How-to: Quantify Uncertainty in Travel Forecasts

APRIL 2018

U.S. Department of Transportation
Federal Highway Administration

TMIP
Better Methods. Better Outcomes.

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Quality Assurance Statement

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. The FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HEP-20-021		

4. Title and Subtitle	5. Report Date
How-to: Quantify Uncertainty in Travel Forecasts	April 2018
	6. Performing Organization Code

7. Authors	8. Performing Organization Report No.
Jason Chen, PhD, Vince Bernardin, PhD, Steve Tuttle, and Steven Trevino	

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
RSG 55 Railroad Row White River Junction, VT 05001	11. Contract or Grant No. DTFH61-12-D-00013, T-12002

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
United States Department of Transportation Federal Highway Administration 1200 New Jersey Ave. SE Washington, DC 20590	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

Travel forecasts inherently embody some level of uncertainty as a result of uncertainties around many of the key inputs, simplifications, and statistical methods that are used to derive those forecasts. This uncertainty translates into risks that decisions based on the forecasts will be misguided – either failing to meet objectives or simply not performing as well as some other alternatives that could have been selected. From an agency perspective these risks can lead to lawsuits, loss of credibility, unwarranted expenditures, or suboptimal allocation of funding. Formal risk analysis includes several dimensions including identifying sources of risk, quantifying their likelihoods of occurrence and estimating the consequences of those occurrences. In the case of travel forecasting, quantifying uncertainty in travel forecasts and its effects on key performance measures is a non-trivial task given the complexity of the factors which can impact travel.

This report provides details on how uncertainty in travel forecasts and related performance measures can be quantified. Formal methods for quantifying risk or uncertainty profiles in key performance measures have been developed and are increasingly common in the context of investment grade traffic and toll revenue studies. This How-to guide will illustrate how these methods can be applied to other performance measures such as system VMT, delay, transit ridership, walk and bicycle mode shares, and greenhouse gases and other emissions. Three different risk analysis approaches are explained and illustrated using an activity-based model for Chattanooga, TN, and a four-step model for Toledo, OH: 1) traditional sensitivity analyses with simple risk profiling (similar to FTA guidance); 2) risk profiling based on univariate sensitivity analysis with Monte Carlo simulation and 3) more robust risk profiling using multivariate response surface methods and Monte Carlo simulation.

17. Key Words	18. Distribution Statement
travel; forecast; model; uncertainty; sensitivity analysis; multivariate response surface; 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	85	N/A

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)

How-to: Quantify Uncertainty in Travel Forecasts

April 2018

Prepared for:

Federal Highway Administration

Table of Contents

1.0 Introduction 1
1.1 Reasons for Quantifying Uncertainty 1
1.2 Methods for Quantifying Uncertainty 3
2.0 Case Studies 6
2.1 Locations 6
2.2 Performance Measures of Interest 10
2.3 Sources of Uncertainty 11
3.0 Univariate Sensitivity Analyses 27
3.1 Telecommuting 27
3.2 Parking Costs 29
3.3 Fuel Prices 32
3.4 Transit Fares 34
4.0 Response Surface Simulation 37
4.1 Experimental Design 37
4.2 Reduced Form Model Estimation 40
4.3 Monte Carlo Simulation 45
4.4 Results 58
5.0 Discussion 72
5.1 Discussion of Example Results 72
5.2 Conclusions 72

List of Figures

Figure 1. Historic accuracy of toll road forecasts. 2
Figure 2. Historic accuracy of ADT forecasts. 3
Figure 3. Toledo model region. 7
Figure 4. Toledo model flow. 8
Figure 5. Chattanooga model region. 9
Figure 6. Chattanooga model flow. 10
Figure 7. Telecommuting growth rate distribution. 16
Figure 8. Equation. Telecommuting share distribution generation. 16
Figure 9. 2045 telecommuting distribution. 17
Figure 10. Inflation rate (parking cost growth rate) distribution. 18
Figure 11. Equation. Parking cost distribution generation. 18
Figure 12. 2045 parking cost distribution. 19
Figure 13. Annual fuel cost growth rate distribution. 21
Figure 14. Equation. Fuel cost distribution generation. 21
Figure 15. Histogram plot of 2045 fuel cost (cents/mile) random draws. 22
Figure 16. 2045 fuel cost (cents/mile) log-normal distribution. 23
Figure 17. Equation. Mode m share distribution generation. 24
Figure 18. 2045 transit share distribution. 24
Figure 19. 2045 bike share distribution. 25
Figure 20. 2045 walk share distribution. 25
Figure 21. 2045 non-auto share distribution. 26
Figure 22. 2045 auto cost distribution (Toledo and Chattanooga). 46
Figure 23. 2045 telecommuting share distribution (Toledo and Chattanooga). 47
Figure 24. 2045 parking cost distribution (Toledo and Chattanooga). 47
Figure 25. 2045 non-auto mode share distribution (Toledo and Chattanooga). 48
Figure 26. 2045 transit fare distribution (Toledo and Chattanooga). 48
Figure 27. Toledo model 2045 urban core population growth rate distribution (2045 vs 2010). 49
Figure 28. Chattanooga model 2045 urban core population growth rate distribution (2045 vs 2010). 49
Figure 29. Toledo model 2045 urban core employment growth rate distribution (2045 vs 2010). 50
Figure 30. Chattanooga model 2045 urban core employment growth rate distribution (2045 vs 2010). 50
Figure 31. Toledo model 2045 halo population growth rate distribution (2045 vs 2010). 51
Figure 32. Chattanooga model 2045 rest of area population growth rate distribution (2045 vs 2010). 51
Figure 33. Toledo model 2045 halo employment growth rate distribution (2045 vs 2010). 52
Figure 34. Chattanooga model 2045 halo employment growth rate distribution (2045 vs 2010). 52
Figure 35. 2045 percentage of Boom City (Toledo and Chattanooga). 53
Figure 36. Toledo model squared error of prediction for VMT. 54
Figure 37. Chattanooga model squared error of prediction for VMT. 54
Figure 38. Toledo model squared error of prediction for delay. 55
Figure 39. Chattanooga model squared error of prediction for delay. 55
Figure 40. Toledo model squared error of prediction for transit ridership. 56
Figure 41. Chattanooga model squared error of prediction for transit ridership. 56
Figure 42. Toledo model squared error of prediction for auto emissions. 57
Figure 43. Chattanooga model squared error of prediction for auto emissions. 57

Figure 44. Cumulative probability distribution of 2045 VMT for Toledo. 59
Figure 45. Cumulative probability distribution of 2045 VMT for Chattanooga. 59
Figure 46. VMT sensitivity for Toledo. 60
Figure 47. VMT sensitivity for Chattanooga..61
Figure 48. Cumulative probability distribution of 2045 delay for Toledo. 62
Figure 49. Cumulative probability distribution of 2045 delay for Chattanooga. 62
Figure 50. Delay sensitivity for Toledo. 63
Figure 51. Delay sensitivity for Chattanooga. 64
Figure 52. Cumulative probability distribution of 2045 transit ridership for Toledo. 65
Figure 53. Cumulative probability distribution of 2045 transit ridership for Chattanooga (original). 65
Figure 54. Cumulative probability distribution of 2045 transit ridership for Chattanooga (revised). 66
Figure 55. Transit ridership sensitivity for Toledo. 67
Figure 56. Transit ridership sensitivity for Chattanooga. 68
Figure 57. Cumulative probability distribution of 2045 car emissions for Toledo. 69
Figure 58. Cumulative probability distribution of 2045 car emissions for Chattanooga. 69
Figure 59. Car emissions sensitivity for Toledo. 70
Figure 60. Car emissions sensitivity for Chattanooga. 71

List of Tables

Table 1. Contribution to forecast uncertainty versus input uncertainty and model sensitivity. 4
Table 2. Toledo 2045 default land use. 12
Table 3. Toledo low growth land use. 12
Table 4. Toledo high growth land use. 13
Table 5. Toledo medium growth land use..13
Table 6. Toledo Boom City scenario. 13
Table 7. Chattanooga 2045 agency land use. 13
Table 8. Chattanooga low growth land use. 14
Table 9. Chattanooga high growth land use. 14
Table 10. Chattanooga medium growth land use. 14
Table 11. Chattanooga boom city scenario. 14
Table 12. Telecommuting trends by year from ACS. 15
Table 13. Telecommuting trends by year from SIPP. 15
Table 14. Telecommuting trends by year from the decennial census. 15
Table 15. Fuel cost annual growth rate. 20
Table 16. Descriptive analysis of ln(2045 fuel cost). 22
Table 17. Non-automobile commuting shares. 23
Table 18. Descriptive analysis of 2045 non-auto preference random draws. 26
Table 19. Change of trip generation in telecommuting scenario for trip-based model. 27
Table 20. Toledo model sensitivity to telecommuting. 28
Table 21. Change of commuting trips in telecommuting scenario for activity-based model. 28
Table 22. Change of tour generation in telecommuting scenario for activity-based model. 29
Table 23. Chattanooga model sensitivity to telecommuting. 29
Table 24. Transit share by purpose in parking cost increase scenario. 30
Table 25. Toledo model sensitivity to parking costs. 30
Table 26. Change of total tours by purpose. 31
Table 27. Change of tour mode. 31

Table 28. Chattanooga model sensitivity to parking costs. 32
Table 29. Change of walk to transit and drive to transit in fuel price increase scenario. 32
Table 30. Toledo model sensitivity to fuel prices. 33
Table 31. Change in tours by purpose in fuel price increase scenario. 33
Table 32. Change of tour mode share in fuel price increase scenario. 34
Table 33. Chattanooga model sensitivity to fuel prices. 34
Table 34. Percent change in Toledo transit shares from fare adjustments. 35
Table 35. Toledo model sensitivity to transit fares. 35
Table 36. Percent change in Chattanooga transit share from fare adjustments. 36
Table 37. Absolute increase in Chattanooga transit trips from fare adjustments. 36
Table 38. Chattanooga model sensitivity to transit fares. 36
Table 39. Levels for each factor. 37
Table 40. Experiment orthogonal fractional factorial design. 39
Table 41. Bivariate correlations. 40
Table 42. Explanatory variables in regression equations for reduced form model estimation. 40
Table 43. Regression results for VMT in Toledo (trip-based) model. 42
Table 44. Regression results for VMT in Chattanooga (activity-based) model. 42
Table 45. Regression results for delay in Toledo (trip based) model. 43
Table 46. Regression results for delay in Chattanooga (activity based) model. 43
Table 47. Regression results for transit ridership in Toledo (trip based) model. 44
Table 48. Regression results for transit ridership in Chattanooga (activity based) model. 44
Table 49. Regression results for auto emissions in Toledo (trip based) model. 45
Table 50. Regression results for auto emissions in Chattanooga (activity based) model. 45

1.0 Introduction

Travel forecasts inherently embody some level of uncertainty as a result of uncertainties around many of the key inputs, simplifications, and statistical methods that are used to derive those forecasts. This uncertainty translates into risks that decisions based on the forecasts will be misguided – either failing to meet objectives or simply not performing as well as some other alternatives that could have been selected. From an agency perspective, these risks can lead to lawsuits, loss of credibility, unwarranted expenditures, or suboptimal allocation of funding. Formal uncertainty analysis must be performed sequentially by identifying sources of uncertainty, quantifying their likelihoods of occurrence and estimating the impact of those occurrences. In the case of travel forecasting, quantifying uncertainty in travel forecasts and its effects on key performance measures is a non-trivial task given the number and complexity of the factors which can impact travel.

