper analysis using a more detailed network-based travel demand model. On the other hand, because strategic planning models are less detailed, it may not be as important to employ robust multiobjective optimization tools, such as those available among the directed search tools.

Interesting clusters of inputs and outputs from a strategic model highlighted by scenario discovery tools can form the basis for deeper analysis using a more detailed network-based travel model.

Visualization Tools

Scatter Plot Matrix Visualizations

Visualization tools are an important part of the exploratory modeling process. They allow analysts and other stakeholders to conceptualize the potentially complex relationships represented in transportation models.

The simplest set of visualization tools illustrated in TMIP-EMAT are scatter plot matrices. Sometimes referred to by the abbreviation "SPLOM," a scatter plot matrix is a collection of two-dimensional plots, each showing a contrast between two factors. The two factors are often an input parameter (i.e., an uncertainty or a policy lever) and an output performance measure, although it aso is possible to plot inputs against inputs or outputs against outputs. To facilitate reading the plots, they are arranged in a grid (i.e., a matrix), where all the plots in a single row share a common Y-axis, and all the plots in a single column share a common X-axis.

A portion of an example SPLOM, as generated using the *display_experiments* tool in the emat.analysis subpackage from TMIP-EMAT, is shown in figure 13. The *display_experiments* tool can automatically create a scatter plot matrix that crosses every parameter with every measure, simply by providing the exploratory scope and the results. By default, plots that display policy levers are shown in blue, and plots that show exogenous uncertainties are in red, but this colorization can be overridden or deactiviated if desired. Although each panel of the SPLOM displays only two dimensions, the entire SPLOM in aggregate can be used to gain a general understanding of modeled relationships across multiple dimensions simultaneously. For example, in figure 13, it is easy to spot the strong and clear relationship between no build time and input flow (first row, third column) and to see that relationship is mirrored, but less well defined on the other performance measures (other rows).

Reviewing experimental results in this way can be instructive not only for exploratory analysis, but also for validation of the results from the core model. An analyst can quickly see the direction, magnitude, and shape of various parametric relationships in the model, and easily detect any parameters that are giving unexpected results.

The visualization tools in TMIP-EMAT also demonstrate an alternative use for scatter plot matrix figures: comparing two different sets of model results. To use a scatter plot matrix visualization to contrast two set of experiments, they will need to be derived from the same (or substantially similar) scopes, so that the various factors embedded in the rows and columns of the scatter plot matrix appear in both sets of results. This is particularly valuable to evaluate the performance of metamodels that are derived from core models, as they naturally share design scopes, as we can generate scatter plot matrices that show experiments from both the core and meta models. The contrast_experiments tool in the emat.analysis subpackage from TMIP-EMAT, demonstrates this capability using different colors to denote different sets of experiments, instead of denoting different kinds of inputs. An example of this is shown in figure 17, which contrasts the actual core model experimental runs against similar results generated by a metamodel based on those runs.

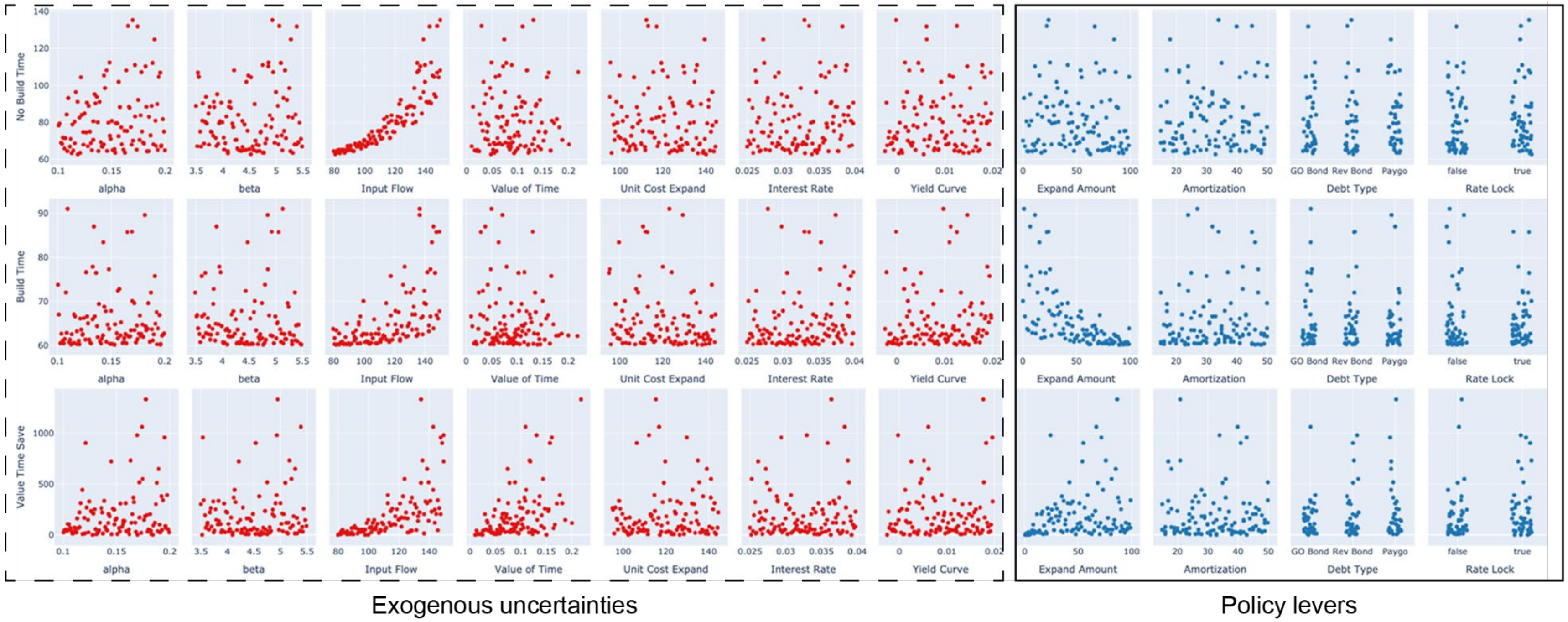


Figure 13. Graph. An example scatter plot matrix.

(Source: TMIP-EMAT Sample Output.)

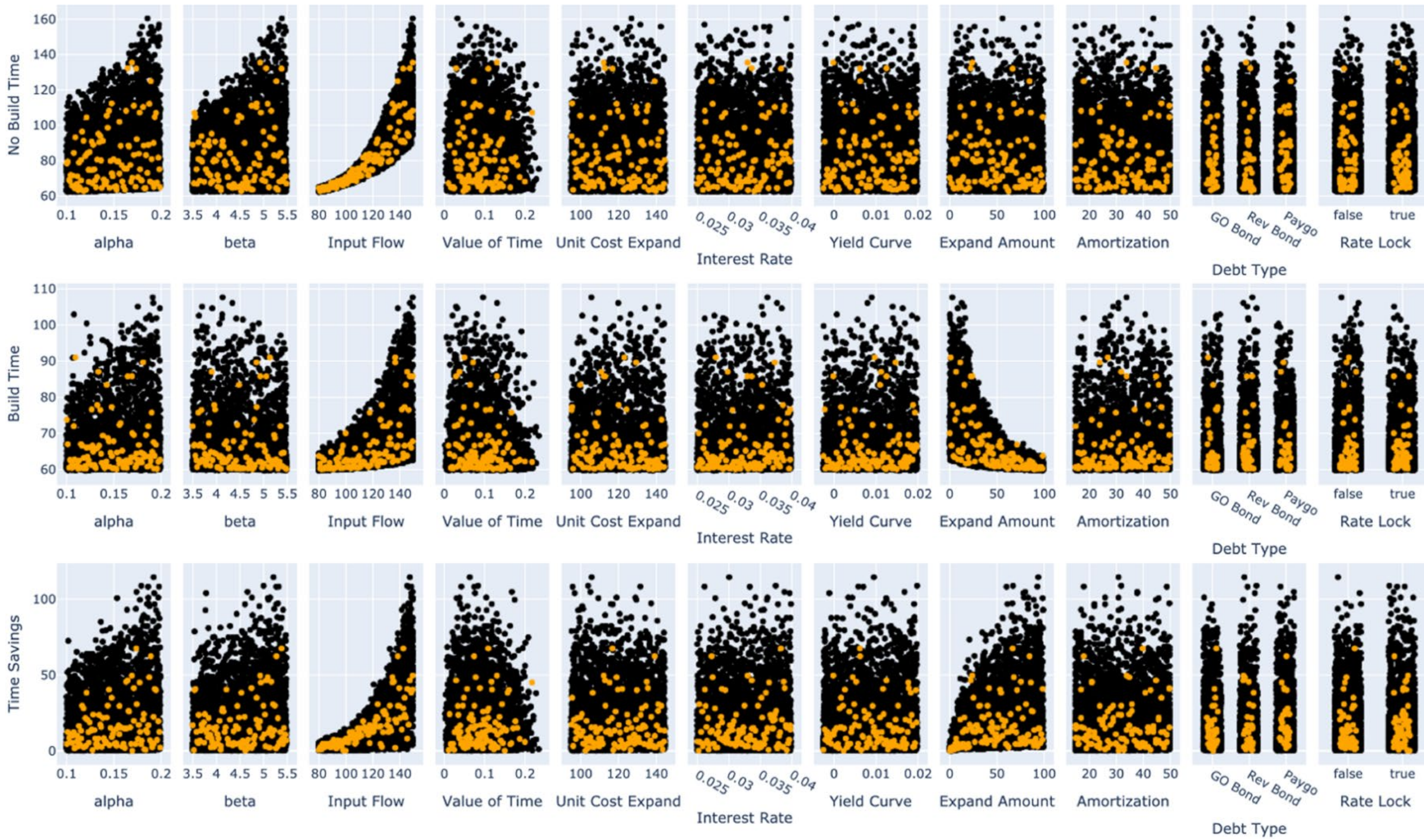


Figure 14. Graph. An example scatter plot matrix contrasting the results of core model runs and metamodel runs.

(Source: TMIP-EMAT Sample Output.)

Interactive Visualizations of Experimental Designs

The various scatter plot matrix visualizations described above offer a static view of a set of exploratory modeling results, perfectly suitable for inclusion in printed reports and other fixed representations. When used in a *Jupyter notebook* environment, TMIP-EMAT also allows for the use of dynamic, *interactive visualization* tools.

The basic interactive visualization interface in TMIP-EMAT is inspired by a similar tool provided with the VisionEval package. To use the interactive visualizer, an analysis will provide the results from design of experiments. These results can be a modest number of experimental model runs executed with a slow-running core model, or a large number of experimental model runs executed with a fast-running core model or a metamodel. The default visualization then renders each dimension of analysis (policy lever, exogenous uncertainty, or performance measure) as a histogram showing the distribution of that dimension across the various experiments included in the analysis.

Although a central tenet of exploratory modeling is a transition away from "point forecast" single model runs, the use of point forecasts is widely ingrained in transportation modeling and analysis, and TMIP-EMAT provides the possibility to add a single point forecast as a reference point against which as spectrum of other experimental results can be displayed in the interactive visualization tools. Such a default result is shown in figure 15, where the reference model inputs and result are displayed as black dotted lines on each figure.

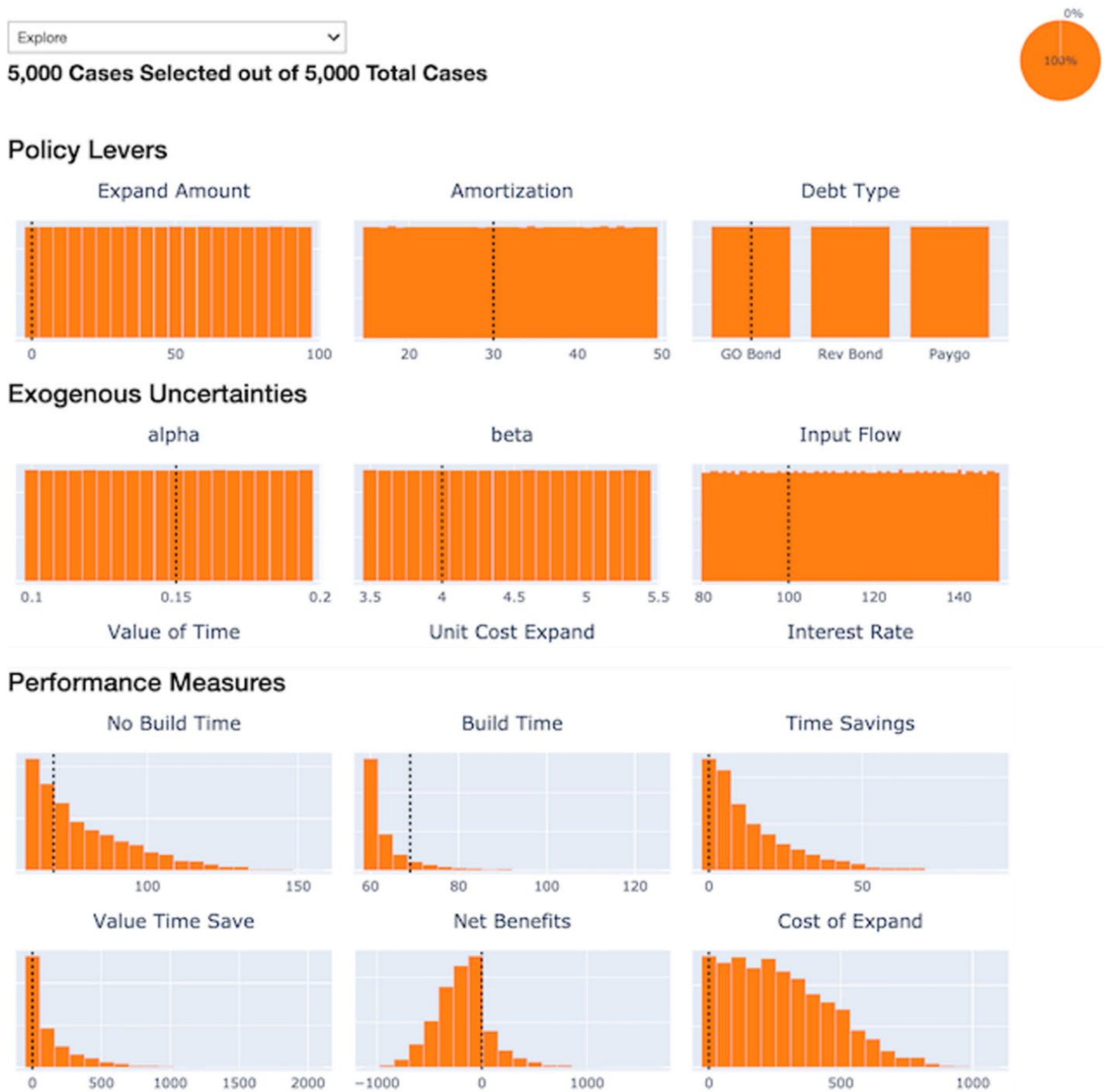


Figure 15. Screenshot. An interactive exploratory modeling visualization before any selections have been made.

(Source: TMIP-EMAT Sample Output.)

The real power of this visualization in TMIP-EMAT is the ability to make subset selections on any histogram displayed, and observe the corresponding subsets across all the other dimensions of analysis. For example, using the analysis shown in figure 15, the analysis can choose to focus on scenarios where the “Expand Amount” policy lever is set to 50 or greater, and the “Debt Type” is set to general obligation bonds. In each case, these selections can easily be made by a simple click-and-drag selection on the relevant histogram, or by making a selection programmatically using Python. Given any particular selections, the entire set of dynamically linked figures will update to show those given selections. Given the example selection described above, the figures will update to appear like the versions shown in figure 15 (note that only a small number of the histograms is displayed as illustration). The "active" selection boundaries are displayed with green rectangles on those dimensions that include the selection criteria. The selected cases represent only about 1/6 of the total number of experimental runs in this design, and the highlighted solid lighter orange region may appear small in some figures, so a dotted line echo of the shape of the selection is also displayed. In this example, we can see that the shape of the distribution in the "Net Benefits" panel is shifted leftward from the overall set of scenarios.