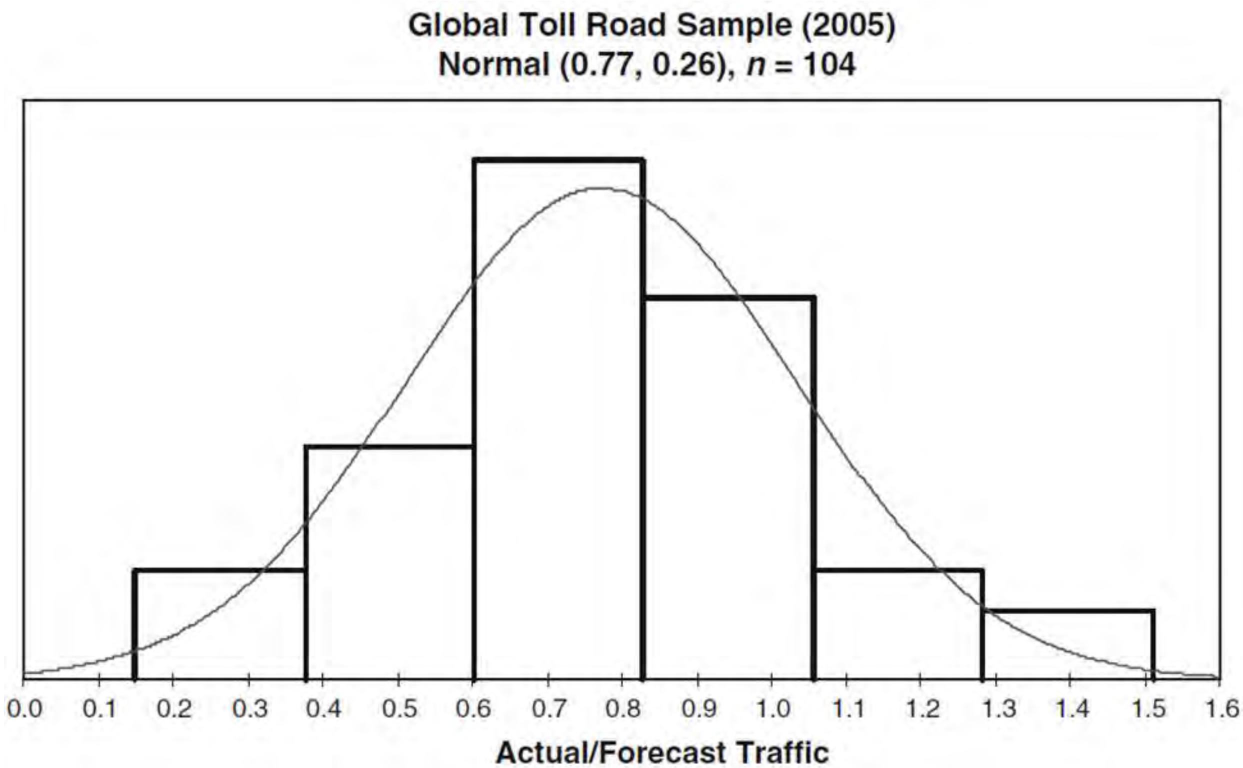
This installment of TMIP's How-To series builds on the TMIP white paper on *Managing Uncertainty and Risk in Travel Forecasting* by providing details on how uncertainty in travel forecasts and related performance measures can be quantified. Formal methods for quantifying risk or uncertainty profiles in key performance measures have been developed and are increasingly common in the context of investment grade traffic and toll revenue studies. This How-to guide illustrates how these methods can be applied to other performance measures such as system VMT, delay, and greenhouse gases or other emissions as representative of performance measures commonly used for metropolitan transportation planning.

1.1 Reasons for Quantifying Uncertainty

While central value or point forecasts of travel demand may be suitable for some applications, such as when filtering a long list of alternative project designs, there are many instances where modelers, policy makers, or outside stakeholders would greatly benefit from a careful examination of forecast uncertainty. Quantifying uncertainty could enhance model forecasts in many ways, including by:

- Providing comprehensive results
- Evaluating the risk to key stakeholders
- Describing how changes to key assumption could affect the outcome
- Accounting for highly uncertain assumptions

Historical evidence shows that actual travel demand can deviate considerably from point forecasts, often due to overly optimistic model assumptions. In a seminal study from the 1990s, Pickrell (1992) observed significant deviation between transit forecasts and actual ridership, with the forecasts generally exceeding actual ridership. Over a decade later, Flyvbjerg et al. (2003) documented analogous forecast deviations for a wide set of transit and toll road projects. Bain (2009) reviewed additional toll road studies and observed approximately normal forecast errors along with a systematic optimism bias.



Source: Bain, R. (2009) Error and optimism bias in toll road traffic forecasts. *Transportation*, Vol. 36, No. 5, pp. 469-482.

Figure 1. Historic accuracy of toll road forecasts.

Even if optimism biases are not uncommon in forecasting, producing a range of outcomes may compel modelers and policy makers to carefully review the key model assumptions and consider a wider and more realistic set of possibilities.

Beyond generally enhancing forecasts, there are specific situations where quantifying uncertainty is crucial. One such case is for projecting toll road or fare box revenue. Public agencies need to balance expected costs against expected revenue and government subsidies. Private parties may be reluctant to take either a debt or equity stake in a transportation infrastructure investment without having a clear understanding of the risks. Private debt holders may be especially sensitive to downside risk and care greatly about revenue forecasts for the worst 25% of cases. Private equity holders would also care about the potential for a high return on investment.

Understanding forecast uncertainty may also be crucial for policy makers who need to set appropriate targets for performance measures. If government funding depends on meeting those targets, then policy makers may want to define conservative but not overly pessimistic goals. When less is at stake, policy makers can set ambitious but attainable targets.

There are also situations where a key model assumption is so uncertain that it would not be credible to forecast, or at least acknowledge, anything but a range of outcomes. A topical example is modeling the use of autonomous vehicles. Among myriad uncertainties, future market penetration rates for “shared” and “private” autonomous are not well established. The analyst may have to bound the forecasts by testing extreme cases of the uncertain assumption.

1.1.1 Qualitative Treatment of Uncertainty

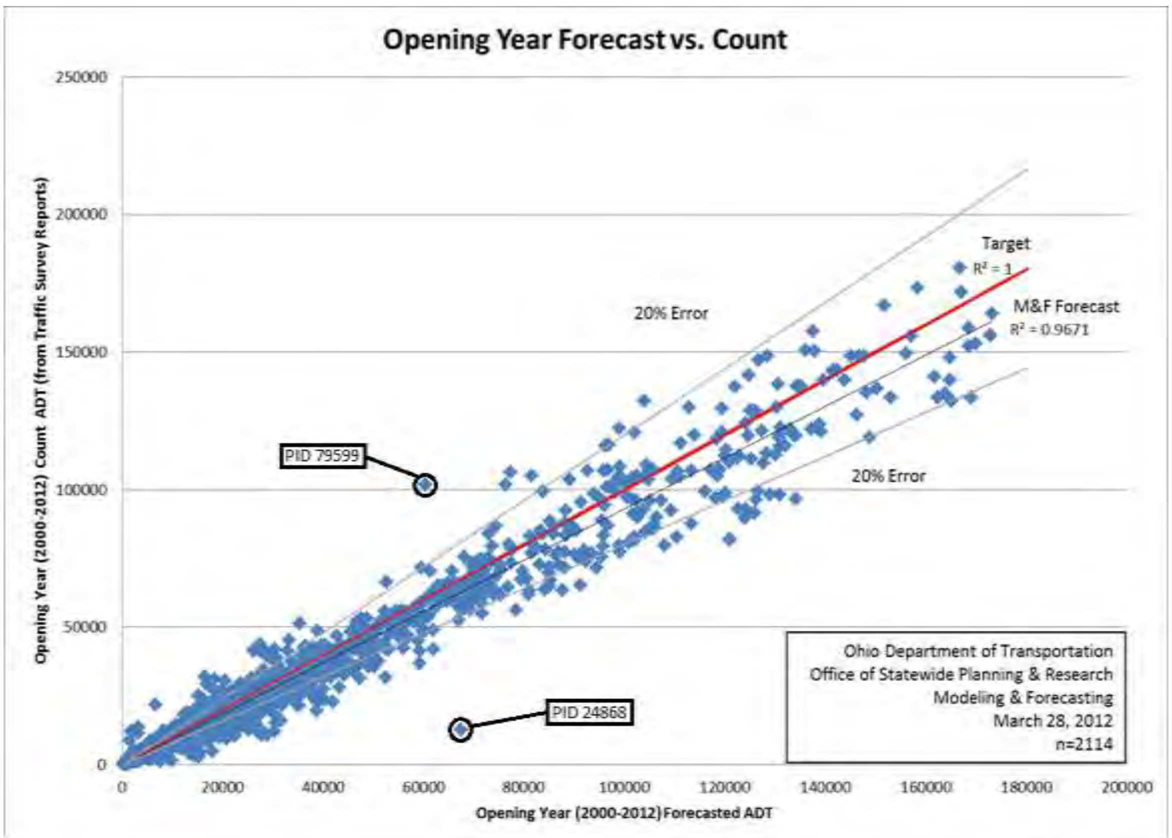
While there are clear benefits to quantifying uncertainty, there are also qualitative approaches to avoid conveying false precision or a false sense of confidence in point forecasts. One common approach is to round the forecasts to one or two significant figures. Another approach is for the analyst, based on professional judgment, to indicate whether the forecasting error is expected to be relatively "small" or "large" or otherwise verbally communicate an approximate level of confidence.

1.2 Methods for Quantifying Uncertainty

Forecast uncertainty can be quantified in historical terms, without explicitly considering the details of each new forecast, or it can be quantified analytically based on the unique set of inputs and methods used in the forecast. Both approaches have advantages and disadvantages, and the preferred approach may depend on whether the new study is unusual or routine and how far the forecast extends into the future.

1.2.1 Historical / Retrospective

Some agencies, including the Ohio Department of Transportation (ODOT), have made a recent effort to record forecast deviations. So far, ODOT has focused on comparing opening year forecasts, but they are progressively building clear, quantitative evidence of forecast uncertainty for routine studies.



Source: Improving Project Level Traffic Forecasts by Attacking the Problem from all Sides, presented by Greg Giaimo and Mark Byram at the 2013 TRB Transportation Planning Applications Conference.

Figure 2. Historic accuracy of ADT forecasts in Ohio.

Retrospective analysis has also become more common in toll road forecasting, with analyses by bond rating houses such as Standard and Poor's and consultants such as Robert Bain (see Figure 1). These analyses provide a valuable complement to analytic methods of estimating uncertainty and have the benefit of being grounded in relevant data. However, they have only limited use in understanding the uncertainty of forecasts for particular projects or for particular regions with characteristics that may deviate from the norm or the relevant observable past. For example, retrospective analyses may underestimate uncertainty related to new factors such as autonomous vehicles or in regions experiencing dynamic growth (unless the historic observations come from the same or a similar region).

1.2.2 Analytic

Analytic methods can be used to evaluate how model inputs contribute to forecast uncertainty. If a probability function can be assumed for each input, then advanced analytic methods can be used to formally estimate the distribution of outcomes.

Univariate

Univariate "sensitivity" analyses can provide a partial quantitative description of how the forecast depends on individual inputs. If the forecast is very sensitive to a relatively uncertain input, then the overall uncertainty is likely to be high.

Table 1. Contribution to forecast uncertainty versus input uncertainty and model sensitivity.

	Low Model Sensitivity to Input	**High Model Sensitivity to Input**
Low Variance Input	LOW Contribution	LOW Contribution
High Variance Input	LOW Contribution	HIGH Contribution

Univariate "sensitivity tests" are not uncommon. In univariate testing, the analyst measures how much the results change in response to either a moderate or substantial change to one of the inputs. This response allows the analyst to estimate a forecast elasticity for each input. The analyst can review elasticities of key inputs to evaluate which variables most affect the results and by roughly how much.

Univariate analyses cannot capture complex interactions among several inputs and are far less useful for evaluating how the results vary with simultaneous changes to multiple inputs which may be likely in many forecasting situations.

Decomposition

Relative to base conditions, forecast scenarios commonly include major changes to both transportation supply and demand. These changes are generally modeled together in a single scenario, but they can also be modeled incrementally, starting with near certain changes and advancing to (somewhat) less certain changes. This decomposition technique approximates the effect of individual changes to supply and demand, and it also provides several forecasts, ranging from very pessimistic to optimistic outcomes.

Starting from base conditions, the minimum "build" network is typically added first since this service is virtually guaranteed to exist once the project is approved. More conservative land use changes can then be added, followed by more optimistic growth or higher levels of service. This approach of building up a series of forecasts starting with a hypothetical scenario as if the facility/service were open today and layering on growth, etc., has become a common approach in transit forecasting, and seen some use in other contexts as well.

Scenario Testing

"Scenario testing" typically focuses on modeling distinct optimistic, middling, and pessimistic visions of the future. At its best, this method is a relatively quick and direct way to produce a meaningful range of forecasts.

The distinct scenarios generally include changes to several inputs and may loosely indicate multivariate effects, however, the scenarios are often too few or inefficiently designed for rigorous analysis of input-output relationships. Further, the likelihood of each scenarios is often unknown, limiting their ability to quantify risks to key stakeholders.

Response Surface Simulation

Response surface simulation can be used to carefully evaluate how forecasts depend on their inputs, describing how the results depend both on individual inputs and on complex interactions among inputs. Further, if the set of possible input values and their likelihoods can be stated in probability functions, then this technique can produce a complete probability distribution of outcomes. This distribution of outcomes can help answer challenging policy questions, such as the likelihood of meeting performance or revenue targets.

Response surface simulation has only recently been introduced to travel forecasting by Adler et al. (2014), and has since been used for several major studies. Tillman and Adler (2015) applied the technique to quantify uncertainty in a traffic and revenue study of express lanes on I-4 near Orlando. Cambridge Systematics (2016) applied a version of the technique in a revenue risk analysis of the proposed California High-Speed Rail.

The technique generally includes three major steps:

- Identifying the key uncertainties or risks
- Specifying a probability distribution for the uncertain or risky inputs
- Estimating the probability distributions of model outputs

The set of key risks typically varies from study to study and might include topics such as demographic trends, fuel prices, or the use of autonomous vehicles. Some risks can be modeled directly by varying an obvious input or assumption. Other risks must be modeled indirectly by varying a proxy input or assumption.

Specifying probability distributions for the uncertain inputs may be the most challenging step. For some inputs, such as population growth, it may be possible to directly estimate the full probability distribution or, at least, the mean and variance of the distribution. For other inputs, such as market penetration rates for autonomous vehicle, the true probability distribution may be unobservable and the analyst may have to assume a distribution on the basis of very little information. A simple uniform or triangular distribution might be used when the true distribution is unknown. For simplicity, inputs are often assumed to be independently distributed, but the analyst may want to estimate joint distributions in important cases. There is some evidence, for example, that fuel prices and employment are negatively correlated (with a lag effect). Accounting for these types of interaction effects can be challenging but can further improve the robustness of an analysis.

Completing the final step may require many model runs and significant calendar time, but it is generally straightforward. The analyst should select an experimental design, such as fractional factorial design, that efficiently tests different combinations of inputs. After running each scenario, the analyst can estimate a regression model where model output is the dependent variable and model inputs, including interactions between inputs, are the explanatory variables. A Monte Carlo simulation can then be used to produce the distribution of outcomes by "drawing" many thousands (or more) combinations of inputs and feeding the inputs into the regression model.

2.0 Case Studies

Techniques for quantifying uncertainty are illustrated through two case studies. The case studies demonstrate methods for defining uncertain inputs and estimating relationships between those uncertain inputs and forecast variables.

This chapter begins with a brief overview of the two case studies, reviewing their travel supply and demand characteristics and the available forecasting tools. The key forecasting variables, or performance measures, are then discussed, followed by detailed qualitative and quantitative definitions of the uncertain inputs.

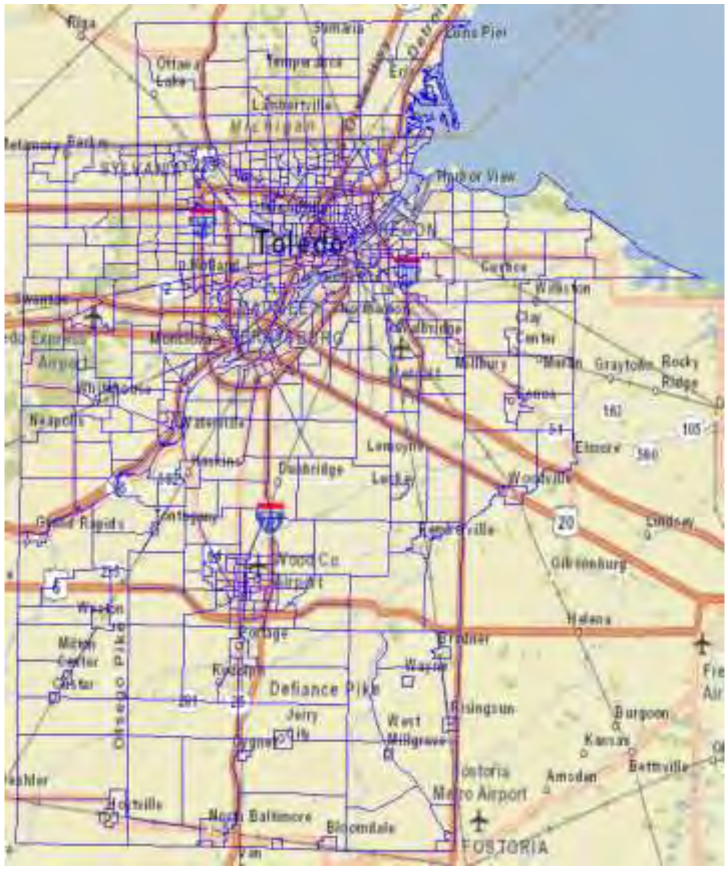
2.1 Locations

The study focuses on two mid-sized areas: Toledo, Ohio, and Chattanooga, Tennessee. The study areas have seen moderate changes in travel supply and demand in recent years, and neither area is expecting rapid growth or transformative infrastructure investment. The locations were chosen in part in the hopes that they were large and dynamic enough to have important uncertainties without being so large as to make the example analysis unduly complex or difficult.

Beyond basic differences in travel supply and demand, another distinction is that Toledo uses a traditional trip-based model and Chattanooga uses an activity-based model and each uses a different network modeling software, making the examples broadly applicable regardless of model design or software.

2.1.1 Toledo

The Toledo model includes Toledo, Ohio, and some neighboring communities within 10-30 miles (see Figure 3). The base model is calibrated to 2010 conditions, when there were 620,746 people and 344,197 jobs in the region. The model includes 239,319 households, 276,777 workers, and 145,885 children.

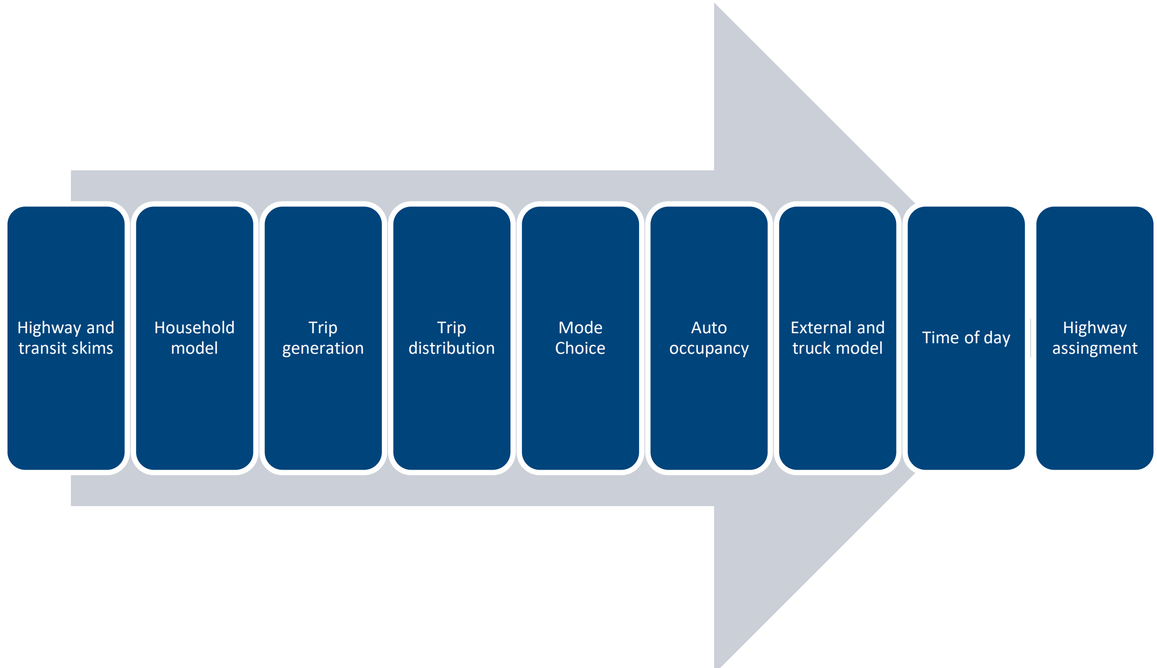


Source: © 2017 Google Maps®

Figure 3. Toledo model region.

Toledo uses a traditional trip-based model, implemented in Cube (see Source: Ohio Medium/Small MPO Model System.

Figure 4). The user runs a convergence check after each iteration and then manually initiates a feedback loop if one is required. Only two iterations are generally needed for convergence. The run times vary based on computer specifications but are generally very fast, with two iterations only requiring 12 to 15 minutes on a machine with four cores.



Source: Ohio Medium/Small MPO Model System.

Figure 4. Toledo model flow.

2.1.2 Chattanooga

The Chattanooga model covers Hamilton County in the state of Tennessee and Catoosa County and portions of Walker and Dade counties in northern Georgia (see Source: © 2017 Google Maps®

Figure 5). The base model is calibrated to 2014 conditions, when there were 445,876 people and 249,320 jobs in the region. The model includes 178,047 households, 183,079 workers, and 126,474 children.

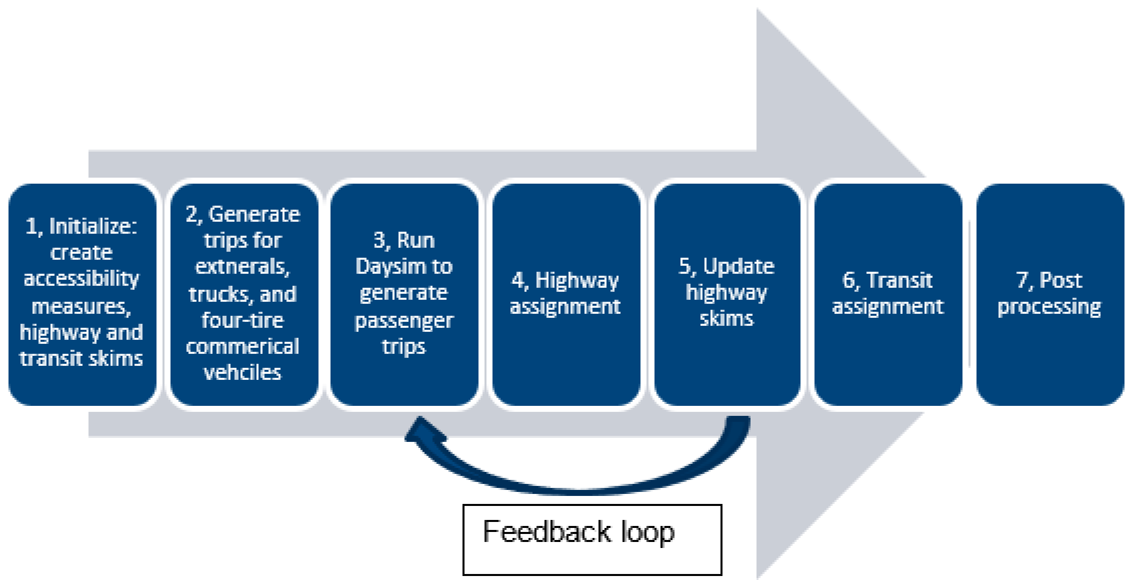
Source: © 2017 Google Maps®

Figure 5. Chattanooga model region.

Chattanooga uses an activity-based model, implemented in TransCAD 7, a GIS-based travel demand modeling software, and DaySim, an open-source software for activity-based modeling.

DaySim performs daily activity generation and scheduling, tour and trip level destination and mode choices for various travel purposes; however, the overall model flow is broadly similar to a traditional model (see Source: Chattanooga-Hamilton County Regional Planning Agency.

Figure 6). Truck and external demand are not microsimulated, but estimated by trip-based model components.



Source: Chattanooga-Hamilton County Regional Planning Agency.

Figure 6. Chattanooga model flow.

The Chattanooga model has an automatic feedback feature. The model checks for convergence based on changes in the highway skims, and it will generally converge in two iterations but quits if the maximum of four iteration is reached. The run time varies based on computer specifications, with a 32-core computer completing two iterations in about 2 hours and 30 minutes.

2.2 Performance Measures of Interest

Most model outputs depend on each uncertain input to at least a very minor degree, but it would be impractical to precisely quantifying how every output depends on every uncertain input. Thus, the analyst generally must focus on a set of key model outputs, or performance measures of interest.

The two case studies in this report focus on five performance measures:

- Vehicle Miles Traveled (VMT)
- Delay (hours)
- Transit ridership
- Walk and bike share (chapter 3 only)
- Auto emissions (chapter 4 only)

Although the case studies include several key performance measures for automobiles and trucks, the dominant modes of transportation, the case studies also include performance measures for transit and active travel modes which may have less certain future demand.