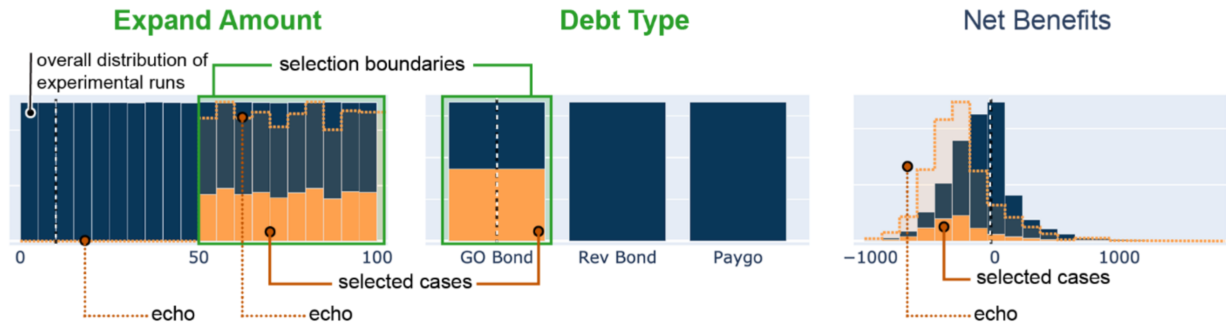


Figure 16. Histogram. Some interactive exploratory modeling visualizations after some selections have been made.

(Source: TMIP-EMAT Sample Output.)

The TMIP-EMAT *interactive visualizer* also can create an interactive two-dimensional scatter plot, linked to the same selections used in the one-factor histograms. This allows the user to specify the variables for both the x and y axis, and either can be any policy lever, exogenous uncertainty, or performance measure. These dimensions can be changed interactively later as well. The resulting scatter plot is linked to the same selection of experiments in the interactive one-dimensional figures shown above, and by default the same experiments are highlighted in the same color scheme in all of these related figures.

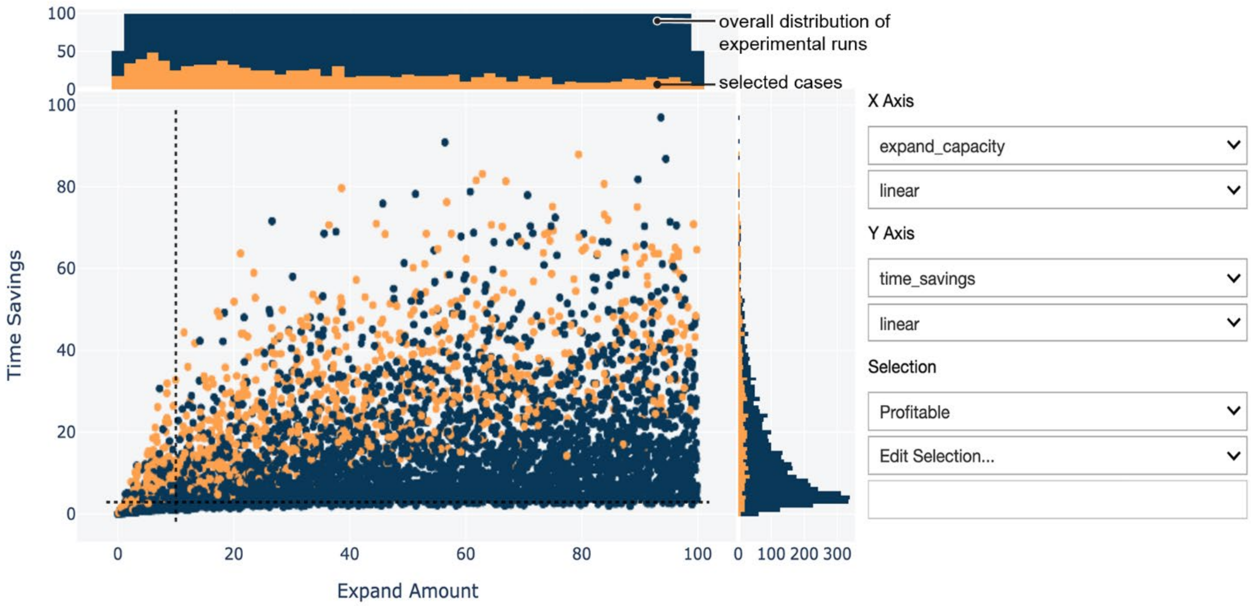


Figure 17. Graph. A two-dimensional interactive scatter plot.

(Source: TMIP-EMAT Sample Output.)

One useful feature of the interactive scatter plot is the ability to manually "lasso" a selection of data points. This lasso selection does not need to be anything like a rectangular box selection, as we have seen so far. Once a lasso selection of data points is made in the figure above, you can choose "Use Manual Selection" from the *Edit Selection…* menu at right, which will create a new Visualizer selection from the selected data. The highlight color changes to signify that this is not an editable rectangular box, and the selected data will be highlighted in all figures linked together, including other scatter plots and one-factor histograms.

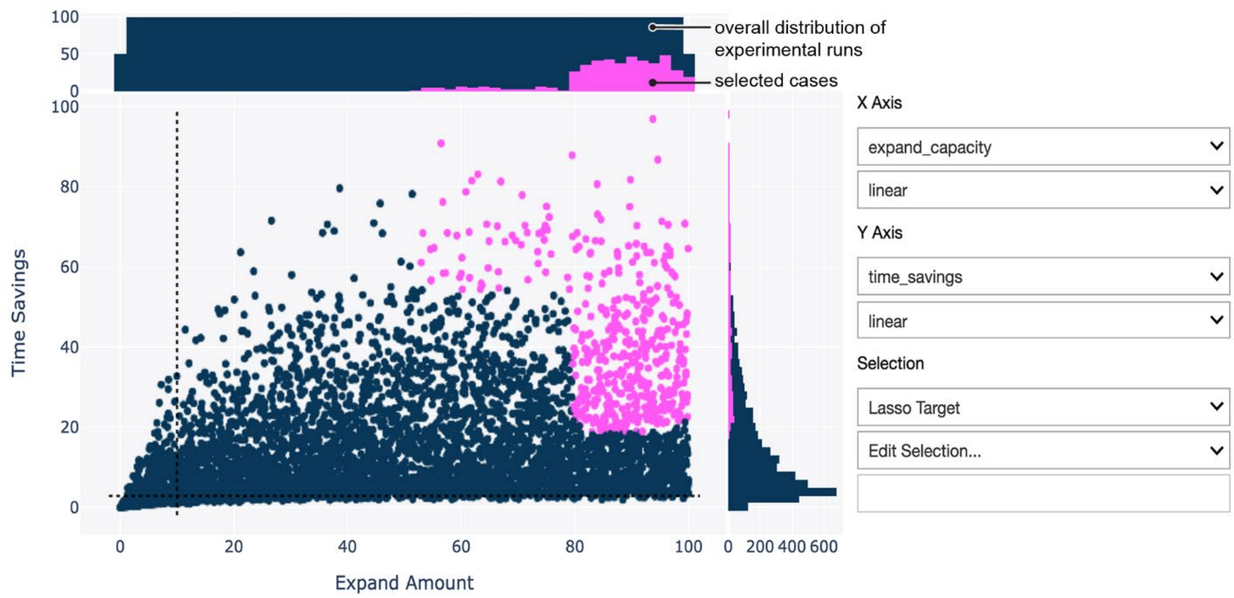


Figure 18. Graph. A two-dimensional interactive scatter plot with a nonrectangular selection.

(Source: TMIP-EMAT Sample Output.)

In addition to the single scatter plot, which offers a feature-packed view of two dimensions at a time, there also is a scatter plot matrix option, which displays a configurable matrix of similar two dimensional views. This matrix has fewer interactive features than the single scatter plot, but still dynamically updates in response to interactive changes in the other figures.

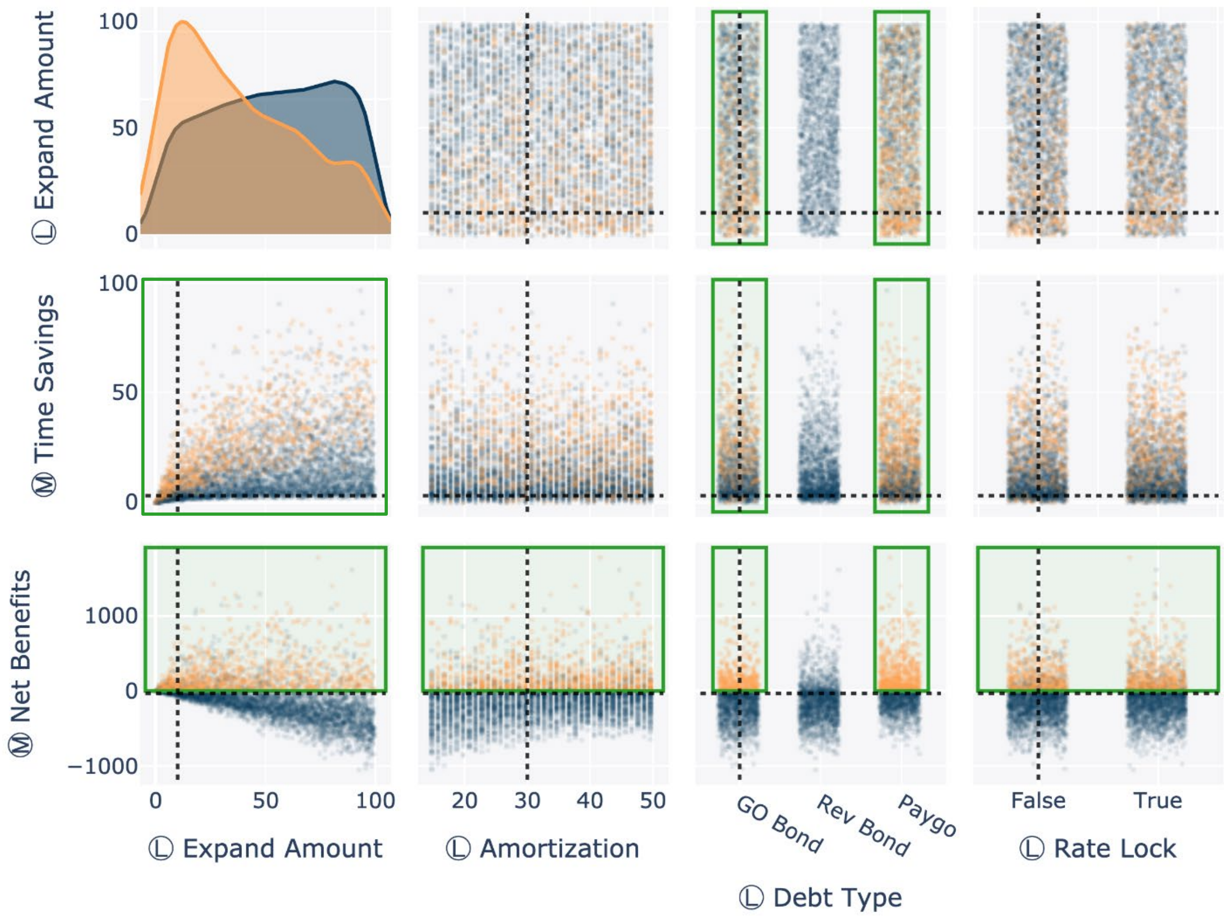


Figure 19. Graphs. A dynamically updating interactive scatter plot matrix.

(Source: TMIP-EMAT Sample Output.)

Scenario Discovery

While visualization tools are focused on letting the analyst see and understand patterns and relationships in model results, scenario discovery tools focus on using computational algorithms to uncover those patterns for us. Both types of analysis are potentially valuable across a wide range of applications, but scenario discovery tools typically become even more useful as the dimensionality of our models (i.e., the number of policy levers and exogenous uncertainties considered) increases.

Feature Scoring

Feature scoring is a technique that allows an analyst to identify the inputs (in machine learning terminology, “features”) that have the largest impacts on particular performance measures. The relationships measured are not necessarily linear, but rather can be any arbitrary linear or nonlinear relationship. The feature scores are determined based on a design of model experiments, and are therefore a product not only of the model, but also the domain (i.e., the range) of the applied input factors.

For example, consider the function *Y(A,B,C) = A/2 + sin(6πB)* + *ε*, where *A, B,* and *C* are input features; and ε is random white noise. We can readily tell from the functional form that the *B* term is the most significant when all parameter vary in the unit interval, as the amplitude of the sine wave attached to *B* is 1 (although the relationship is clearly nonlinear), while the maximum change in the linear component attached to *A* is only one-half, and the output is totally unresponsive to *C*. If we use the TMIP-EMAT visualization tools to generate a scatter plot matrix of 5,000 samples from this model as in figure 20, we could observe visually that *B* is the dominant factor in determining the output measure.

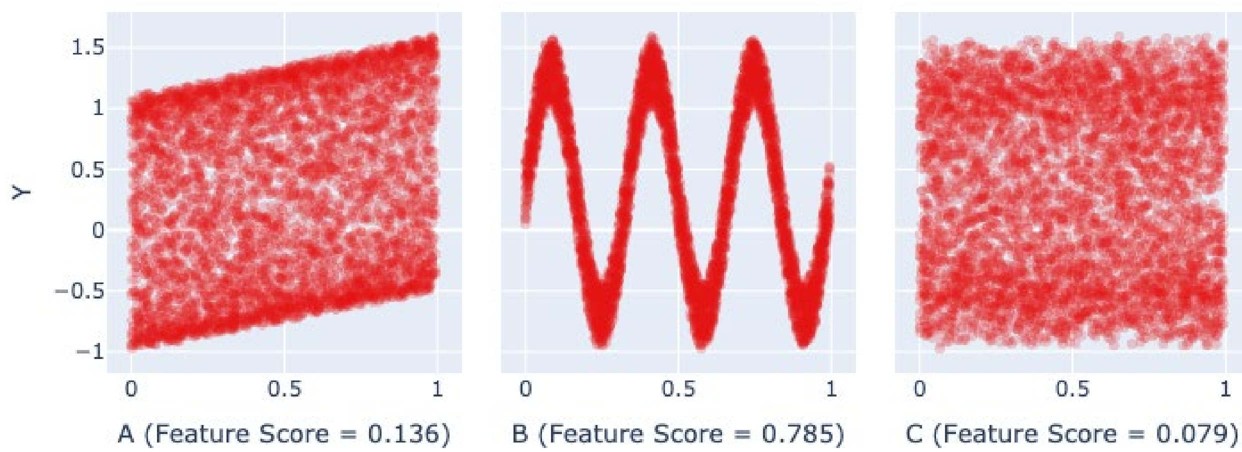


Figure 20. Graphs. Scatter plots visualizing feature scoring for the example function in the unit domain.

(Source: TMIP-EMAT Sample Output.)

In contrast, examining the exact same model under a broader domain yields different feature scores. Figure 21 illustrates this, applying the same model over an input range five times wider. The oscillations of *B* are still visible, but now look much more like white noise, while the effect of *A* is more pronounced. The computed feature scores follow this change with the bulk of the relative importance shifting from *B* to *A*.