These case studies therefore differ from previous applications of the response surface simulation method in which it has generally been applied to understand the uncertainty in a single critical model output, such as toll road revenues or high-speed rail ridership. These pilot studies intentionally looked more broadly at a wider set of performance measures since agencies interested in applying the technique for general planning purposes would presumably also be interested in a number of performance measures.

2.3 Sources of Uncertainty

Identifying the key sources of uncertainty is a crucial step in both simple and more advanced approaches for quantifying uncertainty. Virtually all forecast year model inputs are uncertain to some degree, and the judgment of both local and technical experts is typically needed to identify which sources are important.

For advanced approaches, such as response surface simulation demonstrated in chapter 4, the project team must define a probability distribution for each input that states the likelihood the input will assume each possible value in its domain. For some inputs, the probability distribution, including the shape or mean and variance, can be estimated from historical data. For other inputs, including inputs that depend on emerging technologies or presumed laws and regulations, the distribution may have to be asserted based on professional judgment.

The case studies include six key sources of uncertainty:

- Land Use (chapter 4 only)
- Telecommuting
- Parking Costs
- Transit Fare
- Fuel Costs
- Generational Modal Preferences (chapter 4 only)

No list of travel forecast uncertainties can be truly exhaustive and include every source of uncertainty, however, some major uncertainties might also have to be excluded based on the scope and budget of the study. For example, this list includes several key determinates of future automobile use, but it does not include autonomous vehicles use since modeling this demand could greatly complicate the demonstration exercise. For the same reason, the study also does not include changes in freight and truck growth rate due to factors like globalization, modernization of production technologies, and trade agreements. Both of these may, however, be valuable factors in uncertainty to consider, and although they are not treated directly in the examples here, the methods illustrated here can be directly extended to deal with these issues as well.

Most of the uncertain inputs are assumed to be independently distributed. For example, the future transit fare is assumed not to be correlated with future automobile parking costs. The independence assumption obviates estimating joint probability distributions; however, inputs are often at least weakly correlated. For example, denser land uses may be correlated with higher parking costs. The response surface simulation method demonstrated later in this guide can account for these correlations, but only if they can be properly specified.

2.3.1 Land Use

The spatial distribution of households and employment is among the most important assumptions in travel forecasting. For rapidly transforming areas, these future land use assumptions may be the largest source of uncertainty.

The 2045 Toledo and Chattanooga land use assumptions were based on expert judgment from local agencies and observed population and employment growth from 2000 to 2010, according to the US Census. The Chattanooga land use scenarios were also informed in part through the use of a land use visioning tool.

The land use scenarios focused on changes to three mutually exclusive areas:

- “Urban Core”

- "Boom City," which is a peripheral site that experienced relatively large growth from 2000 to 2010, and while unlikely, could conceivably develop densely on a large scale
- "Halo Area," which includes all land outside of the "Urban Core" and the "Boom City," primarily suburban and rural areas

Distinct 2045 land use scenarios were developed for Toledo and Chattanooga. Each scenario generally had different assumptions for population and employment growth in the "Urban Core," "Halo Area," and "Boom City." The distinct scenarios served two purposes. First, the scenarios were used to develop land use inputs to measure each model's response to changes in the spatial distribution of population and employment. Section 4.1 defines which specific combinations of land use scenarios were used to estimate the reduced form equation for response surface simulation. Second, the distinct scenarios were used to develop continuous distributions of future land use through interpolation. The distinct scenarios include:

- **Default**: The 2045 scenario provided by local agencies
- **High Growth** (and no Boom City): High growth rate in all areas.
- **Low Growth** (and no Boom City): Low growth rate in all areas
- **Medium Growth** (and no Boom City): Medium growth rate in all areas
- **Boom City** (and either low, medium, or high growth): Very high population and employment growth in the boom city. The Urban and Halo growth may either be "low," "medium," or "high."

The default 2045 Toledo land use scenario originally provided is shown in Table 2. For this scenario, the total population decreases by about 5% while total employment increases by 33%, indicating a large imbalance in overall growth.

Table 2. Toledo 2045 default land use.

Area	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	196,322	181,782	-14,540	102,535	137,688	35,153
Boom City	11,781	13,126	1,345	6,653	9,252	2,599
Halo	412,643	397,284	-15,359	235,009	312,672	77,663
Total	620,746	592,192	-28,554	344,197	459,612	115,415

The "low" growth scenario, shown in Table 3, uses the population growth rates by area from the default scenario. The employment growth rate was assumed to be half the population growth rate, leading to -2.5% decrease in total employment.

Table 3. Toledo low growth land use.

Area	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	196,322	181,782	-14,540	102,535	98,738	-3,797
Boom City	11,781	13,126	1,345	6,653	7,033	380
Halo	412,643	397,284	-15,359	235,009	230,635	-4,374
Regional Total	620,746	592,192	-28,554	344,197	336,406	-7,791

The "high" growth scenario, shown in Table 4, uses the employment growth rates by area from the default scenario. Population by areas was assumed to grow at the same rate as employment.

Table 4. Toledo high growth land use.

Area	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	196,322	263,629	67,307	102,535	137,688	35,153
Boom City	11,781	16,383	4,602	6,653	9,252	2,599
Halo	412,643	549,008	136,365	235,009	312,672	77,663
Regional Total	620,746	829,020	208,274	344,197	459,612	115,415

The "medium" growth scenario was assumed to be the average of the high and low growth scenarios (see Table 5).

Table 5. Toledo medium growth land use.

Area	Base Population	Future Population	Population Growth	Base Employment	Future Employment	Employment Growth
Urban Core	196,322	222,705	26,383	102,535	118,213	15,678
Boom City	11,781	14,755	2,974	6,653	8,142	1,489
Halo	412,643	473,146	60,503	235,009	271,654	36,645
Regional Total	620,746	710,606	89,860	344,197	398,009	53,812

The Boom City population and employment growth was derived in a multiple-step process. Regional population and employment were controlled to the levels in the medium growth scenario. The urban core population and employment were set to the average of the low and medium growth scenarios. The "halo" population and employment were set to the average of the medium and low growth scenarios, weighting the medium growth scenario by 2/3 and the low growth by 1/3. Then, all remaining growth was assumed to occur in the Boom City (see Table 6). The Urban and Halo growth were assumed to be independent and may be either low, medium, or high if Boom City Exists.

Table 6. Toledo Boom City scenario.

Area	Base Population	Future Population	Population Growth	Base Employment	Future Employment	Employment Growth
Boom City	11,781	60,504	48,723	6,653	31,553	24,900

The land use data for the Chattanooga model is also developed based on a similar sort of approach. However, since the agency's default scenario was more balanced, the medium growth scenario resembles it more.

Table 7. Chattanooga 2045 agency land use.

	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	119,722	133,482	13,760	101,326	112,300	10,974
Boom City	5,385	9,090	3,705	799	1,625	826
Halo	320,692	423,731	103,039	120,053	175,221	55,168
Regional Total	445,799	566,303	120,504	222,178	289,146	66,968

Table 8. Chattanooga low growth land use.

	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	119,722	128,247	8,525	101,326	106,171	4,845
Boom City	5,385	7,945	2,560	799	1,530	731
Halo	320,692	378,890	58,198	120,053	159,721	39,668
Regional Total	445,799	515,082	69,283	222,178	267,422	45,244

Table 9. Chattanooga high growth land use.

	Base Population	Year 2045 Population	Population Growth	Base Employment	Year 2045 Employment	Employment Growth
Urban Core	119,722	136,510	16,788	101,326	116,978	15,652
Boom City	5,385	9,758	4,373	799	2,863	2,064
Halo	320,692	448,995	128,303	120,053	190,009	69,956
Regional Total	445,799	595,263	149,464	222,178	309,850	87,672

Table 10. Chattanooga medium growth land use.

	Base Population	Future Population	Population Growth	Base Employment	Future Employment	Employment Growth
Urban Core	119,722	133,482	13,760	101,326	109,066	7,740
Boom City	5,385	9,090	3,705	799	1,625	826
Halo	320,692	423,731	103,039	120,053	169,723	49,670
Regional Total	445,799	566,303	120,504	222,178	280,414	58,236

Table 11. Chattanooga boom city scenario.

Area	Base Population	Future Population	Population Growth	Base Employment	Future Employment	Employment Growth
Boom City	5,385	9,090	3,705	799	1,625	826

As discussed further in chapter 44.0 the distinct land use scenarios were also used to develop continuous land use distributions for the response surface simulation. Urban, halo, and boom city growth rates were assumed to be independently distributed. The urban and halo growth were assumed to be normally distributed with mean equal to the growth in the “medium” scenario. The high growth and low growth scenarios were assumed to be about two standard deviations above and below the mean respectively. The percentage of the boom city growth was assumed to be distributed between 0 and 1 by a triangular distribution with likeliest value 0.2.

2.3.2 Telecommuting

Telecommuting continues to grow in the United States. The percentage of people working exclusively at home has increased from 4.8% in 1997 to 6.6% in 2010, according to the Census. This trend is likely to continue with the advance of communication and information technologies, eliminating some motorized commuting trips, particularly during traditional peak period hours.

Telecommuting was assumed to grow from 2010 to 2045 at a rate generally consistent with historical trends. Table 12 shows annual growth rates in telecommuting based on samples of US residents from three sources: the American Community Survey (ACS), Survey of Income and Program Participation (SIPP), and the Decennial Census.

Table 12. Telecommuting trends by year from ACS.

Year	Total	Work at Home	%	Growth Rate
2000	127156	4160	3.27%	
2005	132383	4793	3.62%	2.13%
2006	137295	5301	3.86%	6.64%
2007	138282	5567	4.03%	4.27%
2008	142544	5794	4.06%	0.97%
2009	137312	5812	4.23%	4.13%
2010	135906	5815	4.28%	1.09%

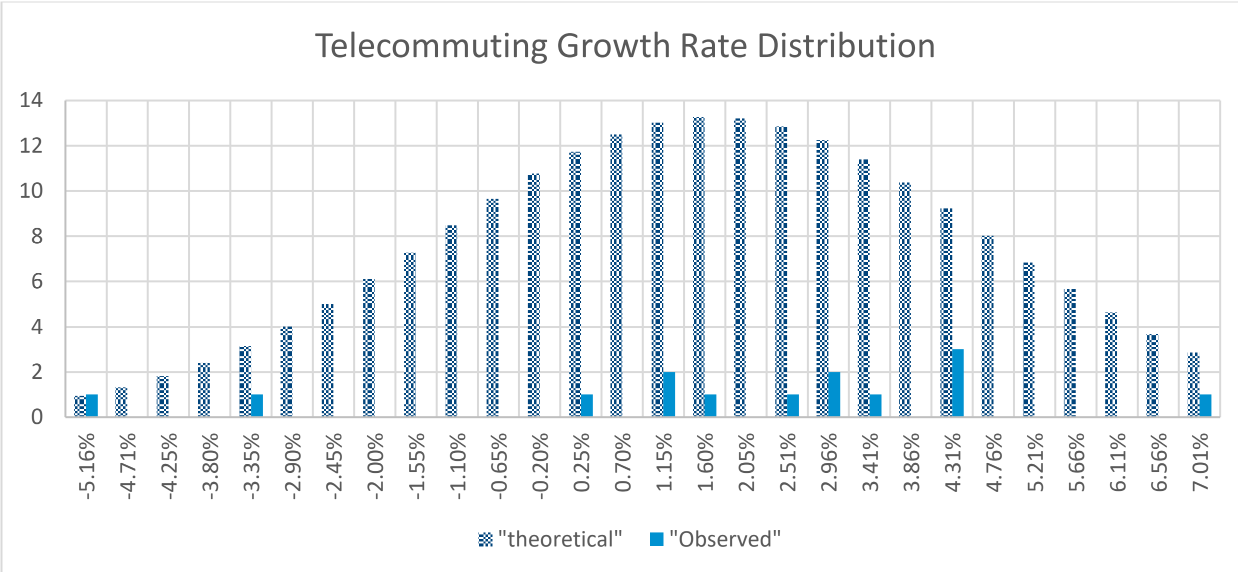
Table 13. Telecommuting trends by year from SIPP.

Year	Total	Work at Home	%	Growth Rate
1995	125925	8340	6.62%	
1997	132229	9241	6.99%	2.76%
1999	135955	9477	6.97%	-0.13%
2002	137930	10393	7.53%	2.70%
2005	144557	11313	7.83%	1.29%
2010	141646	13401	9.46%	4.18%

Table 14. Telecommuting trends by year from the decennial census.

Year	Total	Work at Home	%	Growth Rate
1960	64656	4663	7.21%	
1970	76852	2685	3.49%	-5.16%
1980	96617	2178	2.25%	-3.55%
1990	115070	3406	2.96%	3.13%

An annual telecommuting growth rate of 1.75% with a standard deviation of .03% was estimated from the data in Table 12. The theoretical growth rate was assumed to be normally distribution with this same mean and variance. The historical and theoretical growth rate distributions are shown in Figure 7.



Source: FHWA

Figure 7. Telecommuting growth rate distribution.

The distribution of percent telecommuting for 2045 was simulated by growing the 2010 percentage using annual growth rates randomly drawn from the theoretical distribution shown in Figure 7, per the formula:

$$Pct_{2045} = Pct_{2010} * (1 + G_{10-11}) * (1 + G_{11-12}) * (1 + G_{12-13}) * \dots (1 + G_{44-45})$$

Figure 8. Equation. Telecommuting share distribution generation.

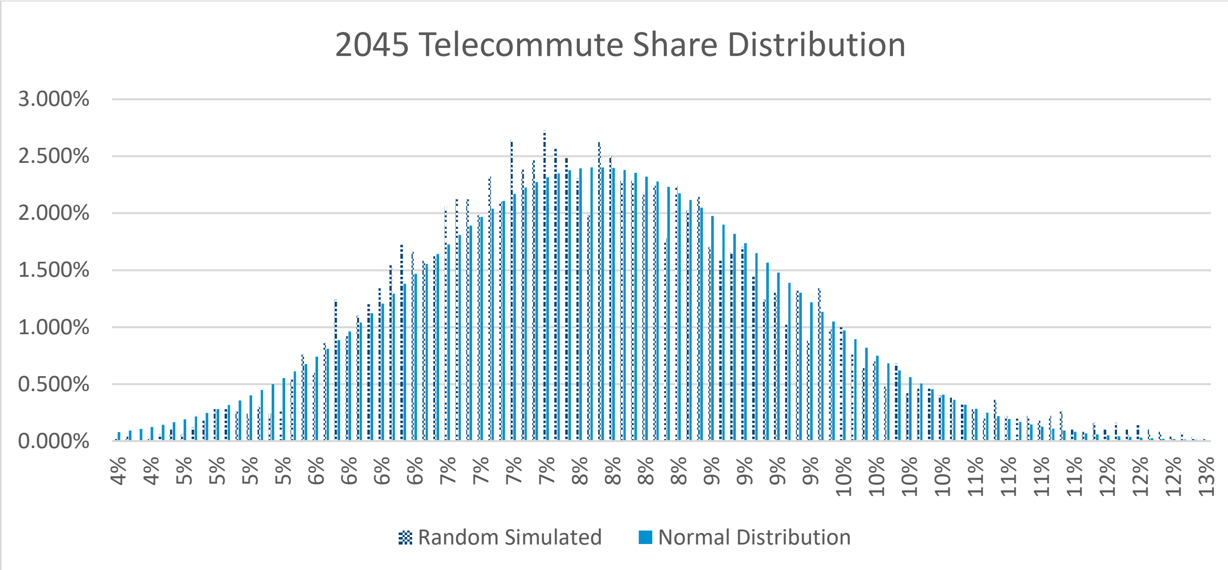
Where:

Pct_n: Year *n* telecommuting percentage

G_{n-n+1}: Random growth rate for year *n* to year *n+1*

The simulated distribution included 5000 values for the 2045 telecommuting percentage. It had a mean of 7.8% and standard deviation of 1.4%.

The final 2045 probability distribution was assumed to be normal (7.8%, 1.4%), with mean and variance based on the simulated distribution (see Figure 9).



Source: FHWA

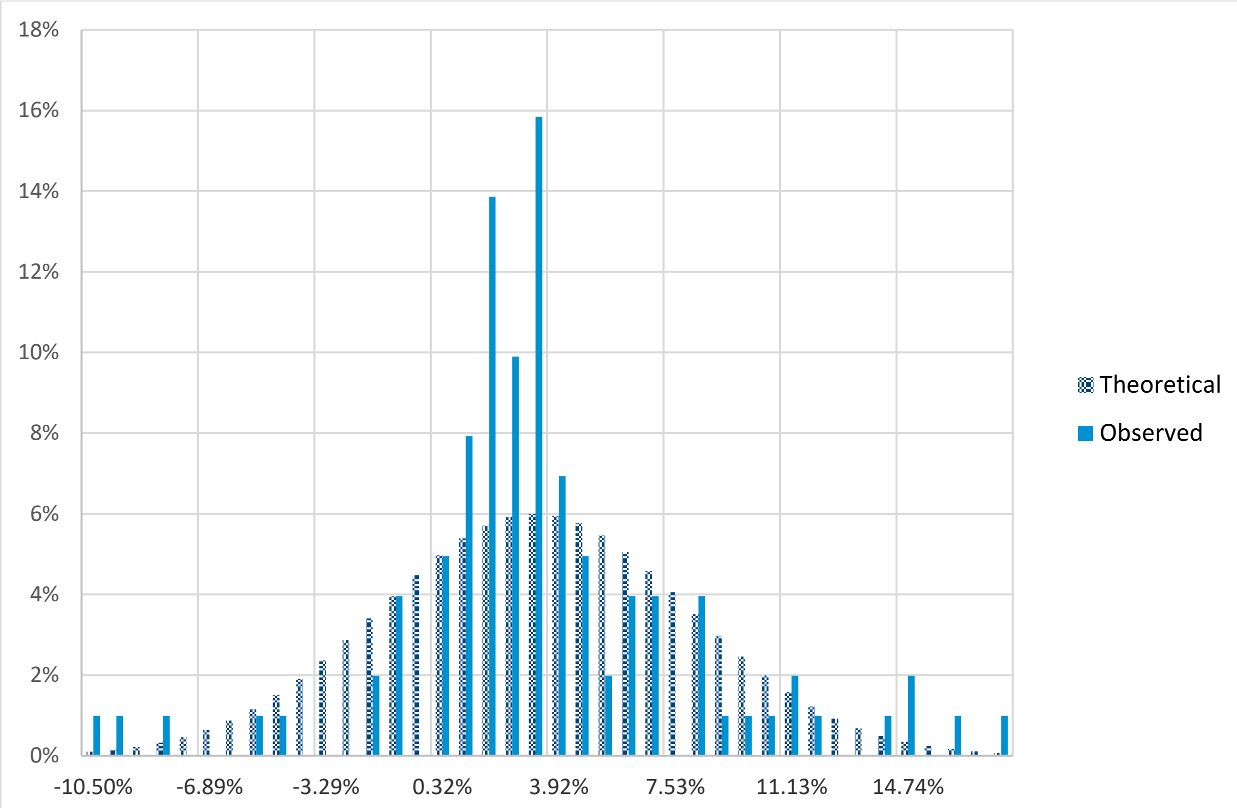
Figure 9. 2045 telecommuting distribution.

The Toledo model does not explicitly represent telecommuting. Instead, telecommuting can be represented indirectly by reducing the initial number of work trips. The Chattanooga model includes a work-from-home option in the work location choice model. Its telecommuting share can therefore be shifted by adjusting the choice model constants.

2.3.3 Parking Cost

Studies have shown that increasing parking costs will decease parking demand. Wilson (1992) found that fewer vehicles were used to drive to work when drivers had to pay for parking compared to when parking is free. Hess (2001) found that both drive-alone and shared-ride trips decrease while public transit increases if free parking is changed to just $1 parking. In general, parking costs could influence many important household decisions, including where to live, how to travel, and where to travel.

Parking cost was assumed to grow from 2010 to 2045 at a rate generally consistent with historical trends. However, historical parking costs were not immediately available for either study area, and the growth rate was instead assumed to be consistent with historical inflation. Inflation data from 1914 to 2015 are plotted in Figure 10. The observed distribution has an average annual growth of 3.3% and standard deviation of 4.8%. The theoretical growth rate was assumed to be normally distributed with the same mean and variance.



Source: FHWA

Figure 10. Inflation rate (parking cost growth rate) distribution.

A 2045 distribution of parking costs was simulated by growing the 2010 parking costs with annual inflation rates randomly drawn from the theoretical distribution in Figure 10, per the formula:

$$Cost_{2045} = Cost_{2010} * (1 + G_{10-11}) * (1 + G_{11-12}) * (1 + G_{12-13}) * \ldots (1 + G_{44-45})$$

Figure 11. Equation. Parking cost distribution generation.

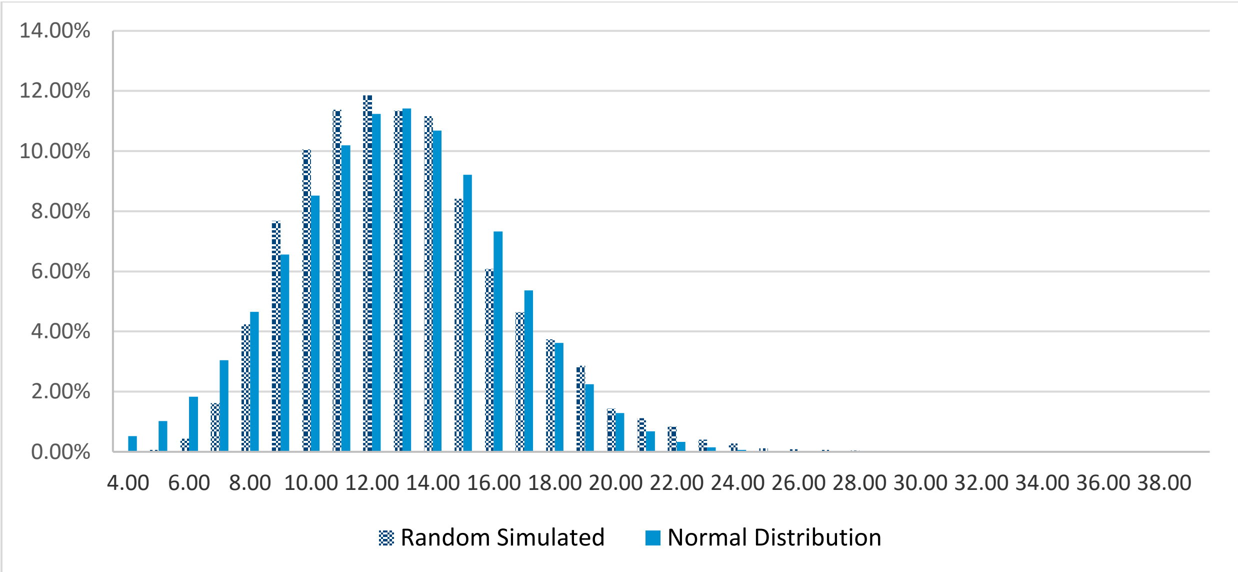
Where:

$Cost_n$: Year *n* parking cost

G_{n-n+1}: Inflation growth from year *n* to year *n+1*

The simulated distribution included 5000 values for the 2045 parking cost and had a mean of X and standard deviation of Y.

The final 2045 probability distribution was assumed to be normal (X, Y), with mean and variance based on the simulated distribution (see Figure 12).



Source: FHWA

Figure 12. 2045 parking cost distribution.

Parking costs can be adjusted in the Toledo model by changing the zonal attribute "AVG_PARK." Costs can be adjusted in the Chattanooga model by changing the hourly and daily parking prices in the microzone file.

2.3.4 Transit Fare

Higher transit fares generally lead to reduced demand. While some studies have found a higher elasticity for transit level of service than for transit fare, fare elasticities are generally significant, often falling between -0.35 and -0.65.

Transit fares are often driven by policy and may be weakly correlated with economic factors. Fares can remain constant for long periods and then be adjusted due to regional policy changes.

For both cases studies, the transit fare distribution was assumed to be discretely distributed with an 80% chance of being equal to the base value and a 20% chance of being half the base value.

Transit fares are represented in the Toledo model by adjusting boarding penalties. Fare are represented directly in the Chattanooga model by adjusting a field in the transit route system file.