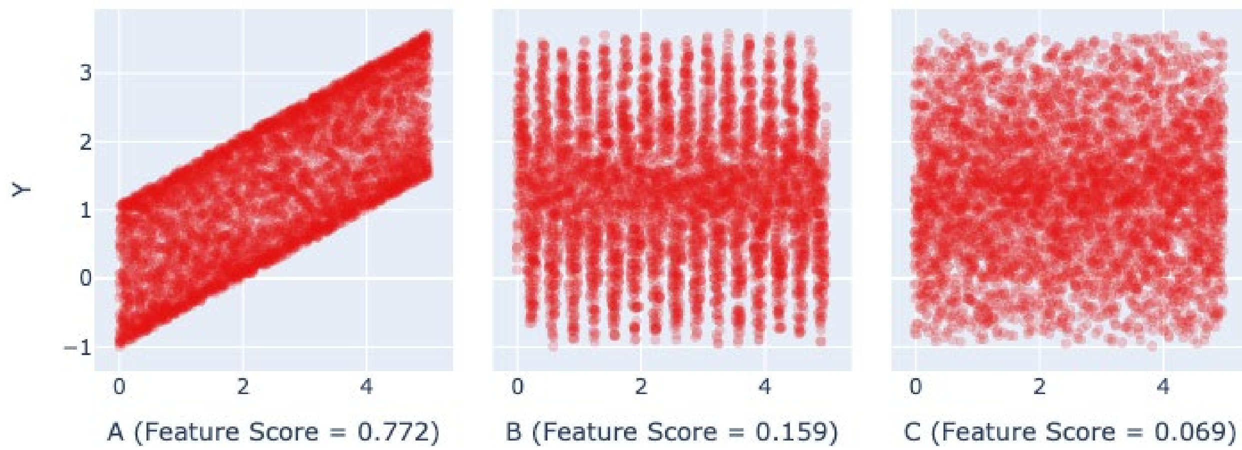


Figure 21. Graphs. Scatter plots visualizing feature scoring for the example function in a larger domain.

(Source: TMIP-EMAT Sample Output.)

The *feature scoring* tools in TMIP-EMAT can automatically calculate and display feature scores for the full set of performance measures defined for a model scope, or a subset thereof. An example of the feature scores table produced by TMIP-EMAT is shown in figure 22. The colors on the returned table highlight the most important input features for each performance measure (i.e., in each row). The yellow highlighted cell indicates the most important input feature for each output feature, and the other cells are colored from yellow through green to blue, showing high-to-low importance.

	alpha	beta	input_flow	value_of_time	unit_cost_expansion	interest_rate	yield_curve	expand_capacity	amortization_period
no_build_travel_time	0.075293	0.068182	0.538393	0.046090	0.043536	0.045446	0.042912	0.041798	0.039350
build_travel_time	0.046962	0.047576	0.294707	0.048980	0.051786	0.061001	0.053346	0.274514	0.047741
time_savings	0.082390	0.089531	0.437443	0.046821	0.051698	0.050382	0.046855	0.086983	0.048425
value_of_time_savings	0.080125	0.064977	0.296455	0.196913	0.043270	0.051546	0.051619	0.086728	0.045814
net_benefits	0.073350	0.048033	0.274585	0.116004	0.048006	0.043684	0.058289	0.189976	0.065641
cost_of_capacity_expansion	0.037718	0.034734	0.048601	0.042841	0.084262	0.039434	0.044483	0.494295	0.103566
present_cost_expansion	0.038055	0.025013	0.029004	0.029836	0.086967	0.039247	0.035163	0.649856	0.025137

high importance
low importance

Figure 22. Screenshot. Example feature scores for the road test model.

(Source: TMIP-EMAT Sample Output.)

Threshold scoring provides a set of feature scores that do not relate to the overall magnitude of a performance measure, but rather whether that performance measure is above or below some threshold level. TMIP-EMAT includes a threshold_feature_scores function that computes such scores for a variety of different thresholds to develop a picture of the relationship across the range of outputs for a particular performance measure.

In figure 23, we can see that *expand_capacity* is important in determining the magnitude of negative outcomes, but for understanding whether we will have positive or negative outcomes, we should focus more on *input_flow*; and if we are interested in the magnitude of positive outcomes, we should look to *input_flow* and to a lesser, but still meaningful extent also the *value_of_time*.

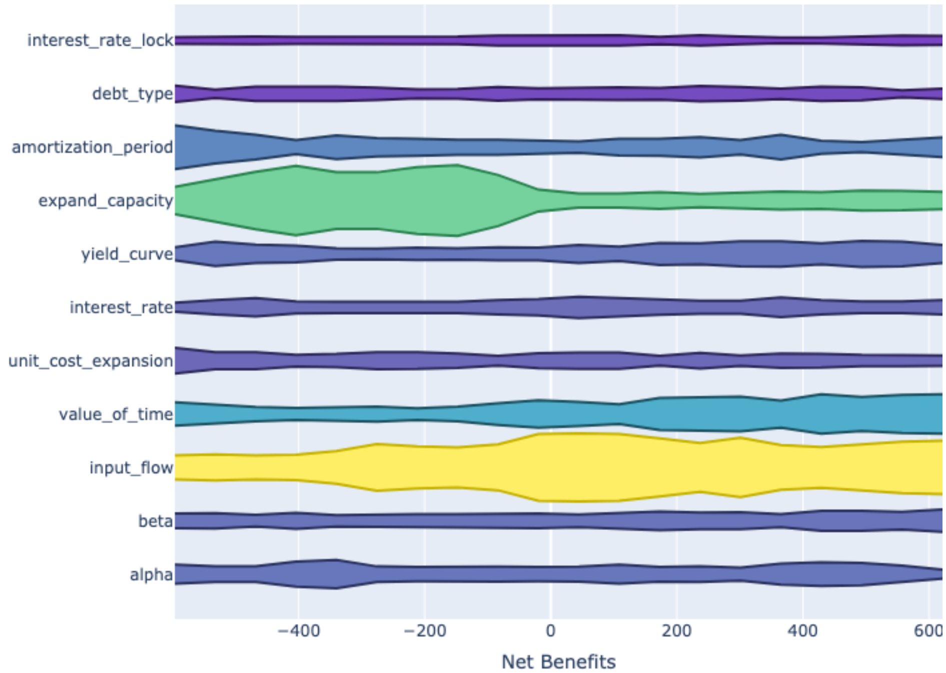


Figure 23. Graph. Example threshold feature scores for net benefits in the road test model.

(Source: TMIP-EMAT Sample Output.)

Patient Rule Induction Method

The *Patient Rule Induction Method* (PRIM) is a scenario discovery algorithm that operates on an existing set of data with model inputs and outputs (i.e., you have already designed and run a set of experiments using either a core model or a metamodel). Several slightly different versions of PRIM can be constucted, but TMIP-EMAT provides a tool that implements the PRIM algorithm based on the EMA Workbench implementation. Generally, a decently sized set of experiments (hundreds or thousands) is used to describe the solution space, although no minimum number of experiments is formally required.

PRIM is used for locating areas of an outcome space that are of particular interest, which it does by reducing the data size incrementally by small amounts in an iterative process as follows:

- Candidate boxes are generated. These boxes represent incrementally smaller sets of the data. Each box removes a portion of the data based on the levels of a single input variable.
  - For categorical input variables, there is one box generated for each category with each box removing one category from the data set.
  - For integer and continuous variables, two boxes are generated—one box that removes a portion of data representing the smallest set of values for that input variable and another

box that removes a portion of data representing the largest set of values for that input. The step size for these variables is controlled by the analyst.

- For each candidate box, the relative improvement in the number of outcomes of interest inside the box is calculated, and the candidate box with the highest improvement is selected.
- The data in the selected candidate box replaces the starting data and the process is repeated.

The process ends based on a stopping criteria. For more details on the algorithm (see Friedman and Fisher (1999) or Kwakkel and Jaxa-Rozen (2016). The PRIM algorithm is particularly useful for scenario discovery, which broadly is the process of identifying particular scenarios of interest in a large and deeply uncertain dataset. In the context of exploratory modeling, scenario discovery is often used to obtain a better understanding of areas of the uncertainty space, where a policy or collection of policies performs poorly because it is often used in tandem with robust search methods for identifying policies that perform well (Kwakkel and Jaxa-Rozen, 2016).

In the context of exploratory modeling, scenario discovery is often used to obtain a better understanding of areas of the uncertainty space where a policy or collection of policies performs poorly because it is often used in tandem with robust search methods for identifying policies that perform well.

In order to use PRIM for scenario discovery, the analyst must first conduct a set of experiments. This includes having both the inputs and outputs of the experiments (i.e., the model or metamodel has already been run). The analyst also must identify what constitutes a case that is "of interest." This is essentially generating a True/False label for every case, using some combination of values of the output performance measures, as well as (possibly) the values of the inputs. Some examples of possible definitions of "of interest" might include the following:

- Cases where total predicted VMT (a performance measure) is below some threshold.
- Cases where transit farebox revenue (a performance measure) is above some threshold.
- Cases where transit farebox revenue (a performance measure) is above 50 percent of budgeted transit operating cost (a policy lever).
- Cases where the average speed of tolled lanes (a performance measure) is less than free flow speed, but greater than 85 percent of free flow speed (i.e., bounded both from above and from below).
- Cases that meet all of the above criteria simultaneously.

The salient features of a definition for "of interest" is that: 1) it can be calculated for each case if given the set of inputs and outputs, and 2) that the result is a True or False value. When conducted using the tools in TMIP-EMAT, the PRIM algorithm will generate a number of different possible boxes along a (heuristically) optimal trajectory, trading off coverage against density.

Coverage is percentage of the cases of interest that are in the box (i.e., number of cases of interest in the box divided by total number of cases of interest). The starting point of the PRIM algorithm is the unrestricted full set of cases, which includes all outcomes of interest, and therefore, the coverage starts at 1.0 and drops as the algorithm progresses. Density is the share of cases in the box that are case of interest (i.e., number of cases of interest in the box divided by the total number of cases in the box). As the box is reduced, the density will increase (as that is the objective of the PRIM algorithm). For the statistically minded, this tradeoff also can be interpreted as the tradeoff between Type I (false positive) and Type II (false negative) errors. High coverage minimizes the false negatives, while high density minimizes false positives.

By default, the PRIM algorithm sets the “selected” box position as the particular box at the end of the peeling and pasting trajectory, which has the highest density, but generally the smallest or close to the smallest coverage. Figure 24 shows the tradeoff curve as a static plot of points mapping coverage versus density, as created using the show_tradeoff command. The colors along the trajectory indicate the number of restricted dimensions used to define each box.

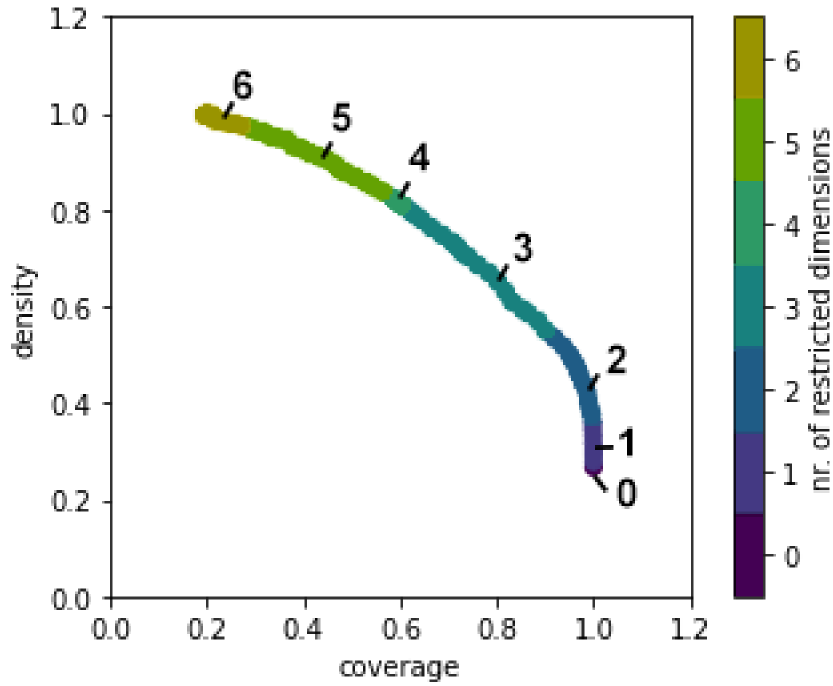


Figure 24. Chart. Example Patient Rule Induction Method trajectory.

(Source: TMIP-EMAT Sample Output.)

The results of a PRIM analysis can be used in conjunction with the interactive exploration tools to find interesting scenarios within a large multidimensional input space, and to identify different combinations of policy strategies that can achieve a portfolio of different goals.

Classification and Regression Trees (CART)

Classification and Regression Trees (CART) also can be used for scenario discovery, and TMIP-EMAT provides a basic tool to do so. This algorithm partitions the explored space (i.e., the scope) into a number of sections, with each additional partition being added in such a way as to maximize the difference between observations on each side of the newly added partition divider, subject to some constraints (e.g., there must be some minimum number of observations in each partition).

Directed Search

The scenario discovery tools outlined above are focused on exploring parameters and outcomes across a preset sample of model runs. This sample can be small or quite large, but the tools only consider cases that have already been evaluated. In the suite of TMIP-EMAT directed search tools, an analyst will find tools that will propose and execute new cases.

Policy Contrast

The *policy contrast* viewer, also known as the “AB” viewer, allows an analyst to compare the outcomes of two different sets of policies to get an understanding of how they differ. The tool runs the model across a distribution of inputs, and displays the resulting distribution of performance measure outputs. By default, uncertainties are modeled using the distributions contained in the model scope, while policy levers are each manipulated to be a specific value. Two sets of model runs are generated by making random draws from all the relevant distributions, and then running the model for every combination of random draws and each set of specific-value inputs. For example, if the number of background random draws is set at 250, then 500 model runs are conducted, 250 each for the 2 different sets of selected policy levers. The exact same random draws are used for both groups of model runs, so that any variation in the performance measures can be unambiguously linked to the changes in the specific-value inputs, instead of being a result of input stochasticity.

The policy contrast viewer has two principal parts:

1. The **interface** (shown in figure 25) has interactive controls for the analyst to control the tool. Each exogenous uncertainty and policy lever is represented by a row of controls. The left-most toggle, next to the parameter name, controls whether each input is in “distribution” or “specific value” mode. This allows, for example, setting one uncertainty to particular high and low values, to visualize threats and opportunities.
   a. When a parameter is set to distribution mode, a readout of the shape and extent of the distribution is provided. There are no other controls for parameters in this mode.
   b. When a parameter is set to specific value mode, an additional set of controls appears in the row. Two sliders are shown along with value read outs. The left and right sliders control the "A" and "B" settings, respectively.