2.3.5 Fuel Cost

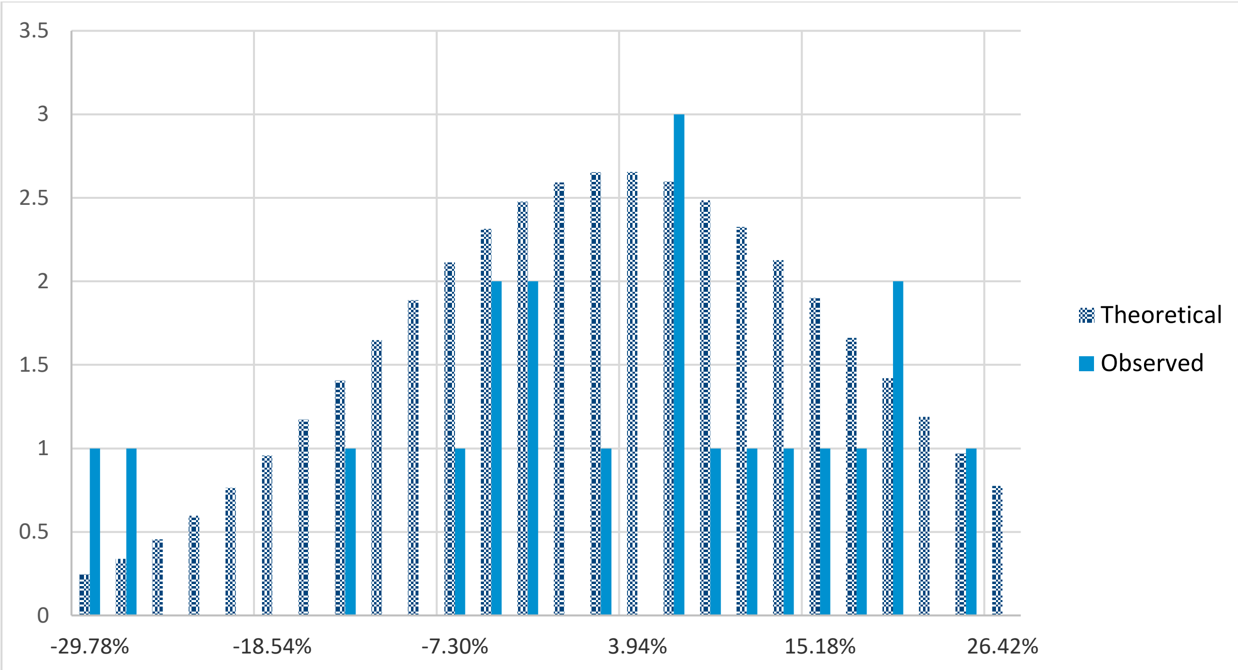
Fuel cost can influence long-term household decisions on purchasing vehicles and short-term decisions on where and how far to travel.

Fuel cost was assumed to grow from 2010 to 2045 at a rate generally consistent with historical trends. For this study, the annual rate of growth in fuel prices was estimated using 1994 to 2014 retail gasoline prices from the U.S. Energy Information Administration (see Table 15). After normalizing prices to 1994 levels, the annual growth rate was calculated to have a mean value of 2.89% and standard deviation of .15%. The theoretical growth rate distribution was assumed to be normal with the same mean and variance. Figure 13 compares the 1994 to 2014 growth rates to the theoretical distributions assumed for this study.

Table 15. Fuel cost annual growth rate.

Year	Retail Gasoline Prices[1]	Inflation Factor	1994 Prices	Annual growth rate
1994	$1.08	144.5	$1.08	
1995	$1.16	148.2	$1.13	4.7%
1996	$1.25	152.4	$1.18	4.6%
1997	$1.24	156.9	$1.15	-3.0%
1998	$1.07	160.5	$0.97	-15.8%
1999	$1.18	163.0	$1.04	8.0%
2000	$1.52	166.6	$1.32	26.7%
2001	$1.46	172.2	$1.23	-7.3%
2002	$1.39	177.1	$1.13	-7.7%
2003	$1.60	179.9	$1.29	13.9%
2004	$1.90	184.0	$1.49	15.6%
2005	$2.31	188.9	$1.77	18.9%
2006	$2.62	195.3	$1.94	9.4%
2007	$2.84	201.6	$2.04	5.2%
2008	$3.30	207.3	$2.30	12.9%
2009	$2.41	215.3	$1.61	-29.8%
2010	$2.84	214.5	$1.91	18.3%
2011	$3.58	218.1	$2.37	24.1%
2012	$3.68	224.9	$2.36	-0.2%
2013	$3.58	229.6	$2.25	-4.8%
2014	$3.44	233.0	$2.13	-5.3%
2015	$2.52	236.7	$1.54	-27.9%

[1] Dollars per Gallon, U.S. All Grades



Source: FHWA

Figure 13. Annual fuel cost growth rate distribution.

A 2045 distribution of fuel costs was simulated by growing the 2010 prices with growth rates randomly drawn from the theoretical distribution in Figure 7, per the formula:

$$Cost_{2045} = Cost_{2010} * (1 + G_{10-11}) * (1 + G_{11-12}) * (1 + G_{12-13}) * \dots (1 + G_{44-45})$$

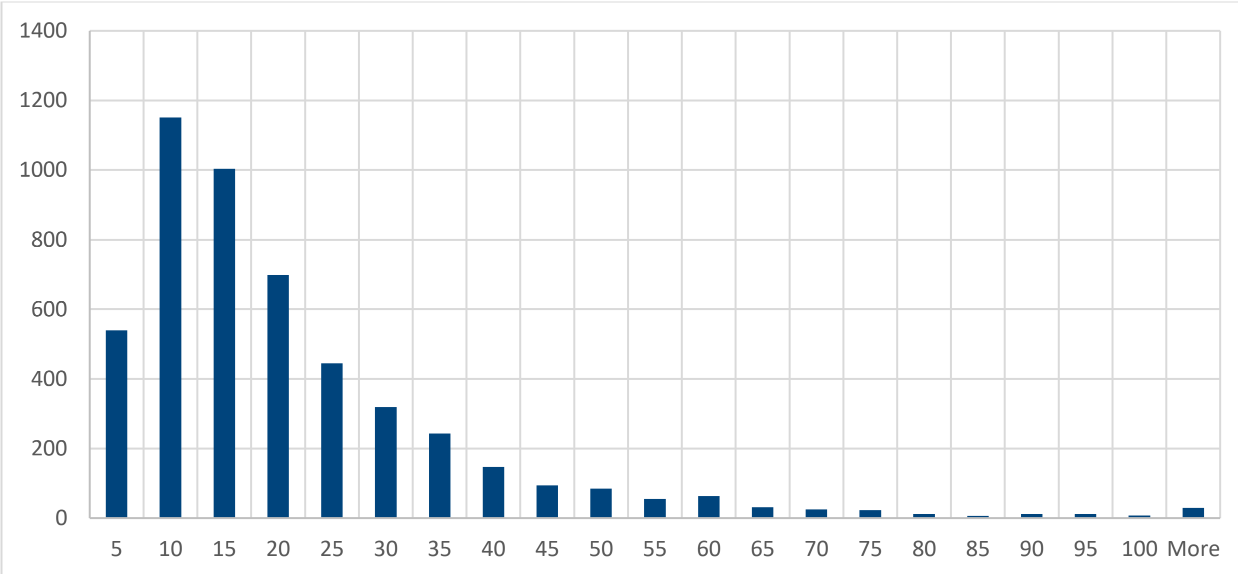
Figure 14. Equation. Fuel cost distribution generation.

Where:

$Cost_n$: Year *n* fuel cost

G_{n-n+1}: Cost growth from year *n* to year *n+1*

The simulated distribution included 5000 values for the 2045 fuel price (see Figure 15). The simulated distributed was postulated to be lognormal. Table 16 shows descriptive statistics for a lognormal transformation of the simulated distribution. These statistics generally support the lognormal assumption since the mean and median values are similar and the skewness is small.



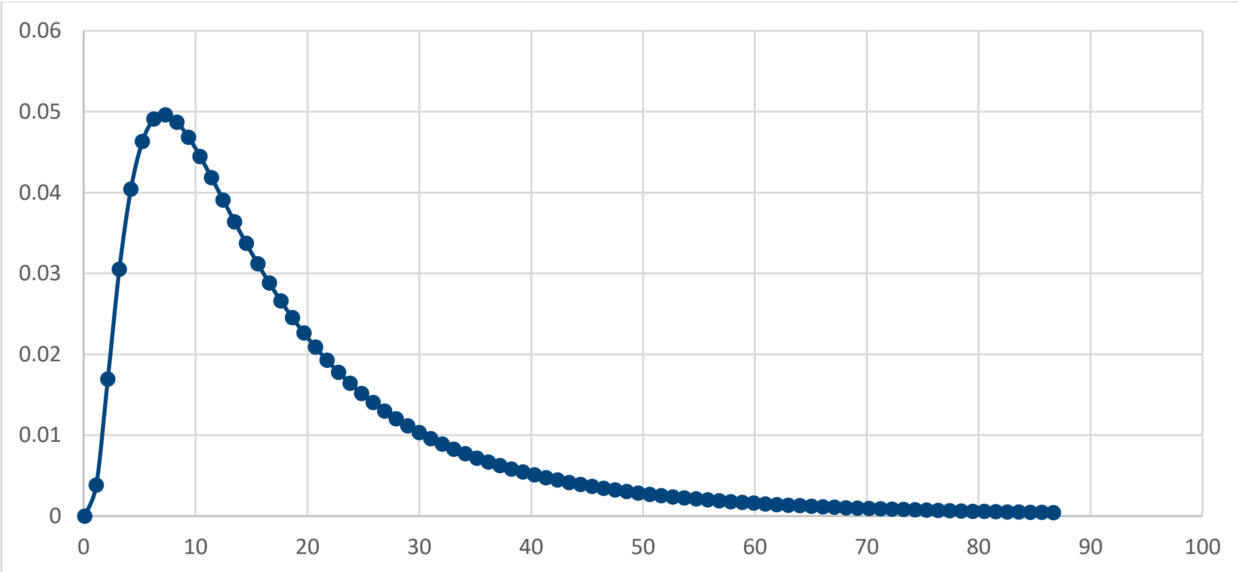
Source: FHWA

Figure 15. Histogram plot of 2045 fuel cost (cents/mile) random draws.

Table 16. Descriptive analysis of ln(2045 fuel cost).

Measure	Value
Mean	2.621641709
Standard Error	0.011489021
Median	2.633970452
Mode	#N/A
Standard Deviation	0.812396458
Sample Variance	0.659988006
Kurtosis	0.046430264
Skewness	-0.11118976
Range	5.948147631
Minimum	-0.611372582
Maximum	5.336775049
Sum	13108.20854
Count	5000

The final 2045 probability distribution of fuel cost was assumed to be lognormal (X, Y) with mean and variances based on the simulated distribution (see Figure 16).



Source: FHWA

Figure 16. 2045 fuel cost (cents/mile) log-normal distribution.

2.3.6 Generational Modal Preferences

Passenger travel in the United States is dominated by private automobile trips, however, the share of non-automobile trips has slowly increased in recent years. This trend may be due to recent investment in transit and walk/bike infrastructure or greater public awareness of the environmental and health benefits from less automobile use. For example, the US census indicates that the number of daily bike commutes has increased from 488,000 in 2000 to about 786,000 in 2008-2012.

Non-automobile shares were assumed to grow from 2010 to 2045 at a rate generally consistent with historical trends. Table 17 shows historical non-automobile commuting shares, according to the Bureau of Transportation Statistics. Average growth rates and standard deviations were estimated for the transit (0.80%, 2.68%), bike (4.72%, 5.11%), and walk (1.04%, 6.63%) shares. The growth rate distributions were assumed to be normal.

Table 17. Non-automobile commuting shares.

Year	Transit	Bike	Walk
2002	4.8	0.4	2.5
2003	4.7	0.4	2.3
2004	4.6	0.4	2.4
2005	4.7	0.4	2.5
2006	4.8	0.5	2.9
2007	4.9	0.5	2.8
2008	5.0	0.5	2.8
2009	5.0	0.6	2.9
2010	4.9	0.5	2.8

Year	Transit	Bike	Walk
2011	5.0	0.6	2.8
2012	5.0	0.6	2.8
2013	5.2	0.6	2.8
2014	5.2	0.6	2.7

The 2045 distributions of transit, walk, and bike mode shares were simulated by growing 2010 shares using growth rates randomly drawn from their respective distributions, per the formula:

$$Share_{m,2045} = Share_{m,2010} * \left(1 + G_{m,10-11}\right) * \left(1 + G_{m,11-12}\right) * \left(1 + G_{m,12-13}\right) * \dots \left(1 + G_{m,44-45}\right)$$

Figure 17. Equation. Mode m share distribution generation.

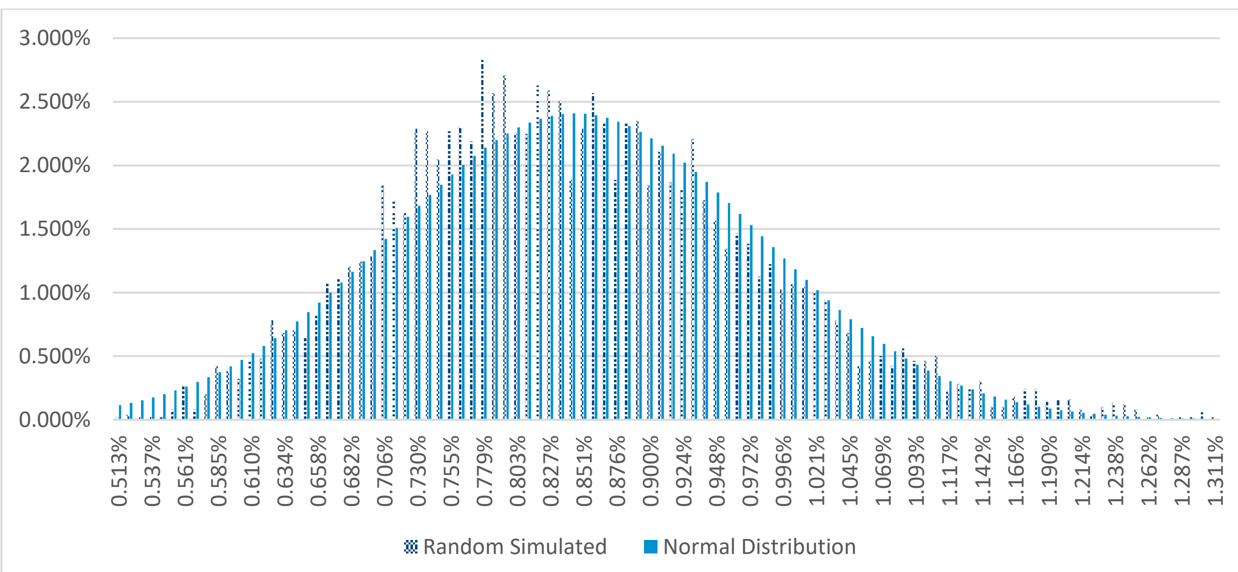
Where:

$Share_m$: Year *n* transit share for mode *m*

$T_{m,n-n+1}$: Growth rate from year *n* to year *n+1* for mode *m*

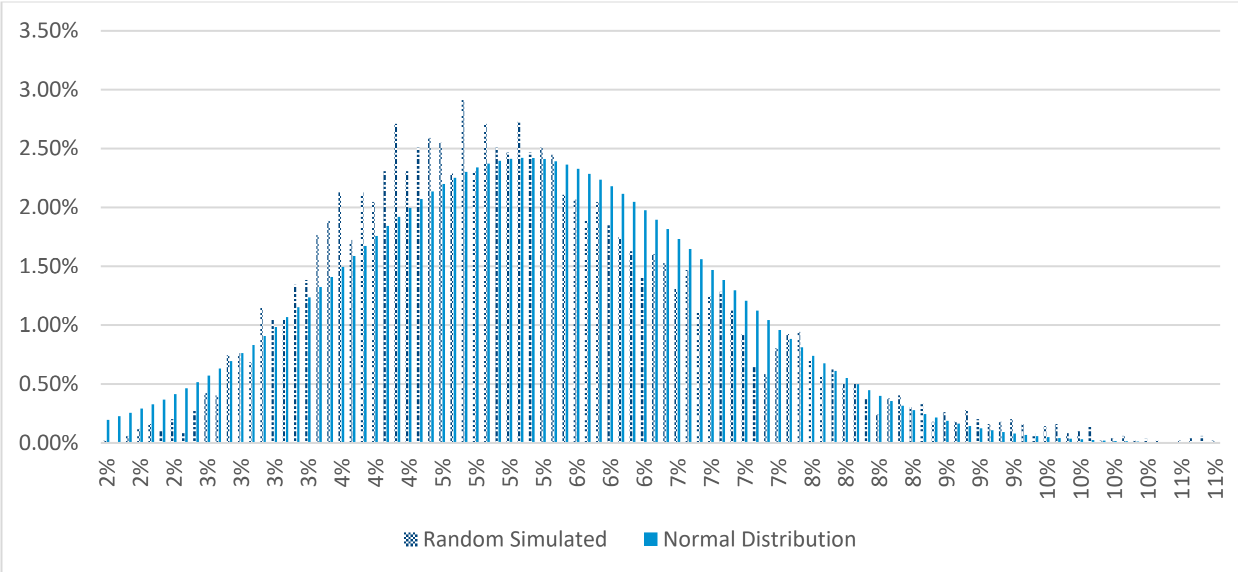
$m \in \{transit, walk, bike\}$

The simulated 2045 transit, bike, and walk distributions each included 5000 values. A mean and standard deviation was computed for the simulated transit (0.84%, 0.13%), bike (5.27%, 1.53%), and walk (4.08%, 1.66%) distributions. The theoretical distributions were assumed to be normal. Figure 18 to Figure 20 compares the simulated and theoretical distributions.



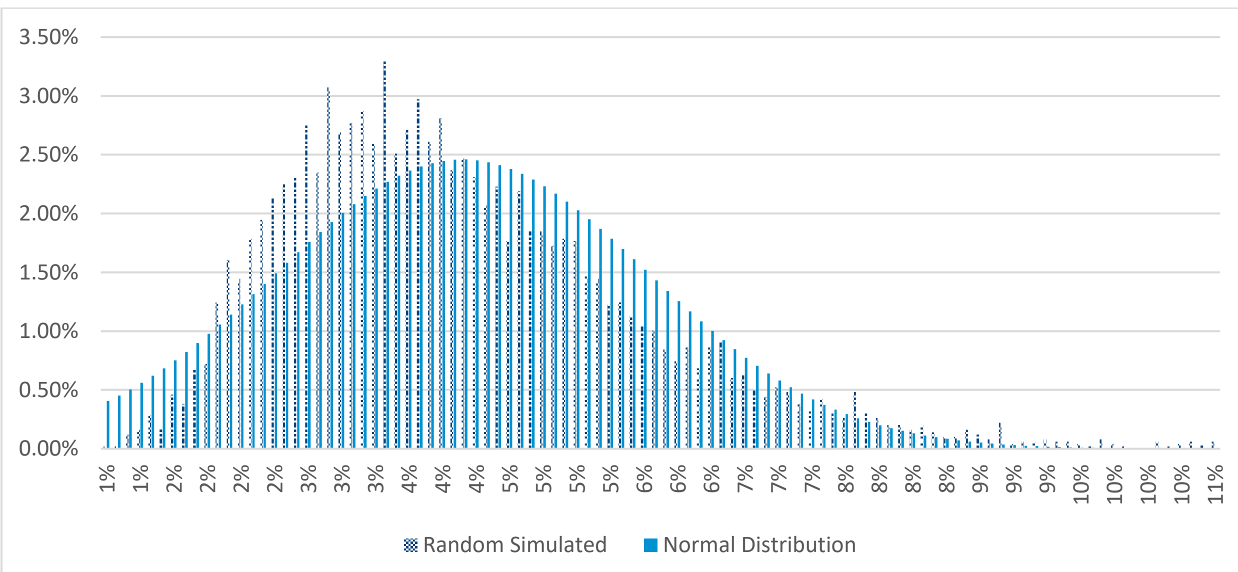
Source: FHWA

Figure 18. 2045 transit share distribution.



Source: FHWA

Figure 19. 2045 bike share distribution.



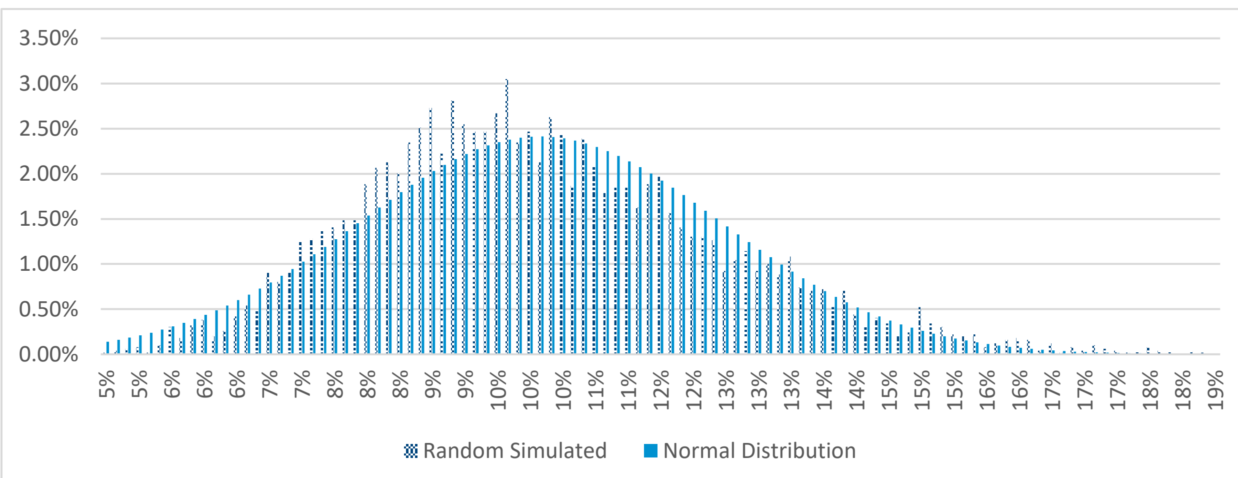
Source: FHWA

Figure 20. 2045 walk share distribution.

Table 18 shows descriptive statistics for the simulated 2045 transit, bike, walk distributions as well as statistics for the overall non-automobile distribution. Although the walk share distribution shows some positive skewness, the overall non-automobile distribution has small positive skewness, indicating that it may be appropriate to assume a normal shape. The theoretical distribution of 2045 non-automobile shares was assumed to be normal with a mean of 10.19% and standard deviation of 2.24% (Figure 21).

Table 18. Descriptive analysis of 2045 non-auto preference random draws.

Statistic	Transit	Bike	Walk	Non-Auto
Mean	8.44E-03	5.27E-02	4.08E-02	1.02E-01
Standard Error	1.90E-05	2.16E-04	2.35E-04	3.17E-04
Median	8.32E-03	5.09E-02	3.79E-02	9.93E-02
Standard Deviation	1.34E-03	1.53E-02	1.66E-02	2.24E-02
Sample Variance	1.80E-06	2.33E-04	2.76E-04	5.02E-04
Kurtosis	7.80E-01	9.14E-01	2.98E+00	8.34E-01
Skewness	5.88E-01	7.77E-01	1.26E+00	6.74E-01
Range	1.08E-02	1.13E-01	1.49E-01	1.64E-01
Minimum	5.13E-03	1.85E-02	9.20E-03	4.84E-02
Maximum	1.59E-02	1.32E-01	1.58E-01	2.13E-01
Sum	4.22E+01	2.63E+02	2.04E+02	5.10E+02
Count	5.E+03	5.E+03	5.E+03	5.E+03



Source: FHWA

Figure 21. 2045 non-auto share distribution.

As discussed in chapter 4, the theoretical 2045 non-automobile distribution shown in Figure 21 was not used in the response surface simulation for two main reasons. First, the simulated non-automobile distribution was based on aggregate trends for the United States, but Toledo and Chattanooga may not reach a double-digit non-auto share by 2045. Second, the initial response surface equations yielded non-intuitive results when extrapolating beyond a small change in the non-auto share, particularly for Toledo.

Instead of using the distribution from Table 28, the absolute growth in non-automobile preferences from 2010 to 2045 was initially assumed to be normal with mean 0.50% and standard deviation 0.1%. However, this assumption may have been too conservative, and was relaxed for Chattanooga in a second experiment, as described in Section 4.4.

3.0 Univariate Sensitivity Analyses

Univariate "sensitivity" analyses can provide a partial quantitative description of how forecasts depend on individual inputs. To conduct these analyses, the analyst will adjust the value of a single input and record how much the forecast changes. An elasticity can be computed from the size of the output change relative to the size of the input change.

At their best, univariate analyses are quick and effective, describing basic input and output relationships without requiring the analyst to estimate input probability distributions or conduct many model runs. However, univariate analyses do not indicate complex multivariate interactions and generally do not reveal the likelihood of different outcomes.