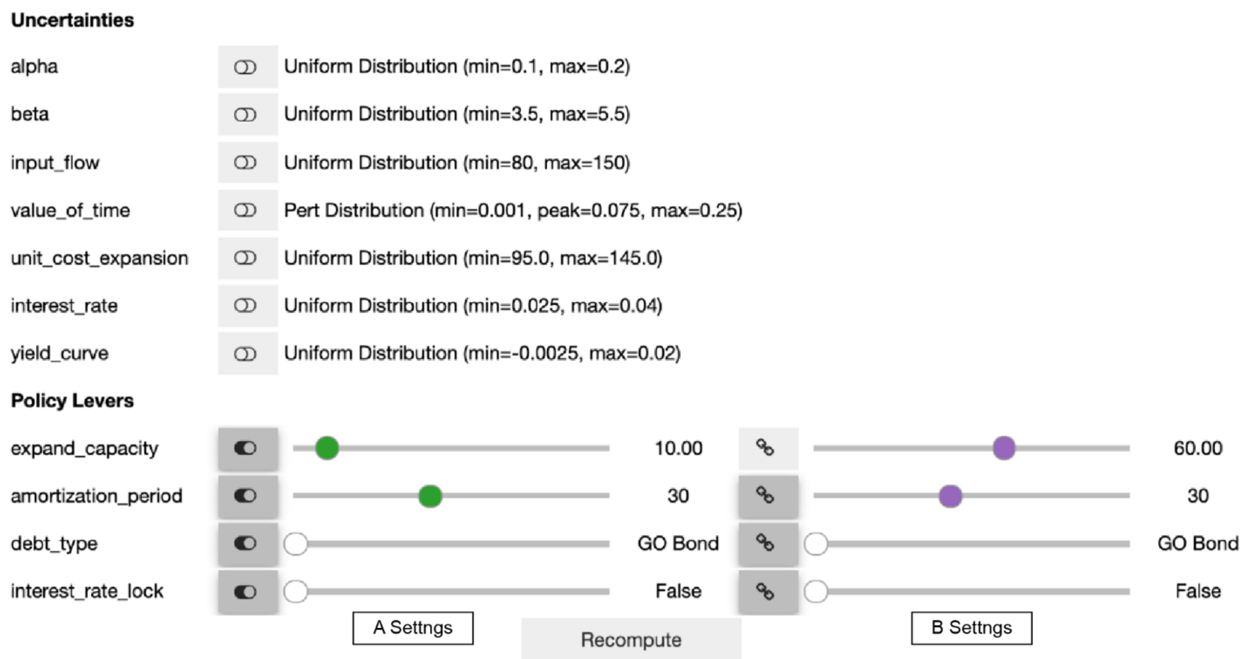


Figure 25. Screenshot. The Travel Model Improvement Program-Exploratory Modeling and Analysis Tool policy contrast interface.

(Source: TMIP-EMAT Sample Output.)

2. The **figure viewer** (figure 26) can be shown using the `get_figures` command. This command allows for naming a curated list of specific performance measure outputs, instead of generating figures for all performance measures. Each figure contains two plots. The left plot is an asymmetric violin plot, showing the distribution for that performance measure under the A (green, upper half) and B (purple, lower half) policies. The right-hand figure shows in red the distribution of the *paired* differences between the A and B results. These differences are not aggregate differences for A and B, but rather the distribution of the differences for the same random draw in each group.

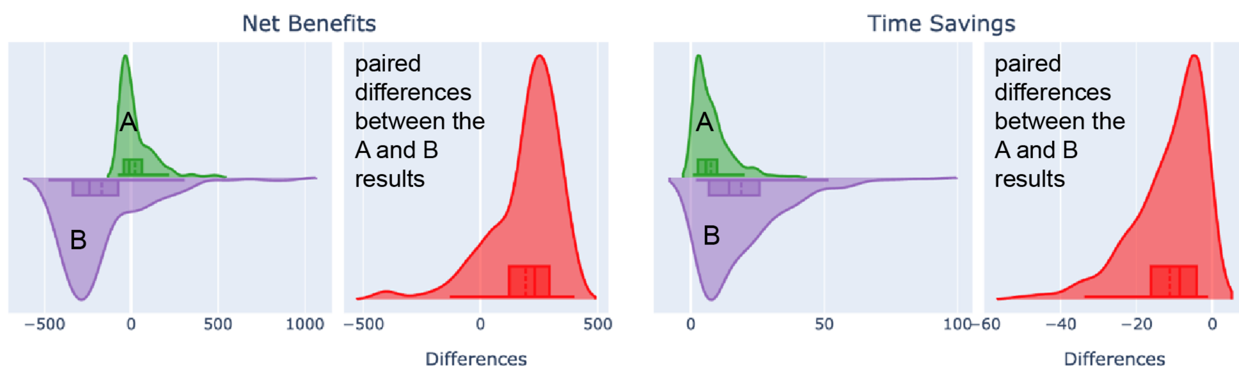


Figure 26. Charts. Examples of results from the Travel Model Improvement Program-Exploratory Modeling and Analysis Tool policy contrast tool.

(Source: TMIP-EMAT Sample Output.)

There are several notable features of the policy contrast tool:

- Using the TMIP-EMAT policy contrast viewer interactively requires a model that is rerunnable in near real time, as every time the “recompute” button is pushed, the model will need to be run hundreds of times before the figures can be updated. This can be a very simple model or a metamodel of a more complex model.
- The analyst also specifies the number of background model runs that will be conducted to simulate the distribution of parameters set to “distribution” mode. Generally, a few hundred runs will be sufficient to get a reasonable high-level overview of the distributions.

Optimization

Typically, transportation planning models will be used to try to find policies that provide the “best” outcomes. In a traditional analytical environment, that typically means using models to find *optimal* outcomes for performance measures. Transportation models as used in the TMIP-EMAT framework are generally characterized by two important features: they are subject to significant exogenous uncertainties about the future state of the world, and they include numerous performance measures for which decisionmakers would like to see good outcomes. Therefore, optimization tools applied to these models should be flexible to consider *multiple objectives*, as well as be *robust* against uncertainty.

One approach to managing a multiobjective optimization problem is to distill it into a single objective problem by assigning relative weights to the various objectives. This can be difficult to accomplish in public policy environments that are common in transportation planning for a variety of reasons, including the following:

- Multiple stakeholders may have different priorities and may not be able to agree on a relative weighting structure.
- Certain small improvements in a performance measure may be valued very differently if they tip the measure over a regulated threshold (e.g., to attain a particular mandated level of emissions or air quality).

Instead of trying to simplify a multiobjective into a simple-objective one, an alternate approach is to preserve the multiobjective nature of the problem and find a set or spectrum of different solutions; each of which solves the problem at a different weighting of the various objectives. Within a set of solutions for this kind of problem, each individual solution is “Pareto optimal,” such that no individual objective can be improved without degrading at least one other objective by some amount. Thus, each of these solutions might be the “best” policy to adopt, and exactly which is the best is left as a subjective judgment to decisionmakers, instead of being a concretely objective evaluation based on mathematics alone.

Optimization tools in exploratory modeling are not limited trying to discover the best solution; sometimes optimization is used to find the worst outcomes as well. TMIP-EMAT offers several different approaches to optimization; each of which can offer different insights into the policymaking process for stakeholders and analysts alike.

The simplest optimization tool available for TMIP-EMAT users is a *search over policy levers*, which represents multiobjective optimization, manipulating policy lever values to find a Pareto optimal set of solutions, holding the exogenous uncertainties fixed at a particular value for each uncertainty (typically at the default values). This is often a useful first step in exploratory analysis, even if your ultimate goal is to eventually undertake a more robust optimization analysis. This less complex optimization can give insights into tradeoffs between performance measures and reasonable combinations of policy levers.

To conduct an optimization search over levers, the analyst can use the optimize method of the TMIP-EMAT model class, setting the search_over argument to 'levers'. In a Jupyter notebook environment, an analyst can monitor convergence visually in real time in the figures that will appear automatically when optimizing. An example of these displays is shown in figure 27.

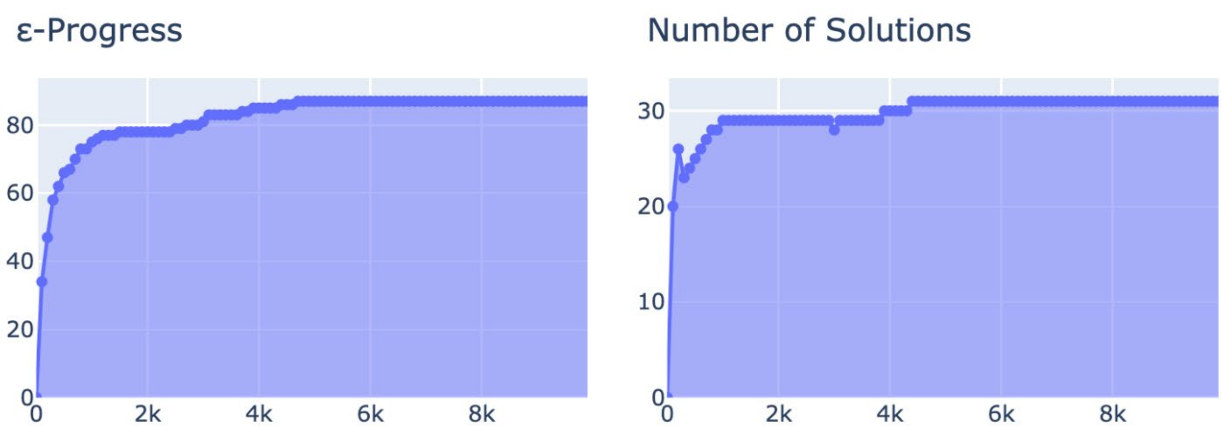


Figure 27. Charts. Example convergence displays.

(Source: TMIP-EMAT Sample Output.)

Once converted, TMIP-EMAT returns potentially not just one solution, but a Pareto optimal set of different resulting solutions, as well as some information about how they were derived. One way that TMIP-EMAT users can visualize the set of solutions is by using a parallel coordinates plot (figure 28), which is composed of a number of vertical axes, one for each column of data in the results table. By default, the axes representing performance measures to be minimized are inverted in the parallel coordinates, such that moving up along any performance measure axis results in a “better” outcome for that performance measure.

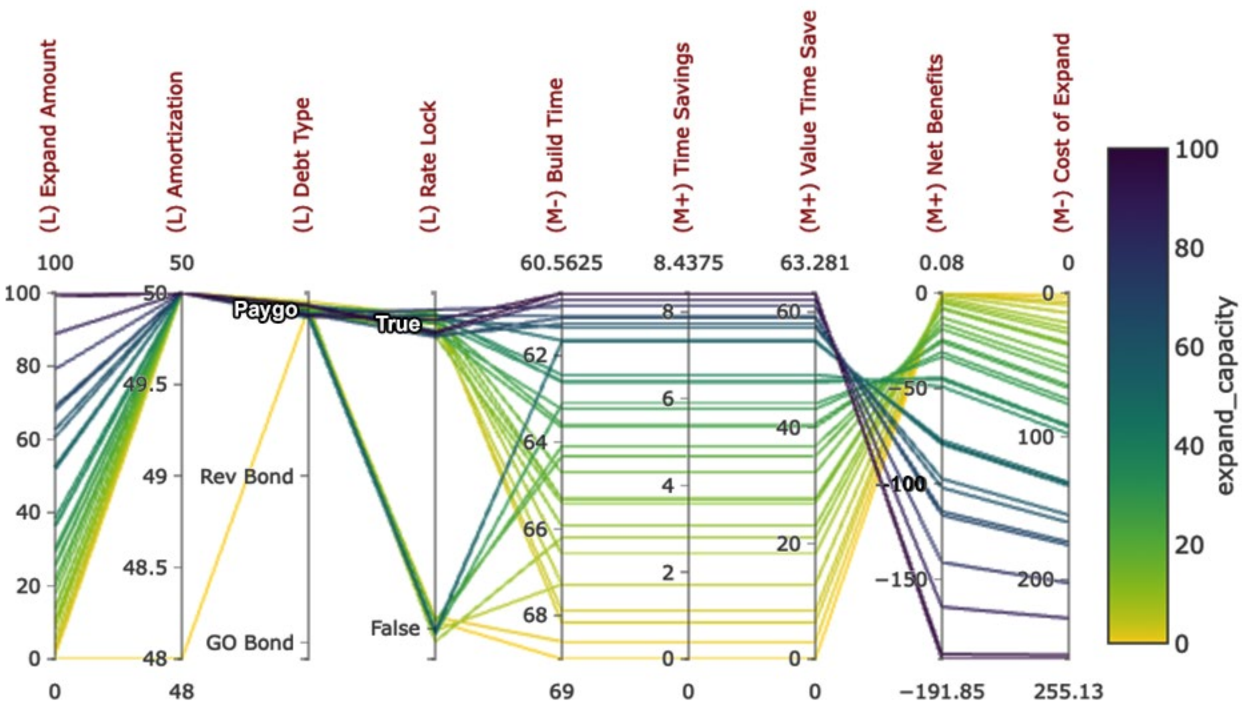


Figure 28. Graph. An example parallel coordinates plot of results from a search over levers using the example road test model.

(Source: TMIP-EMAT Sample Output.)

In the example shown in figure 28, the four policy levers appear as the first four vertical axes. Several observations can be made:

- Nearly all of the Pareto optimal policy solutions for our reference scenario share an amortization period of 50 years, and all share a debt type of "Paygo."
- The set of optimal solutions include multiple different values for the expand capacity lever, ranging from 0 to 100. These different values offer possible tradeoffs among the performance measures: lower levels of capacity expansion (shown in yellow) will maximize net benefits and minimize the cost of the project, but they will also fail to provide much travel time savings. This tradeoff is visually cued by a characteristic "twist" in the chords that appears towards the right side of figure 28. It is left up to the analysts and decisionmakers to judge what tradeoffs to make between these conflicting goals.

We can apply the same multiobjective optimization tool in reverse to undertake a *worst case discovery*. In such an analysis, generally the analyst will flip from manipulating levers given fixed values of the uncertainties to manipulating uncertainties given a particular set of policy lever settings. In addition, the optimization engine is reversed, to search for the worst outcomes instead of the best ones.

An example set of results from a worst-case analysis are shown in figure 29, where a large road capacity expansion project is undertaken. The large number of chords in this figure represents the very broad number of ways that things can go very badly for this model. The colorization can assist in interpreting the results, as it highlights the characteristic tradeoff "twist" so often visible in such results. In this example, we can observe two basic sets of worst case problems: we can do badly because high traffic flows cause congestion, or we can do badly because low traffic flows mean there would have been little congestion anyway, and our very expensive public works project delivered minimal benefit.

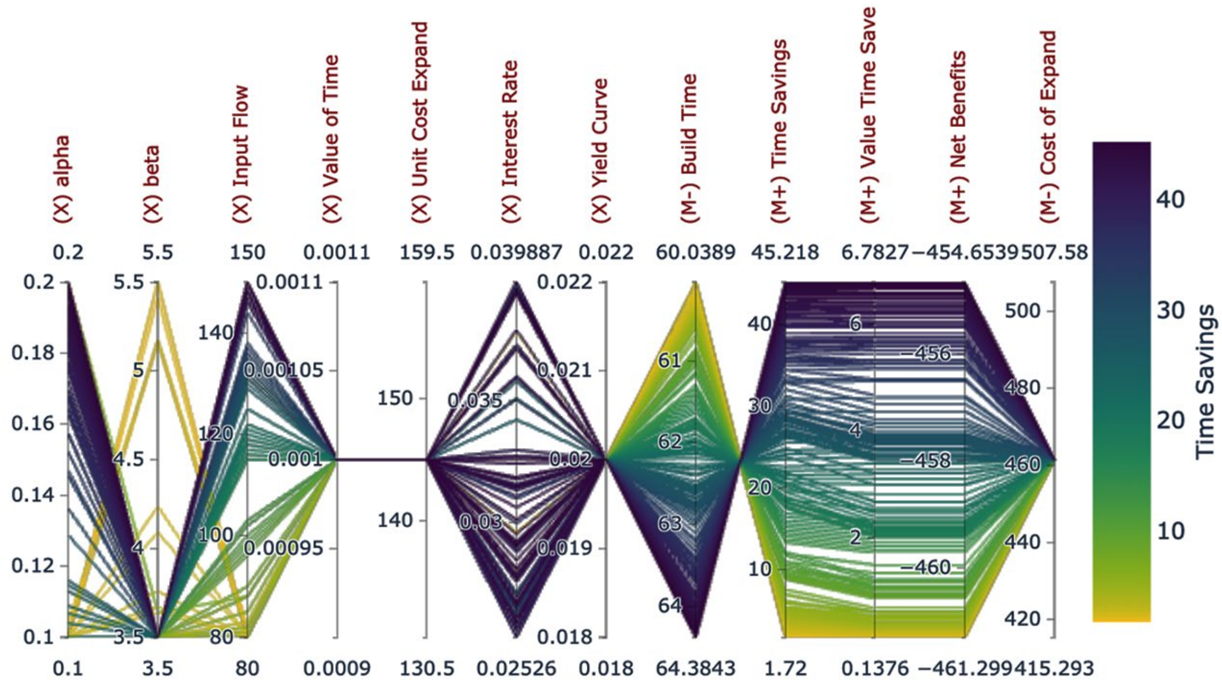


Figure 29. Graph. An example parallel coordinates plot of results from a worst-case analysis using the example road test model.

(Source: TMIP-EMAT Sample Output.)

Robust Optimization

In addition to the multiobjective nature of optimization in TMIP-EMAT, analysts may wish to optimize not only to find a solution that is ideal for one particular scenario, but rather a robust solution. Robust optimization is a variant of the more traditional optimization problem, where policies that yield good outcomes across a range of possible futures are sought, rather than the best outcome for a particular future. As it has some important implementation features that are distinct from other optimization problems, *robust optimization* is implemented in TMIP-EMAT as a distinct tool.

To perform robust optimization using TMIP-EMAT, in addition to providing a core model (or metamodel) with defined input and output parameters, an analyst will need to define one or more functions called "robustness measures" that define what a robust measure represents.

For example, consider the example policies shown in figure 30. In this example, we could compute a varitey of robustness measures shown on the right-hand side of the figure. In this example, policy 3 performs best if the robustness measure is to select the maximum possible outcome of the performance measure, while policy 2 performs best if the robustness measure is selected as the mean, median, or 90th percentile of the performance measure across values of the exogenous uncertainty.

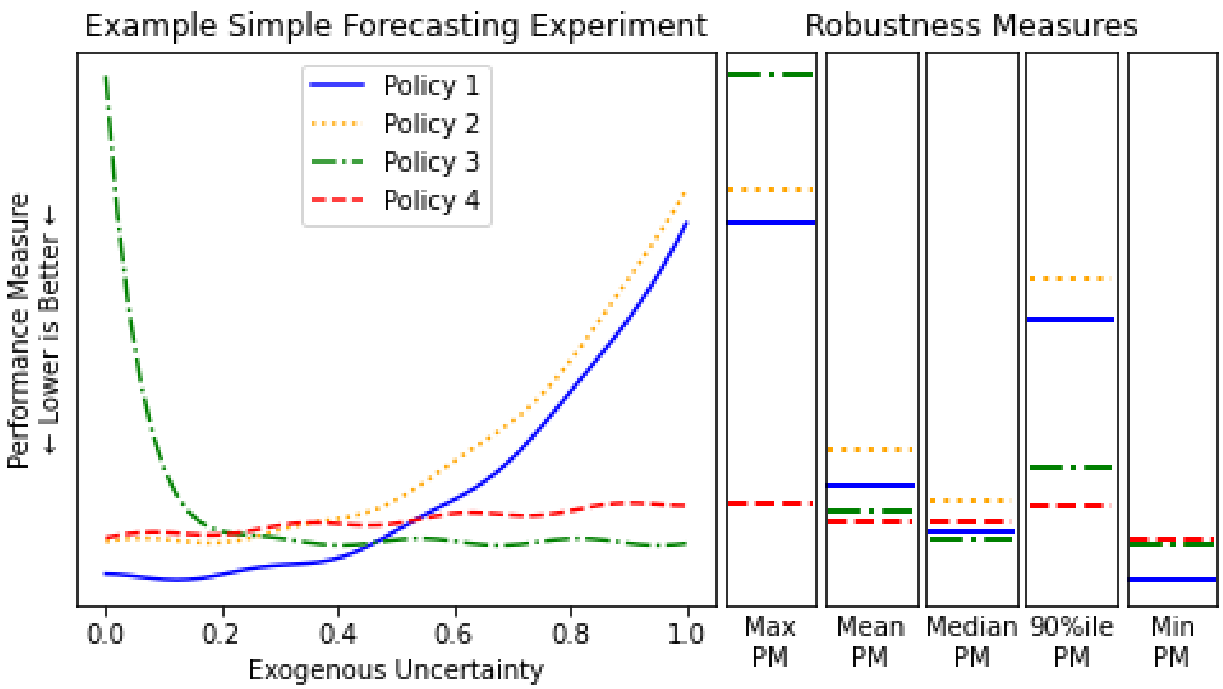


Figure 30. Chart. Example policy outcomes and robustness measures.

(Source: TMIP-EMAT Sample Output.)

Implementing robust optimization in TMIP-EMAT requires an analyst to define the relevant robustness functions. As illustrated above, the functional form of robustness functions can be many different things depending on the particular application, risk tolerances of stakeholders, and expectations about the future, so TMIP-EMAT does not implement a mechanism to generate them automatically. Instead, it is left to the analyst to develop a set of robustness functions that are appropriate for each application.

Several features of the robust optimization tool in TMIP-EMAT include the following:

- A robust measure is created in TMIP-EMAT using the same *Measure* object types used for performance measures that are direct model outputs.
- Robust measures have two important additional attributes: a *variable_name*, which names the underlying performance measure upon which this robust measure is based, and a *function* that describes how to aggregate the results. Typical functions used for robust optimization include the minimum, maximum, and percentiles.

- More abstract robustness measures also can be created. For example, we can compute the percentage of scenarios where the net benefits of a project are positive, or where greenhouse gas emissions are below a particular statutorily mandated level.
- The robust optimization process in TMIP-EMAT can be constrained to only include solutions that satisfy certain constraints.
- The robust optimization process generally is much slower than other multiobjective optimizations in TMIP-EMAT, as each iteration of the process entails running the underlying model not once, but many times, to evaluate the various robustness measures against a sample of numerous different sets of exogenous uncertainties.

Once completed, the results of a robust optimization look very similar to those of the other optimization processes outlined above and can be analyzed in many of the same ways. Figure 31 illustrates a parallel coordinates plot showing the various pareto-optimal robust outcomes. Since each possible set of policy levers is evaluated against not just one set of uncertainties, but a large pool of different values, the performance measures shown in this figure are not the raw performance measures from the actual model runs, but instead the robustness measures are displayed. Figure 31 again shows a characteristic twist on the right side indicating a tradeoff: better values for 95th percentile project cost must be traded off against better values for the expected travel time savings. Again, TMIP-EMAT makes no assertion about how to value this tradeoff, but defers these value judgments to analysts and stakeholders.

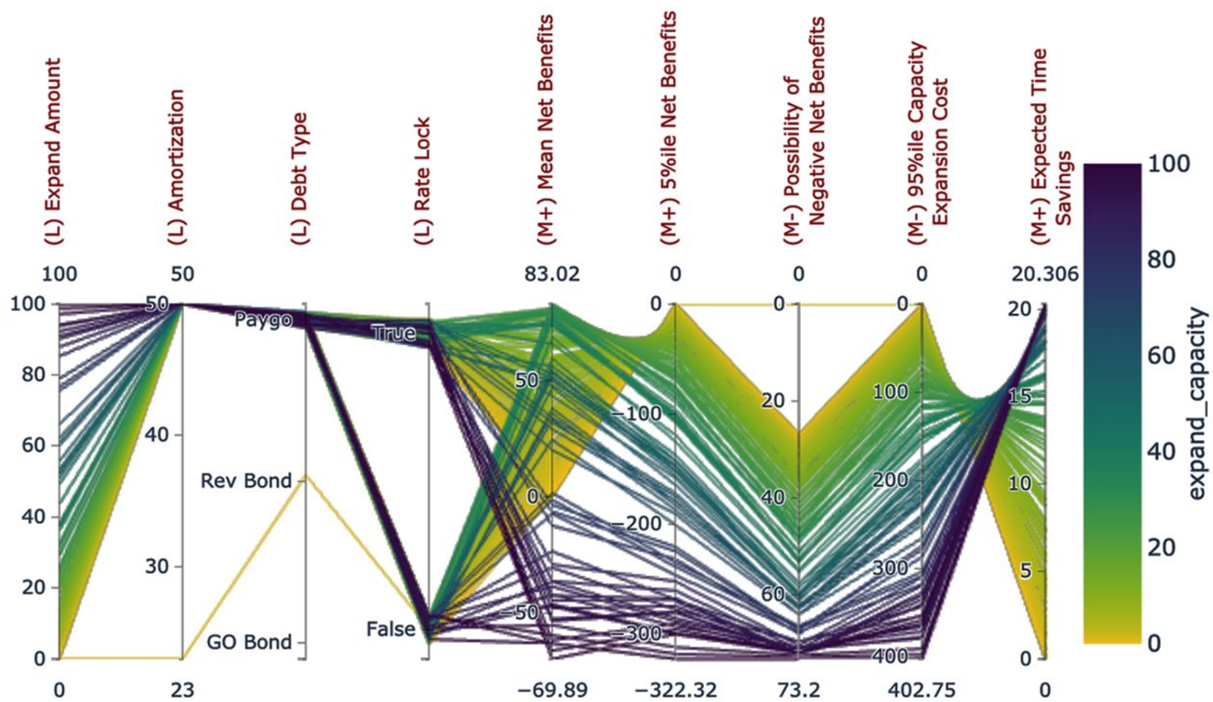


Figure 31. Graph. Example robust optimization results.

(Source: TMIP-EMAT Sample Output.)

3.5 Using the Results

The tools provided by TMIP-EMAT, and by exploratory modeling more generally, are meant to be descriptive, and not prescriptive. Each tool is built to provide insight into the relationships between policy levers, uncertainties, and performance measures. The end goal of the analysis is not to provide explicit guidance about what decisions are "best," but to allow decisionmakers and other stakeholders to better grasp the tradeoffs between different courses of action, and help to facilitate and focus discussion about policy actions.

In particular, TMIP-EMAT's tools can help analysts craft and visualize narratives around the quantitative results of travel demand models. The multidimensional and multiobjective nature of the tools pairs well with the nuanced and multifaceted nature of the transportation system. Rather than simply putting the interactive tools in the hands of stakeholders and inviting them to explore for themselves, analysts may find it more useful to explore and vet ideas on their own, and then use the interactive tools to walk stakeholders through the "story" of the model, pointing out and discussing key relationships and tradeoffs. In these discussions, often the simplest exploratory analysis tools can be the most compelling: the feature scoring tools, which succinctly describe the relative importance of each model input on each output, have generally triggered the most discussion about the exploratory modeling results, especially in discussions with stakeholders outside the core team of model developers who are intimately familiar with the mathematical structure of the core model.

Just as a good story can be enhanced by working with an editor to iteratively make revisions and enhancements, so too can exploratory modeling be made more effective by reworking scopes and models. TMIP-EMAT applications are best conducted as iterative processes with the expectation that the set of policy levers, exogenous uncertainties, and performance measures will be refined in response to the results of earlier exploratory analyses. Ideally, the scoping exercises outlined at the beginning of this section are not a one-time event at the beginning of the modeling processes, but can be repeated again after some results are available, allowing all stakeholders to have a voice in how the exploratory scope can or should be refined or expanded. For example, sensitivity tests may show an unreasonable or insubstantial response to a changing input, or the model response may inspire an interest to investigate another aspect of the outputs more closely. It is common for users to discover that their core model behaves in a manner somewhat different from a priori expectations, and to want to drop some inputs and add some others after an initial analysis. This is especially common when modifying core models to represent changes in technology or new modes, such as automated or connected vehicles. Since the nature and magnitude of these changes is so deeply uncertain, reviewing a first attempt at modeling these effects can easily inspire new ideas from a variety of stakeholders about other possible impacts that might be modeled.

Incorporating TMIP-EMAT into model development early in the planning process also can be useful, as the core model can be designed with intention to enable the kinds of manipulations that TMIP-EMAT facilitates. The interface used to programmatically connect TMIP-EMAT with a core model is fairly flexible, but the development and use of that interface can be streamlined if the core model is built specifically to accommodate it. One especially convenient approach is to consolidate all of the relevant hooks or settings that TMIP-EMAT might manipulate into a single consolidated configuration file, or a limited number of such files arranged in a well-structured and stable environment, so that each policy lever or exogenous uncertainty can be changed easily. One particularly challenging situation that should be avoided if possible is having TMIP-EMAT alter a model parameter in a file that is regularly overwritten by the core model during its operation, as this may make it difficult to ensure the change sticks during the entire core model run.

Incorporating TMIP-EMAT into model development early in the planning process also can be useful, as the core model can be designed with intention to enable the kinds of manipulations that TMIP-EMAT facilitates.

TMIP-EMAT tools can effectively present information from the core model and encourages analysts to think about the implications of the model assumptions on the results. The results from this level of modeling interactivity tend to encourage analysts and other stakeholders to think critically about the assumptions embedded in travel demand models—both the assumptions explicitly addressed as exogenous uncertainties, and other assumptions that are baked into the design and application of the core models. By explicitly expressing a number of exogenous uncertainties, stakeholders may be prompted to raise questions about other uncertainties and become hesitant to draw real conclusions from the results because of other assumptions. The limitations of these assumptions can become more obvious than they would have been with a simpler scenario planning approach. This also highlights the need to proceed with exploratory analysis iteratively, as once the importance of these particular important assumptions becomes clear, it can be helpful to return to the model and incorporate them explicitly in the analysis.

4.0 Case Studies of Travel Model Improvement Program-Exploratory Modeling and Analysis Tool Scoping and Analysis

This chapter puts the exploratory analysis steps into greater focus using specific examples from the Beta Test Report (Milkovits et al., 2019) to illustrate how to choose specific features of the analysis. Additional details regarding these examples can be referenced from the Beta Test Report. Here, the focus is on translating the high-level scope of the exploratory analysis to help frame the specific policies and uncertainties that are explored, setting up the exploratory experiments, aggregating and interpreting results, and applying those results to address policy goals.

4.1 Case Study 1—Oregon Department of Transportation

Oregon Department of Transportation (ODOT) was motivated to test TMIP-EMAT to support analysis of future technologies where little or no observed data exist to estimate and validate models, using the Southern Oregon Activity-Based Model. This example was originally documented in the Beta Test Report.

Scoping

The first step of the scoping process is defining goals and objectives of the analysis. In ODOT's scoping workshop, the group started with the goal of providing equitable and accessible transportation system for all income groups. With this overarching goal in mind, the group outlined the following policy options for consideration:

- Transit system enhancements through the investment in fixed-route system and/or a collaboration with private TNC services.
- Incentivizing transit-oriented development (TOD).
- Pricing mechanisms on roadways (road user charge system, toll/managed lane facilities) on TNC/auto-based mobility services, on transit through fare subsidies, or through parking fees.
- Investment in active transportation modes, possibly through micromobility programs.
- Mobility as a Service (MaaS).
- Incentives for electric vehicles.

In order to evaluate the efficacy of these various policy options, a number of metrics was discussed and considered in the ODOT scoping workshop. The metrics were identified on the basis of the policies discussed above and how well they relate to the underlying goals behind these policies. Metrics that were considered included the following:

- Accessibility measures, ideally segmented by income, that capture the travel time to employment and services with a multimodal lens.

- Measures of congestion.
- Standard metrics of vehicle miles traveled, person miles traveled, and vehicle hours traveled.
- Mode shares by demographic segment.
- Household expenditures on transportation by income segment.
- Total time spent traveling.
- Out of home activities (number and duration).
- Revenue from user fees/transit.
- Transit ridership.
- Safety/reliability/exposure.

Lastly, a number of exogenous variables was identified that could impact the way in which or whether the various policies would be effective. These are the uncertainties that are outside the analyst or decisionmaker's control. Those considered in the ODOT scoping workshop include the following:

- Vehicle technology (autonomous and connected vehicles).
- Supply side impacts on capacity through more (or less) efficient use of existing roadways.
- A changing auto ownership model that could support more family sharing and fewer autos per household or even a fully fleet shared paradigm.
- The demand side disutility of in-vehicle time may decrease as autonomous vehicle (AV) passengers are able to use their time productively.
- Cost changes (vehicle operating and parking).
- User cost and availability of TNCs.
- Land use and demographic changes, including total growth, shifting income and age distributions, spatial distribution, and density/zoning changes.
- Changes in the larger economy that would impact household spending power, travel costs, and work habits (more telecommuting).
- Management of curb facilities to facilitate local delivery.
- Freight operational changes for local delivery, as well as long/medium haul.

Once these high-level scoping elements were identified, a key piece of the scoping process is fitting these elements within the context of what the model is capable of analyzing. This requires an understanding of the model and the inputs and variables it uses. Table 2 provides the specific model elements that were considered with respect to policy levers by the ODOT group.

Once these high-level scoping elements were identified, a key piece of the scoping process is fitting these elements within the context of what the model is capable of analyzing. This requires an understanding of the model and the inputs and variables it uses.

Table 2. Oregon Department of Transportation model variable identification by lever.

Lever	Potential Model Variables to Represent Lever
Transit System Enhancements	• Transit lines. • Transit headways. • Transit travel times. • Transit bias coefficients. • Transit fares. • Increase park-and-ride availability. • Restructure walk-connection for mid-range to represent micromobility availability. • Synthesize transit skim to represent TNC collaboration.
Pricing	• Income-specific auto operating costs. • Facility-specific tolls by occupancy, area, and time of day. • Transit fare by route, district-district connections, vary by person attributes. • Parking rates for work and nonwork. • New TNC mode alternative. • Park-and-ride lot fee.
Active Transportation	• Increase bike and walk speeds. • Change the maximum distance threshold for nonmotorized modes. • Enhance active network connectivity. • Vary nonmotorized bias constant.
Mobility as a Service	• Allow zero-auto households to use drive-alone modes. • Revise treatment of households with fewer vehicles than workers and/or drivers.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

For most policy levers, multiple model variables were identified, but for a couple, no model variables were identified. These were cases where the model was simply not designed to analyze the impacts of a given type of policy, including incentives for TOD and incentives for electric vehicles.

For each of the identified uncertainties, the ODOT group also identified specific model features and inputs that could be used to analyze those uncertainties, as shown in table 3.

Table 3. Oregon Department of Transportation model variable identification by exogenous uncertainty.

Exogenous Uncertainty	Potential Model Variables to Represent Uncertainty
Vehicle Technology Impacts on Operations	• Capacity by facility type, intersection versus lane capacity.
Vehicle Technology Penetration	• Simulate as part of synthetic population generation. • Incorporate a new model component. • Implement average values proportional to penetration rates.
Zero-Occupancy Vehicles	• Post-processing of trip tables. • Post-processing of aggregate VMT. • Develop new autonomous vehicle routing model.
Vehicle Technology Impacts on Behavior	• Modify time and cost coefficient.
Electric Vehicle Impact on Fuel Costs	• Auto operating costs associated with electric vehicles.
Vehicle Technology Impact on Parking costs	• Factor applied to default parking costs associate with a simulation of automated vehicle availability .
Vehicle Technology Impacts on Operations	• Capacity by facility type, intersection versus lane capacity.
New Mobility Services and Increased Use of TNCs	• Would require substantial changes to model and was dropped.
Land Use	• Zonal employment. • Modify synthetic population (control totals by geography).
Economy	• Modify jobs (zonal employment) and workers (synthetic population). • Vary income distribution in synthetic population. • Vary transit level of service. • Reduce work tours.
Curb Management	• Terminal times. • Parking costs. • Availability and alignment of centroid connectors.
Freight	• Direct changes to heavy truck trip table. • Replace simulated personal shopping trips with truck trips. • Reduce commercial vehicle and personal shopping to represent drone delivery.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

After considering the work required to implement each lever, uncertainty variable, and metric, the group selected the model variables that would be leveraged through TMIP-EMAT. In this process, it was important to recognize that for independent model input that needs to be modified, approximately 10 core model runs are required. The cost of adding additional layers of model runs was weighed against the importance of each additional varying input.

For independent model input that needs to be modified, approximately 10 core model runs are required. The cost of adding additional layers of model runs was weighed against the importance of each additional varying input.

Table 4 summarizes the selected levers and exogenous uncertainties. Originally, nine levers and uncertainties were scoped. Through the development and testing process, two uncertainties were dropped.

Table 4. Oregon Department of Transportation selected levers and uncertainty variables.

Policy-Lever/ Uncertainty Variable	Minimum	Default	Maximum	Distribution (Applies to Exogenous Uncertainties Only)	Unit/Correlations/ Other Notes
Lever: Transit Everywhere (Synthesize transit skim to represent TNC collaboration)	NA	Current transit system	Transit all over replaces fixed-route system	NA	Originally envisioned as a single lever, later segmented into two levers with the Transit level of service (LOS) continuous variable lever changing Transit Alternative-Specific Constant (ASC), and this Boolean lever changing the availability of transit.
Lever: Transit LOS (In-vehicle travel time (IVTT) equivalent change in transit utility)	-10.0	0	10.0	NA	Applies to both base fixed-route transit, as well as transit everywhere.
Lever: Parking rate factor on existing parking	0.5	1.0	20	NA	Existing parking rates are factored up by the factor provided.
Lever: Active transport improvements—factor applied to walk/bike speeds	1	1	2	NA	Changes only to speed, maximum distance is maintained; proxy for micromobility penetration.
Ex. Uncertainty: Interstate (access controlled) Capacity—vehicles per hour per lane	1,500	1,900	3,000	Uniform	A proxy for AV penetration and impact on access controlled facility (Interstate) capacity.

Table 4. Oregon Department of Transportation selected levers and uncertainty variables (continuation).

Policy-Lever/ Uncertainty Variable	Minimum	Default	Maximum	Distribution (Applies to Exogenous Uncertainties Only)	Unit/Correlations/ Other Notes
Ex. Uncertainty: Auto operating cost—cents per mile	1.0	12.4	25.0	Uniform	Low represent electric vehicle efficiency, high represents fleet AVs and higher pricing (tax) structures.
Ex. Uncertainty: Household income multiplier	0.5	1.0	1.5	Uniform	Was used as a simple method to represent changing jobs, job type, household worker mix, etc.
Ex. Uncertainty: Value of time (change in sensitivity to IVTT)	0.5x	1x	1.2x	Would make higher sensitivity less likely	Uncertainty variable was dropped when found to be perfectly correlated with auto operating costs.
Ex. Uncertainty: Household (HH) densification (% shift distance to the center)	0.5x from the core	1x from the core	1.5x from the core	Uniform	Prototype implementation defined rings, chose number of houses to shift by ring, and did a ring jump. Testing following full Latin HyperCube Sampling (LHS) runs showed unreasonable responses and variable was removed from scope.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

Table 5 provides the scoped metrics for ODOT's exploratory analysis during the scoping step. Again, these metrics were selected to measure the effectiveness of the set of policies on the goals of the analysis, which were to provide an equitable and accessible transportation system to all users.

Table 5. Oregon Department of Transportation scoped performance metrics.

Metric
Percentage of population with access to 50k jobs by car within 20 minutes in performance metrics (PM)
Bike and walk mode share
Transit with park-and-ride and kiss and ride mode share
Millions of person miles traveled
Millions of vehicle miles traveled in PM
Millions of auto miles traveled
Millions of truck miles traveled
Millions of vehicle miles traveled
Thousands of vehicle hours traveled in PM
Thousands of auto hours traveled
Thousands of truck hours traveled
Thousands of vehicle hours traveled
Percent of Interstate miles over 90% volume-to-capacity (V/C) ratio during the PM peak
Percent of principal arterial miles over 90% V/C ratio during the PM peak
Percent of minor arterial miles over 90% V/C ratio during the PM peak
Number of autos owned per household
Percent of nonmandatory tours

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

Interfacing between Travel Model Improvement Program-Exploratory Modeling and Analysis Tool and the Core Model

ODOT staff built the API implementation in such a way that interfaced with both R and Python. The structured API that organizes the interaction points between TMIP-EMAT and the core model were built as part of this process. ODOT also leveraged their model development consultant to assist with development of model-side API functionality to enable a programmatic control of model inputs.

Univariate sensitivity tests were conducted using TMIP-EMAT's univariate experimental design process to test that the model was appropriately sensitive to each uncertainty variable and policy lever. ODOT produced a set of R summaries that input the experiment dataframe to comprehensively compare all metrics of each experiment against the baseline metrics, and to compare each metric variation by lever and uncertainty variable. Figure 32 provides an example of these tests, where a capacity increase in freeways was evaluated against the key performance metrics.

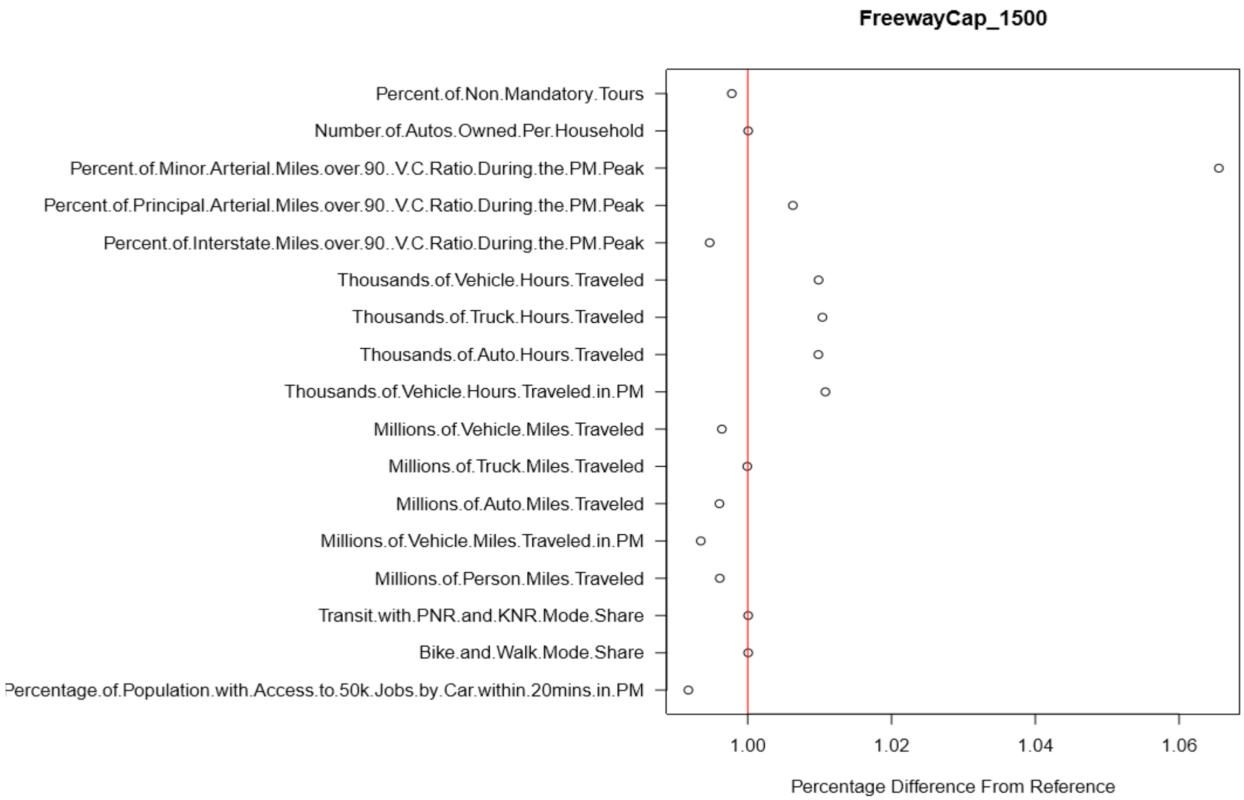


Figure 32. Chart. Comparison of single experiment across all metrics.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

Ultimately, the univariate sensitivity tests were useful as an initial confirmation of the correct operation of the model, and to gauge the degree of variation for each metric. It is after the univariate sensitivity tests are complete that we begin to see if the metric selection is appropriate.

The univariate sensitivity tests were useful as an initial confirmation of the correct operation of the model and to gauge the degree of variation for each metric.

Experimental Design

The TMIP-EMAT software includes an experimental design feature that automatically generates a *Latin Hypercube design of experiments*, which includes the input values for each policy lever and uncertainty variable across the set of model runs to be performed. This method ensures that the range of inputs to the experiments adequately covers the input domain and does not leave large information gaps for certain combinations of inputs. As a result, the analyst can be more confident in the performance metric results and relationships that are output from the core model runs. ODOT's experiments were run across multiple computers with the scope and results saved in a common SQLite database on a single machine.

There was a total of eight policy levers and uncertainty variables in ODOT's scope. For ideal meta-model development, 10 experiments should be run for each policy lever and uncertainty variable. Thus, the experimental design set input values for 80 core model runs. As noted above, the specific input values for the 80 model runs were determined by the TMIP-EMAT software.

It is worth noting that there is an opportunity for errors to be introduced when moving from individual sensitivity test runs to the full set of experiments. ODOT ended up running the full set of experiments three times due to several run and setup issues that occurred:

- In the first run of the full set of experiments, ODOT conducted a review after 20 runs were completed and discovered issues in how the model inputs were set.
- In the second run of experiments, all 80 runs were completed, but the review revealed an issue with the land use density variable, which ultimately led to the removal of this variable from the exploratory analysis. After removing that variable, 10 fewer runs were needed and the experimental design was updated to include only 70 runs, which were completed and used for the analysis workshop.
- The availability of network licenses was an issue at times while running the core model runs. ODOT also needed access to licenses for other ongoing studies, which required management of the license utilization throughout execution of core model runs.
- ODOT also experienced random crashes of the model during the core model execution due to configuration issues with the VISUM model software. The software vendor, PTV, was able to address these issues, which allowed the complete set of experiments to run in the final iteration of the experimental design execution.

There is an opportunity for errors to be introduced when moving from individual sensitivity test runs to the full set of experiments.

As noted above, the land use variable was ultimately removed from the analysis. In this particular example, the land use variable included the development of a mechanism to move development from the downtown areas to the suburban ring areas. While this mechanism passed the initial sensitivity tests that ran at the extreme values, when the model was run with the full set of experiments, the household density produced a nonmonotonic and unreasonable response in several performance measures; most notably the nonmotorized mode share (see results from the *display experiments* tool in figure 33). After several rounds of troubleshooting that, it was ultimately determined that the best path forward was to remove the variable. Issues like this can point to problems with sensitivities in the underlying model that could be addressed at a later point or as part of the TMIP-EMAT work.

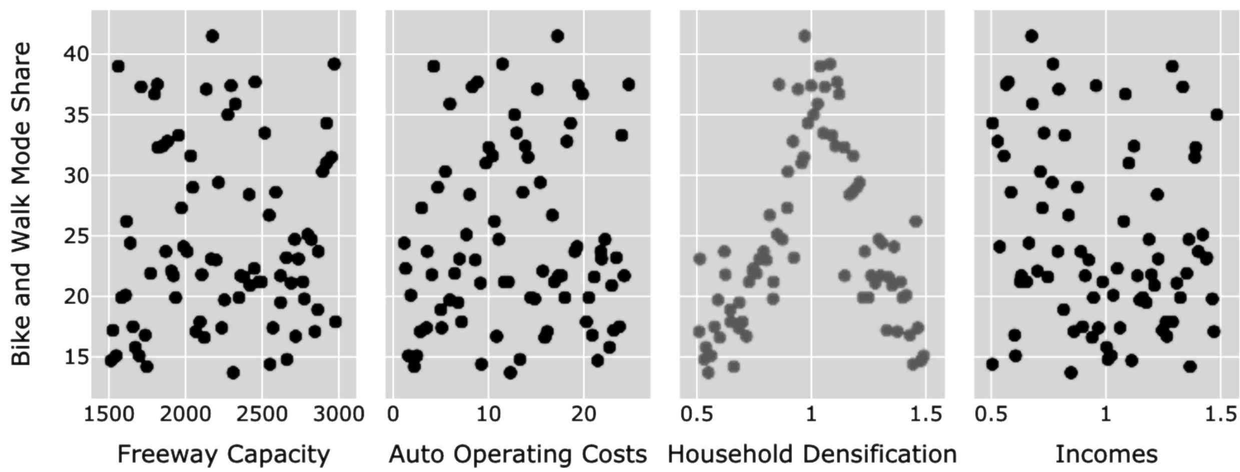


Figure 33. Graph. Oregon Department of Transportation bike and walk share sensitivity to household density.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

As described above, any number of issues can emerge when executing the core model runs using an automated API. While several challenges emerged in ODOT's implementation of the experimental design runs, the use of multiple systems gave ODOT a lot of flexibility and responsiveness to complete the model runs in a timely manner even with these setbacks. Moreover, having overcome those challenges, future implementations of EMAT with the ODOT model will likely proceed much more smoothly.

Exploring Results

Once the experimental design runs of the core model were completed, the results were aggregated and interpreted using some of the TMIP-EMAT tools discussed in section 3.4.

ODOT used the *interactive visualizations* and *feature scoring* to determine whether policy levers used in the analysis were important factors for key metrics. In the visualization exploration, ODOT went back to the original goal (providing accessibility to all groups) and investigated metrics that supported the ultimate goal. For example, reducing congestion is an intermediate goal, but improved accessibility may mean more overall auto travel.

ODOT also used the *multi-objective optimization* and *robust optimization* utilities of EMAT to reveal tradeoffs between nonmotorized and transit mode shares. ODOT noted that the optimization utilities required that solutions optimize at least one performance metric, while ODOT preferred to find robust solutions that performed well across a collection of performance metrics. This led ODOT to consider composite performance metrics that measure the relative performance of the other performance metrics in a single metric. By doing so, the directed search can be transformed into a solution-finding utility that optimizes the system performance in a more ideal way across the set of primary performance metrics that ODOT was actually interested in.

4.2 Case Study 2—Greater Buffalo Niagara Regional Transportation Council

The Greater Buffalo Niagara Regional Transportation Council (GBNRTC) was motivated to use TMIP-EMAT to evaluate investments along a specific corridor in the region using their four-step travel model. This example was originally documented in the Beta Test Report.

Scoping

GBNRTC defined their goals for their TMIP-EMAT analysis on the basis of focus areas in their Regional Transportation Plan (RTP). The goals they defined included the following (the bullet heading denotes one of the focus areas of the RTP, and the subbullets denote specific goals around that focus area):

- Using transportation investments to strengthen communities using existing infrastructure:
  - Increase accessibility to influence land use.
  - Increase access to services for general population and communities of concern (high-poverty zip codes).
  - Increase multimodal access to neighborhood services.
  - Increase active transportation options.
  - System safety for all modes.
  - Improve access to parks, greenways, and waterfronts.
- Creating opportunities for economic development and supporting access for the workforce in the region:
  - Reduce freight delays.
- Improving mobility using technology:
  - Decrease lane-miles with underutilized capacity.
  - Decrease impervious surfaces.
  - Decrease VMT.
- Protecting the natural environment:
  - Increase lane miles of connected corridors.
  - Improve reliability.

During the scoping workshop with GBNRTC, these goals were used as the foundation for developing specific transportation policies. The policies that were considered are shown in table 6.

Table 6. Policy ideas from the Greater Buffalo Niagara Regional Transportation Council scoping workshop.

Scope	Focus Area	Policy
Corridor level	Transit and nonmotorized improvements	• Complete streets with potential higher-transit service. • Mobility hubs. • Transit signal priority (TSP) and bus priority. • Cycle track.
Corridor level	Land use	• Encourage redevelopment. • Densification of land use.
Corridor level	Freight	• Binational, green, autonomous, freight corridor. • Distribution centers.
Corridor level	Roadway improvements	• Adaptive signal control in coordination with highways. • Support for new vehicle technology (e.g., autonomous, mixed, connected vehicles).
Regionwide	All	• Increase pretrip information. • Shared-use mobility services. • Regional cycle network. • MaaS

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

As noted earlier, GBNRTC decided to focus this work on specific corridors. While three corridors were considered, the Bailey Avenue corridor was selected for the exploratory analysis. Within this corridor, the following policy levers were determine to be relevant for the analysis:

- Bus Rapid Transit (BRT) with dedicated lane.
- BRT without dedicated lane.
- Transit mobility hubs scenarios.
- Parking changes (restricted on-street and system controlled).

In order to evaluate these types of policies, a number of metrics was identified in the scoping workshop as well. Given the focus on the Bailey Avenue corridor, corridor-level metrics were considered more appropriate in many cases than regional metrics. The key metrics that were identified for the study include the following:

- Daily VMT, VHT, and delay by time of day.
- Regional trips to/from the Bailey Corridor.
- Daily corridor route ridership and transit mode share.
- Daily nonmotorized mode share.
- Employment within 20 minutes of the Bailey corridor.

- Regional VMT.
- Total transit boardings.
- Regional transit and nonmotorized mode shares by time of day.

Having narrowed the scope of the exercise considerably, the next step in the scoping workshop was to identify uncertainties that may impact the effectiveness of the strategies. The following set of uncertainties were identified:

- Land use: Shift in employment locations and density of new development.
- Demographics: Aging population and income distributions.
- Vehicle technology: Mix of connected, automated technology available and capabilities.
- Mobility services, including TNCs and micromobility.
- Climate and weather impact on nonmotorized modes.
- International travel demand in response to change in the currency exchange and how the border crossings will operate.

The next step in the workshop was to identify the existing or needed model functionality to represent each lever and uncertainty variable in the model. Table 7 and table 8 identify all potential model variables for each lever and uncertainty. As part of the discussion, the group discussed the level of effort involved in developing new model functionality.

Table 7. Greater Buffalo Niagara Regional Transportation Council model variable identification by lever.

Lever	Potential Model Variables to Represent Lever
Roadway reconfiguration	• Highway network geometry. • Lane configuration. • Functional class and capacity attributes. • Speed. • Intersection delay.
Promote nonmotorized modes	• Make links available for walk and bike modes. • Centroid connector number and location. • Nonmotorized speeds (micromobility).
Mobility hubs	• Add park-and-ride availability at key stops.
Transit enhancements	• Travel time improvements (TSP and BRT). • Headway improvements. • Stop frequency. • Access/egress improvements. • Add new transit route.
Parking policies	• Terminal times (represent less on-street parking).

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

Table 8. Greater Buffalo Niagara Regional Transportation Council model variable identification by exogenous uncertainty.

Exogenous Uncertainty	Potential Model Variables to Represent Uncertainty
Land use/ demographics	• Employment level by segment (retail, wholesale, manufacturing, Government, service, office). • Household size and income segmentation. • Number of households by location along corridor. • School/university enrollment. • Development at key areas (e.g., Genesee node on Bailey).
Vehicle technology	• Supply side changes. • Roadway capacity. • Intersection delay. • Parking costs and terminal times to represent self-parking. • Electric vehicle reductions in operating costs. • Demand side changes. • In-vehicle travel time sensitivity. • Zero-occupancy vehicle travel generated as separate trip table. • New mobility services represented through vehicle availability levels.
Climate/weather	• Decrease walk speed. • Nonmotorized distance threshold. • Increase transit in-vehicle time. • Decrease roadway speeds/capacities. • Reduce parking capacity.
International	• Increased shopping trips across border. • Change enplanements at Buffalo Niagara International Airport. • Border crossing availability/capacity/delay.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

After considering the work required to implement each lever, uncertainty variable, and metric, the group selected the model variables that would be used for TMIP-EMAT, keeping in mind that the total number of policy levers and independent uncertainty variables has a linear relationship with the required number of core model runs; for example, 10 policy levers/ uncertainty variables require 100 core model runs, while the number of metrics has no impact on the number of core model runs required.

The scope was revised through the subsequent steps. GBNRTC identified four levers associated with the corridor: 1) improved transit headway, 2) micromobility options, 3) mobility hubs, and 4) reduced parking. They also identified four uncertainties: 1) land-use, 2) self-parking vehicles, 3) shared mobility, and 4) Inclement weather. Through the development and testing process, two levers were dropped because of unreasonable or insubstantial responses in the model. Table 9 summarizes the final selected levers and exogenous uncertainties.

Table 9. Greater Buffalo Niagara Regional Transportation Council selected levers and uncertainty variables.

Policy-Lever/ Uncertainty Variable	Minimum	Default	Maximum	Distribution (Applies to Exogenous Uncertainties Only)	Unit/Correlations/ Other Notes
Lever: Transit headway	True	False		NA	Half the headway on Bailey Avenue routes
Lever: Micromobility		False	True	NA	Improved access to transit stops; higher density of transit stops along Bailey corridor.
Lever: Mobility hubs		False	True	NA	Every other stop on Bailey Avenue is a PNR lot.
Lever: Reduced parking		False	True	NA	Parking on Bailey is moved to side streets by increasing terminal time for auto.
Ex. Uncertainty: Bailey land use	0	0	1	Uniform	0 = base 2025 forecast; and 1 = full build out of vacant lots along corridor.
Ex. Uncertainty: Self-parking	False	False	True	Binary	True = all terminal times are set to zero; and false = base model terminal times related to land-use density.
Ex. Uncertainty: Shared mobility	0	0	1	Uniform	0 = calibrated distribution of zero and insufficient vehicle households; and 1 = all households are treated as having sufficient vehicles.
Ex. Uncertainty: Weather impacts	0	0	1	Binary w/ 90% = 0; 10% = 1	0 = base capacity and walk speed; and 1 = 75% decrease in highway capacity, walk speed.

(Source: TMIP-EMAT Beta Test by Oregon DOT.)

Interfacing between Travel Model Improvement Program-Exploratory Modeling and Analysis Tool and the Core Model

Coordinating changes across all the networks and managing the files was an opportunity for error and required careful testing.

Because the GBNRTC model was utilized in the initial TMIP-EMAT proof-of-concept work, an API had already been developed to interface with the model, so no additional work was needed there. However, additional model functionality was required in order to efficiently update the specific model inputs that would be changed in the set of core model runs that were to be performed. Of particular note, the GBNRTC micromobility, transit TSP, and BRT levers were represented through changes in the transit and highway networks (Note that the transit TSP and BRT levers were removed from the final analysis due to issues with the transit representation.). Coordinating changes across all the networks and managing the files was an opportunity for error and required careful testing.

Univariate Sensitivity tests were conducted to verify that input variables had appropriate effects on performance metrics. The univariate sensitivity tests were useful as a first test in assessing the reasonableness of the levers, uncertainties, and metrics; and in a couple of cases, the univariate tests alerted GBNRTC of elements of the analysis that were not working properly, and they were fixed prior to running the full set of experiments in the next step.

Experimental Design

The TMIP-EMAT software creates the experimental design to determine the set of input variable levels for each core model run. The GBNRTC API was able to automatically run the set of core model runs on a single computer.

In preparing the model results, it was concluded that the performance of the transit TSP and BRT strategies was skewing some of the analysis results, specifically for scenario discovery and *multi-objective optimization* analyses. As a result, these strategies were removed from the analysis, and the set of experiments was rerun using the set of strategies outlined in the previous section.

Exploring Results

Overall, GBNRTC found *feature scoring* to be a useful tool for analysis. The visualizations of the feature scores were useful for showing the broader view of how the levers and uncertainties affected the performance metrics and helped explain what lever/uncertainty dominated the results. Rather than use the feature scoring results directly, GBNRTC used the feature scoring results to inform a larger story that would be told to decisionmakers. Similarly, GBNRTC used the scatterplots and histograms from the *interactive vislualizer* to illustrate the results visually to policymakers and the public.

The “Lasso” feature in the scatter plot of the *interactive visualizer* was used by GBNRTC to select a set of scenarios and scroll through different plots to see where they lie within the range of other metrics, uncertainties, and levers. This feature allowed for assessing best/worst case outcomes and outliers, understand what other uncertainties or levers have the biggest impact on metrics, and assess which levers/uncertainties are driving the interaction with other levers/ uncertainties (i.e., driving outliers).

5.0 Conclusion

TMIP-EMAT tools can effectively present information from the core model and encourages analysts to think about the implications of the model assumptions on the results. The results from this level of modeling interactivity tend to encourage analysts and other stakeholders to think critically about the assumptions embedded in travel demand models—both the assumptions explicitly addressed as exogenous uncertainties and other assumptions that are baked into the design and application of the core models. By explicitly expressing a number of exogenous uncertainties, stakeholders may be prompted to raise questions about other uncertainties and become hesitant to draw real conclusions from the results because of other assumptions. The limitations of these assumptions can become more obvious than they would have been with a simpler scenario planning approach.

Ultimately, TMIP-EMAT was developed to support robust decisionmaking for transportation. The multi-dimensional and multi-objective nature of the tools pairs well with the nuanced and multi-faceted nature of the transportation system. The tools provided by TMIP-EMAT, and by exploratory modeling more generally, are meant to be descriptive, and not prescriptive. Each tool is built to provide insight into the relationships between policy levers, uncertainties, and performance measures. The end goal of the analysis is not to provide explicit guidance about what decisions are “best” but to allow decisionmakers and other stakeholders to better grasp the tradeoffs between different courses of action, and help to facilitate and focus discussion about policy actions.

TMIP-EMAT can be a valuable tool for proactive, iterative, continuous, and comprehensive transportation planning, especially under the conditions of deep uncertainty. It is a tool that planning agencies can use to facilitate community engagement and visioning process, and enable effective communication among technical analysts, planners, various stakeholders and decisionmakers. With appropriate robust core models, TMIP-EMAT empowers communities to conduct a much broader question-driven exploration, leading to decisions that are robust in a wide range of futures at a time when planners must deal with deep uncertainty. With the gained insights of potential, possible, plausible, probable, or preferred futures, policy-makers and stakeholders can effectively plan, prepare, mitigate, adapt and shape their strategies based on the community value and vision.

6.0 References

Adler, T., V. Bernardin, J. Dumont, L. Flake, and H. Sadrsadat (2017). The Promise and Limitations of Locational App Data for Origin-Destination Analysis: A Case Study. TMIP, Federal Highway Administration.

Alexander, L., S. Jiang, M. Murga, and M. Gonzalez (2015). Origin-Destination Trips by Purpose and Time of Day Inferred from Mobile Location Data. Transportation Research Part C, 58, 240–250.

Ban, X., C. Chen, F. Wang, J. Wang, and Y. Zhang (2018). Promises of Data from Emerging Technologies for Transportation Applications: Puget Sound Region Case Study. Travel Model Improvement Program (TMIP), FHWA.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research*, Vol. 41, No. 3, 1993, pp. 435–449.

Bernardin, V. (2018). *How-to: Think About Model Design for Your Region*. Prepared for Federal Highway Administration.

Bureau of Labor Statistics (2019). American Time Use Survey, Tables 6 and 7. https://www.bls.gov/news.release/atus.t06.htm, https://www.bls.gov/news.release/atus.t07.htm.

Cambridge Systematics, Inc. (2010). *Travel Model Validation and Reasonableness Checking Manual: Second Edition*. Prepared for Federal Highway Administration.

Cambridge Systematics, Inc. (2016). Emerging Technology, Demographic Changes, and Travel Behavior.
http://www.fsutmsonline.net/index.php?/user_groups/comments/future_mobility_serpm.

Cambridge Systematics, Inc. (2018). *California High-Speed Rail 2018 Business Plan*. Prepared for California High-Speed Rail Authority. https://hsr.ca.gov/wp-content/uploads/docs/about/business_plans/2018_CA_High_Speed_Rail_Business_Plan_Ridership_and_Revenue_Risk_Analysis.pdf.

Cambridge Systematics, Inc. and Massachusetts Institute of Technology (2018). *NCHRP Report 868: Cell Phone Location Data for Travel Behavior Analysis*. National Cooperative Highway Research Program, Washington, DC, TRB of the National Academies.

Castiglione, J., T. Chang, D. Cooper, J. Hobson, W. Logan, E. Young, B. Charlton, C. Wilson, A. Mislove, L. Chen, and S. Jiang (2016). TNCs Today: A Profile of San Francisco Transportation Network Company Activity. San Francisco County Transportation Authority (SFCTA).

Chen, C., X. Ban, F. Wang, J. Wang, C. Siddique, R. Fan, and J. Lee (2017). Understanding GPS and Mobile Phone Data for Origin-Destination Analysis, TMIP, Federal Highway Administration.

Childress S., B. Nichols, B. Charlton, and S. Coe (2015). Using an Activity-Based Model to Explore the Potential Impacts of Automated Vehicles. Transportation Research Record. 2015;2493(1):99-106. doi:10.3141/2493-11.

DeSilver, D. (2020). Before the Coronavirus, Telework Was an Optional Benefit, Mostly for the Affluent Few. Pew Research, accessed July 2020 from https://www.pewresearch.org/fact-tank/2020/03/20/before-the-coronavirus-telework-was-an-optional-benefit-mostly-for-the-affluent-few/.

Dias, F., P. Lavieri, V.Garikapati, S. Astroza, R. Pendyala, and C. Bhat (2017). A Behavioral Choice Model of the Use of Car Sharing and Ride-Sourcing Services. Transportation 44, 1307–1323 (2017). https://doi.org/10.1007/s11116-017-9797-8

Federal Transit Administration (2008). Mobility Benefits from Changes in Unmeasured Transit Attributes, Appendix A: Technical Methods for Alternatives Analysis, Information for the Case for a Project.

Friedman, J., and N. Fisher (1999). Bump hunting in high-dimensional data. *Statistics and Computing* 9, 123–143 (1999). http://statweb.stanford.edu/~jhf/ftp/prim.pdf.

Friedrich, B. (2016). The Effect of Autonomous Vehicles on Traffic. In: M. Maurer, J. Gerdes, B. Lenz, and H. Winner (eds) *Autonomous Driving*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-48847-8_16.

Gehrke, S., A. Felix, and T. Reardon (2018). Fare Choices: A Survey of Ride-Hailing Passengers in Metro Boston. Metropolitan Area Planning Council.

Gucwa M. (2014). Mobility and energy impacts of automated cars. In *Proceedings of the Automated Vehicles Symposium*, San Francisco.

Guyot, K., and I. Sawhill (2020). Telecommuting Will Likely Continue Long after the Pandemic. Brookings Institute. Accessed July 2020 at https://www.brookings.edu/blog/up-front/2020/04/06/telecommuting-will-likely-continue-long-after-the-pandemic/.

Jiang, J. (2019). More Americans Are Using Ride-Hailing Apps, Pew Research Center. Accessed July 2020 from https://www.pewresearch.org/fact-tank/2019/01/04/more-americans-are-using-ride-hailing-apps/.

Kohli, S., and L. Willumsen (2016). Traffic Forecasting and Autonomous Vehicles. Proceedings of the 2016 European Transport Conference, Barcelona, Spain.

Komanduri, A., Z. Wafa, K. Proussaloglou, and S. Jacobs (2018). Assessing the Impact of App-Based Ride Share Systems in an Urban Context: Findings from Austin. Transportation Research Record, 2672 (7), 34-46.

Kroger, L., T. Kuhnimhof, and S. Trommer (2016). Modeling the Impact of Automated Driving—Private Autonomous Vehicle Scenarios for Germany and the U.S. Proceedings of the 2016 European Transport Conference, Barcelona, Spain.

Kwakkel, J. (2017). The Exploratory Modeling Workbench: An Open Source Toolkit for Exploratory Modeling, Scenario Discovery, and (Multi-objective) Robust Decision-making. *Environmental Modeling and Software*, Vol. 96, 2017, pp. 239–250.

Kwakkel, J., and M. Jaxa-Rozen (2016). Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. *Environmental Modeling & Software*, Volume 79, 2016, pp. 311-321, ISSN 1364-8152. https://doi.org/10.1016/j.envsoft.2015.11.020 (https://www.sciencedirect.com/science/article/pii/S1364815215301092).

LaMondia, J., D. Fagnant, H. Qu, J. Barrett, and K. Kockelman (2016). Long-Distance Travel Mode Shifts Due to Autonomous Vehicles: A Statewide Mode-Shift Simulation Experiment and Travel Survey Analysis, Transportation Research Record, 2566, 1-11.

Lemp, J., J. Zang, X. Y. Chang, A. Komanduri, M. Xyntarakis, D. Von Stroh, and K. Proussaloglou (2019). Assessing the Impact of Mobile Data Processing Algorithms and Expansion Methods on the Inference of Travel Behavior Metrics. Proceedings of the 98th Annual Meeting of TRB.

Lempert, R., S. Popper, and S. Bankes (2003). *Shaping the Next One Hundred Years*, RAND Corporation, Santa Monica, CA, pp. 3-4.

Lempert, R. (2019). Robust Decision-making (RDM). 10.1007/978-3-030-05252-2_2.

Litman, T. (2017). Transportation Cost and Benefit Analysis II—Travel Time Costs. Victoria Transport Policy Institute (VTPI).

Litman, T. (2020). Autonomous Vehicle Implementation Predictions: Implications for Transport Planning, Victoria Transport Policy Institute (VTPI).

Loeppky, J., J. Sacks, and W. Welch (2009). Choosing the Sample Size of a Computer Experiment: A Practical Guide. *Technometrics*, 2009, pp. 366–376.

Maat, K. and R. Konings (2018). Accessibility or Innovation? Store Shopping Trips versus Online Shopping. Transportation Research Record 2018, Vol. 2672(50) 1–10.

McGuckin, N. and A. Fucci (2018). Summary of Travel Trends: Findings from the 2017 NHTS, available at: http://onlinepubs.trb.org/onlinepubs/Conferences/2018/NHTS/McGuckinTravelTrends.pdf.)

Milkovits, M., R. Copperman, J. Newman, J. Lemp, and T. Rossi (2019). *TMIP Exploratory Modeling and Analysis Tool (TMIP-EMAT) Beta Test Results*. Prepared for Federal Highway Administration. https://www.fhwa.dot.gov/planning/tmip/publications/other_reports/emat_beta/.

Mokhtarian, P., and C. Chen (2004). TTB or Not TTB, That is the Question: a Review and Analysis of the Empirical Literature on Travel Time (and Money) Budgets. *Transportation Research, Part A*, 38:643-675.

National Academies of Sciences, Engineering, and Medicine (2018). Updating Regional Transportation Planning and Modeling Tools to Address Impacts of Connected and Automated Vehicles, Volume 1: Executive Summary. Washington, DC: The National Academies Press. https://doi.org/10.17226/25319.

Newman, J., R. Copperman, J. Lemp, D. Kurth, B. Lipkin, and M. Henley (2018). Gaussian Process Regression for Risk Analysis of Travel Demand Forecasts. Presented to Seventh International Conference on Innovations in Travel Modeling, Atlanta, Georgia. http://onlinepubs.trb.org/onlinepubs/Conferences/2018/ITM/JNewman2.pdf.

Rayle, L., D. Dai, N. Chan, R. Cervero, and S. Shaheen (2016). Just a Better Taxi? A Survey-Based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco. *Transport Policy*, 45, 168-178.

Roy, S., A. Komanduri, and K. Proussaloglou (2020). Evolution of Ride-Hailing and Ridesharing in Chicago: 2013–2018. Proceedings of the 99th Annual Meeting of TRB, January 2020.

Sacks, J., W. Welch, T. Mitchell, and H. Wynn (1989). Design and Analysis of Computer Experiments. *Statistical Science*, 1989, pp. 409–423.

Thibault, R. (2020). Uber versus Lyft: The Battle for the U.S. Rages On. SimilarWeb, accessed July 2020 from https://www.similarweb.com/corp/blog/uber-vs-lyft-the-battle-for-the-us-rages-on/.

Transportation Research Board (2007). *Special Report 288: Metropolitan Travel Forecasting Current Practice and Future Direction*. (http://onlinepubs.trb.org/onlinepubs/sr/sr288.pdf).

Transportation Research Board Committee on Travel Survey Methods. Travel Survey Manual. https://tfresource.org/surveymanual/. Accessed March 5, 2021.

Vovsha, P. (2017). Modeling Impacts of Autonomous Vehicles on Travel Demand with an Activity-Based Model. TRB National Transportation Planning Applications Conference, Raleigh, NC, 2017.

Wang, F., and C. Chen (2018). On Data Processing Required to Derive Mobility Patterns from Passively-Generated Mobile Phone Data. Transportation Research Part C, 87, 58–74.

Widhalm, P., Y. Yang, M. Ulm, S. Athavale, and M. Gonzalez (2015). Discovering Urban Activity Patterns in Cell Phone Data. Transportation, 42 (4), 597–623.

Zhou, Y. and K. Kockelman (2017). Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas, Proceedings of the 96th Annual Meeting of the Transportation Research Board (TRB), Washington, DC.

Zmud J., T. Williams, M. Outwater, M. Bradley, N. Lara, and S. Row (2018). NCHRP Report 896: Unpdating Regional Transportation Planning and Modeling Tools to Address Impacts of Connected and Automated Vehicles. Transportation Research Board.

Appendix 1. Alphabetical List of TMIP-EMAT Methodological Tools

- **CART**
  Classification and Regression Trees, or CART, is a simple machine learning technique for predicting a target variable. Within TMIP-EMAT, the CART algorithm is implemented as a scenario discovery method, which can be used to develop interesting boxes for model exploration.
- **Contrast Experiments**
  The contrast experiments method renders two different set of experiments on a common scatter plot matrix. This visualization approach makes it easy to see if the overall shape of the distribution of experiment inputs and outputs is similar or different. It is particularly useful in validating that TMIP-EMAT's automatically generated meta-models are performing correctly.
- **Display Experiments**
  The display experiments method generates a scatter plot matrix that displays model inputs (uncertainties and policy levers) in one dimension and model outputs (performance measures) in the other.
- **Feature Scoring**
  This is a scenario discovery method for identifying what model inputs have the greatest relationship to the outputs by computing a numerical value that summarizes the relative importance of each input in determining the level of the output.
- **Exploratory Scoping**
  While not a methodological approach per se, TMIP-EMAT provides a notational structure to concretely define the manner in which the XLRM framework is to be operationalized for a given travel model (R) and its uncertainties (X), policy levers (L), and performance measures (M).
- **Interactive Visualizer**
  The Interactive Visualizer in TMIP-EMAT provides a set of tools that can display a dynamically generated selection of experiments in a number of visualizations, including histograms, scatter plots, and SPLOMs. The dimensional bounds of the select (the "box") can be manipulated by a user programmatically or by clicking and dragging directly on the figures.
- **Latin Hypercube Design of Experiments**
  A Latin Hypercube is a space-filling mathematical process for making pseudo-random draws from a multi-dimensional space. This kind of design is not formally "random," but approximates a random distribution while ensuing a reasonable coverage across the spectrum of possible values in each dimension. Meta-models for deterministic simulation experiments, such as most transportation models, are best supported by a "space filling" design of experiments such as this.

- **Meta-model Creation**
  A main feature of TMIP-EMAT is the ability to automatically generate meta-models that provide a good approximation of the underlying core model in most situations. By default, metamodels derived through TMIP-EMAT include two stages, a linear regression model to capture overall trends and a gaussian process regression (GPR) model that can capture a wide variety of non-linear effects.
- **Monte Carlo Simulation**
  A Monte Carlo simulation is a simple random (or in more precise computer science terminology, pseudo-random) process for generating a design of experiments. It is not generally an efficient design, but with a large enough sample size efficiency is less relevant and simplicity can be valuable.
- **Multi-objective Optimization**
  With exploratory modeling, optimization is also often undertaken as a multi-objective optimization exercise, where multiple and possibly conflicting performance measures need to be addressed simultaneously. Instead of generating one unique "optimal" solution, this TMIP-EMAT method can be used to find a spectrum of different solutions. Each of them solves the problem at a different weighting of the various objectives. Decisionmakers can then review the various different solutions, and make judgements about the various tradeoffs implicit in choosing one over another.
- **Policy Contrast**
  The Policy Contrast method in TMIP-EMAT allows an analyst to compare the outcomes of two different sets of policies. The tool runs the model across a distribution of inputs, and displays the resulting distribution of performance measure outputs. Two sets of model runs are generated with the same design of experiments for all the noncontrasted distributions, and so any variation in the performance measures can be unambiguously linked to the changes in the specific-value inputs, instead of being a result of input stochasticity.
- **PRIM**
  The Patient Rule Induction Method, or PRIM, is a scenario discovery method. This method is a "bump hunting" technique introduced by Friedman and Fisher (1999), which often provides insightful results for complex models.
- **Reference Experiment**
  A "design of experiments," which contains only a single experiment with all input values set to their default parameters.
- **Robust Optimization**
  Robust optimization is a variant of more traditional optimization problems. Rather than seeking a solution that provides the best outcome, a robust optimization problem is one where we try to find policies that yield good outcomes across a broad range of possible futures. It is common to employ various different criteria for what constitutes "good" or "broad" by also borrowing methods from the Multi-objective Optimization tools.
- **Search over Levers**
  A Search over Levers is a particular style of multi-objective optimization for exploratory modeling in the XLRM framework, where the uncertainties are held constant at some particular value, and only the policy levers are manipulated by the search algorithm.

- **Scatter Plot Matrix**
  The Scatter Plot Matrix, or SPLOM, is a visualization method. It is a collection of two-dimensional scatter plots arranged in a matrix, where each column of plots shares a common x-axis definition, and each row shares a common y-axis definition. The Display Experiments and Contrast Experiments tools in TMIP-EMAT create SPLOMs for one or two sets of experimental data, respectively.
- **Threshold Scoring**
  A variant of feature scoring, where inputs are scored not with just a single numerical value, but with a range of values representing the relative importance of inputs for getting the output to be above or below various possible threshold values.
- **Univariate Sensitivity Testing**
  One of the simplest experimental designs is a set of univariate sensitivity tests. In this design, a set of baseline model inputs is used as a starting point, and then input parameters are changed one at a time to non-default values. Univariate sensitivity tests are excellent tools for debugging and quality checking the model code, as they allow modelers to confirm that each modeled input is (or is intentionally not) triggering some change in the model outputs.
- **Worst Case Discovery**
  Worst Case Discovery is a particular style of multi-objective optimization for exploratory modeling in the XLRM framework. In this analysis, the policy levers are held constant at some particular value, and only the exogenous uncertainties are manipulated by the search algorithm. In addition, the directionality of all objective dimensions is inverted, so that the search algorithm seeks to find values for the input that lead to worse outcomes instead of better ones.

U.S. Department of Transportation
Federal Highway Administration
Office of Planning, Environment, and Realty
1200 New Jersey Avenue, SE
Washington, DC 20590

July 2021

FHWA-HEP-21-032

U.S.Department of Transportation
Federal Highway Administration