Formal probability distributions have already been defined for the uncertain inputs in the case studies; however, these probability distributions are not needed to conduct basic univariate analyses. Rather, these distributions are needed for the final more advanced multivariate analyses discussed in the next chapter. The input probability distributions will not be used in this chapter other than to loosely guide the test values.

This chapter shows how changes to individual model inputs affect the performance measures. Each section discusses a different input.

3.1 Telecommuting

The US Census indicates that telecommuting has increased from 4.8% in 1997 to 6.6% in 2010. This upward trend is likely to continue with advances in communication and information technologies. For the sensitivity tests, the percentage of 2045 telecommuting is increased to about 10%.

3.1.1 Toledo Results

A 10% telecommuting share was indirectly represented in the Toledo model through a 10% reduction in work trips. Table 19 shows the change in work and total trips. The overall number of trips decreased by about 2%, which is a moderate but meaningful change.

Table 19. Change of trip generation in telecommuting scenario for trip-based model.

Purpose	Base Scenario	10% Telecommuting Scenario	Difference
Home based work trips	323,106	290,795	-10%
Non-home based work trips	142,437	128,194	-10%
All Other Trips	1,965,077	1,965,077	0%
Grand Total	2,430,620	2,384,066	-2%

Table 20 shows the sensitivity of Toledo performance measures to a 10% telecommuting share. Transit ridership was the most sensitive measure, decreasing by 6.3%. This result is not unexpected since a relatively large number of transit trips are home-based work (HBW) and more telecommuting directly reduces commuting. VMT and delay decreased by a moderate amount comparable to the percent decrease in overall trips (see Table 19). Delay decreased slightly more than VMT likely because commutes often occur during the most congested periods of the day. The reduction in walk and bike trips was more moderate likely because these are not common HBW modes in Toledo.

Table 20. Toledo model sensitivity to telecommuting.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-2.18%	-.20
Delay	-2.41%	-.22
Transit Ridership	-6.3%	-.58
Walk/Bike Trips	-1.3%	-.12

3.1.2 Chattanooga Results

The regional Chattanooga telecommuting share was increased to a little over 10% by adjusting the work-at-home constant in the work location choice model. Table 21 shows the percentage and absolute number of telecommuting workers in the base and telecommuting scenarios. The overall percentage of employed persons working from home increased from 3.6% to 10.6% while the absolute number increased from 7,105 to 20,803. Some areas saw a relatively large increase in telecommuting; for example, the percentage of Downtown Chattanooga residents working from home increased from 14.9% to 32.6%.

Table 21. Change of commuting trips in telecommuting scenario for activity-based model.

Areas	Base Scenario (Percentage of work at home)	Base Scenario (Number of workers working at home)	10 % Telecommuting Scenario (Percentage of work at home)	10 % Telecommuting Scenario (Number of workers working at home)
Downtown Chattanooga	14.9%	894	32.6%	1,957
Near East Side	12.0%	1,012	27.6%	2,331
East Ridge	1.6%	263	5.1%	850
Red Bank	3.8%	786	12.6%	2,629
Lookout mountain	4.3%	911	13.7%	2,872
Soddy-Daisy	3.4%	767	10.6%	2,410
Middle Valley	1.4%	256	4.5%	837
Ridgeside	1.8%	257	6.3%	879
Harrison	1.4%	264	4.3%	811
Collegedale	3.4%	778	11.0%	2,483
Catoosa	3.5%	917	10.6%	2,744
Total	3.6%	7,105	10.6%	20,803

Table 23 shows the total tours by purpose for the base and telecommuting scenarios. As expected, the number of work and work-related tours decreased in response to the increase of telecommuting. At the same time, the number of non-work related tours increased, which may be due to the new telecommuters starting more tours from home rather than work and using some of the travel time saved from telecommuting to make other trips.

Table 22. Change of tour generation in telecommuting scenario for activity-based model.

Purpose	Base	10% Telecommuting	Difference
work	154,010	140,467	-9%
school	74,838	75,363	1%
escort	70,170	73,676	5%
pers.bus	85,862	88,489	3%
shop	53,021	54,469	3%
meal	30,162	30,890	2%
soc/rec	53,122	54,184	2%
workbased	21,654	18,888	-13%
Total	542,839	536,426	-1%

Table 23 shows the sensitivity of Chattanooga performance measures to a 10% telecommuting share. VMT, delay, and transit ridership all decreased from increased telecommuting as would generally be expected. Even though the overall number of tours only decreased by 1% (Table 22), the nearly 10% decrease in work tours could explain the somewhat larger changes in VMT, delay and transit since work trips tend to be longer, in more congested periods, and more likely to use transit. The decreasing in VMT is similar to that predicted by the Toledo model, while the Chattanooga model predicts a somewhat larger decrease in delay. This may be a result of greater based levels of peak period congestion due to chokepoints in the Chattanooga network due to the mountains and rivers. The decrease in transit ridership is greater than the decrease in trips or VMT but less than the decrease predicted by the Toledo model. This may be due to a difference in the work/non-work split of transit riders in the two cities. Walk and bike trips are the most sensitive to the 10% telecommuting shares in Chattanooga, which is notably different from the Toledo result, where the number of walk and bike trips decreased slightly (Table 20). This may be a result of the fact that the model predicted people living in downtown Chattanooga to be more likely to telecommute. These people may invest the time saved from telecommuting in short non-work walking trips such as walking to lunch.

Table 23. Chattanooga model sensitivity to telecommuting.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-1.73%	-0.18
Delay	-4.56%	-0.47
Transit Ridership	-3.07%	-0.39
Walk/Bike Trips	6.9%	0.72

3.2 Parking Costs

The case studies test an average regional parking cost of $15 per day. Relative to the 2045 base scenarios, the $15 cost represents a major price increase for both Toledo and Chattanooga.

3.2.1 Toledo Model

Average parking costs were increased from $3.64 to $15 by scaling values for the zonal attribute "AVG_PARK."

The total regional trips did not change since the trip generation model is not sensitive to parking costs. Transit shares increased for almost every purpose, with home-based work (HBW) shares increasing the most (see Table 24). The induced transit commuting is intuitive since transit service to downtown areas and employment centers is generally more competitive and could capture some choice riders.

Table 24. Transit share by purpose in parking cost increase scenario.

Purpose	Base Scenario	Parking Cost Increase Scenario	Difference
Home based work trips	2.20%	2.69%	0.49%
Home base shopping	0.43%	0.50%	0.07%
Home based school trips	0.33%	0.35%	0.01%
Home based high school trips	0.37%	0.48%	0.11%
Home based college trips	0.37%	0.37%	0.00%
Home based other trips	0.45%	0.70%	0.24%
Non-home based work trips	0.05%	0.18%	0.13%
Non-home based other trips	0.12%	0.36%	0.24%
TOTAL	0.64%	0.87%	0.24%

Table 25 shows the sensitivity of the Toledo performance measures to $15 parking. Transit ridership is clearly the most sensitive of the performance measures. The number of riders increased by 37.1% and the .14 elasticity suggests that doubling parking costs could induce nearly 15% more riders. The other performance measures, particularly VMT and delay, are largely insensitive to parking prices. Thus, despite a relatively large increase in transit ridership, aggregate VMT and delay statistics only decrease slightly because automobile use is very prevalent in Toledo. Parking cost had little effect on non-motorized trips, but might induce a small number of trips to switch from automobile to walking or biking.

Table 25. Toledo model sensitivity to parking costs.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-.05%	-.00019
Delay	-.07%	-.00028
Transit Ridership	37.1%	.14
Walk/Bike Trips	.19%	.00072

3.2.2 Chattanooga Results

For the Chattanooga parking cost sensitivity test, the average daily parking price was increased from $0.50 to $15 by scaling prices in the microzone file. (The base parking cost should have been coded as $5 and this later corrected, but the sensitivity analysis presents the results with the original scenario as stated.)

The parking cost increase resulted in a very small decrease in total generated tours (Table 26). The share of drive-alone decreased from 41.3% to 40.2%, while the shares of high-occupant vehicle and transit tours increased (see Table 27). The direction of these changes is intuitive.

Table 26. Change of total tours by purpose.

Purpose	Base Scenario	Parking Cost Increase Scenario	Difference
work	154,010	153,230	-0.51%
school	74,838	74,609	-0.31%
escort	70,170	70,245	0.11%
pers.bus	85,862	85,841	-0.02%
shop	53,021	52,940	-0.15%
meal	30,162	30,254	0.31%
soc/rec	53,122	52,775	-0.65%
workbased	21,654	21,575	-0.36%
Total	542,839	541,469	-0.25%

Table 27. Change of tour mode.

Mode	Base Scenario Share	Parking Cost Increase Scenario Share	Base Scenario Trips	Parking Cost Increase Scenario Trips
Drive Alone	41.3%	40.2%	224,051	217,768
Shared Ride 2	28.2%	28.3%	152,905	153,084
Shared Ride 3+	23.7%	24.1%	128,854	130,535
Drive-Transit	0.1%	0.1%	317	308
Walk-Transit	0.5%	0.7%	2,931	3,744
Bike	0.1%	0.2%	678	882
Walk	1.9%	2.3%	10,548	12,379
School Bus	4.1%	4.2%	22,525	22,739
Total	100.0%	100.0%	542,809	541,439

Table 28 shows the sensitivity of the Chattanooga performance measures to $15 parking. VMT decreases by 0.57% in the test scenario while delay decreases by 2.09%. Delay may decrease by more than VMT since the higher parking cost discourages trips to the more congested downtown area. Total transit boardings increased by 83.36%, which greatly exceeds the percent increase in transit tours. Since the CBD area was most affected by the increase in parking cost, the disproportionate increase in total boardings may be due to riders making multi-transfer trips to the downtown area rather than pay the parking costs. Walk and bike trips also increased, although not as much as transit. In general, these changes predicted by the Chattanooga model were similar in direction but greater in magnitude than those predicted by the Toledo model.

Table 28. Chattanooga model sensitivity to parking costs.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-0.57%	-.00002
Delay	-2.09%	-0.00008
Transit Ridership	83.36%	0.0032
Walk/Bike Trips	17.61%	0.00067

3.3 Fuel Prices

Fuel price increases can lead to less VMT and delay and cause substitution of transit for private automobile. The case studies tested an increase in 2045 fuel prices to $0.27 per mile (or roughly $8/gallon assuming around a 30 mile per gallon fleet fuel efficiency).

3.3.1 Toledo Results

For the Toledo sensitivity test, auto operating cost was increased from $0.10 to $0.27 by adjusting a mode choice parameter. Thus, trip generation and other model steps occurring before mode choice were not directly affected.

In response to the fuel price increase, walk-to-transit mode shares increased slightly for each trip purpose, however, drive-to-transit shares decreased for each purpose (Table 29). The overall increase in walk-to-transit trips slightly outweighed the overall decrease in drive-to-transit trips, but the opposite was true for the home-based work purpose.

Table 29. Change of walk to transit and drive to transit in fuel price increase scenario.

	Base Scenario		Fuel price increase Scenario		Absolute Difference	
PURPOSE	**Walk to Transit**	**Drive to Transit**	**Walk to Transit**	**Drive to Transit**	**Walk to Transit**	**Drive to Transit**
Home based work trips	1.570%	0.629%	1.848%	0.315%	0.278%	-0.314%
Home base shopping	0.393%	0.034%	0.483%	0.018%	0.091%	-0.016%
Home based school trips	0.307%	0.024%	0.352%	0.013%	0.044%	-0.011%
Home based high school trips	0.340%	0.030%	0.393%	0.017%	0.053%	-0.013%
Home based college trips	0.340%	0.033%	0.421%	0.015%	0.081%	-0.018%
Home based other trips	0.419%	0.035%	0.501%	0.019%	0.081%	-0.017%
Non-home based work trips	0.051%	0.000%	0.058%	0.000%	0.008%	0.000%
Non-home based other trips	0.120%	0.000%	0.135%	0.000%	0.015%	0.000%
TOTAL	0.512%	0.125%	0.607%	0.063%	0.094%	-0.062%

Table 30 shows the sensitivity of the Toledo performance measures to the fuel price increase. Transit ridership increased by about 4.9%. Although it was the most sensitive of the performance measures, the .029 transit elasticity indicates that a doubling of fuel prices might only increase demand by about 3%. VMT and delay decreased by a very small amount in response the fuel price increase. Their elasticities suggest that a doubling of prices might reduce VMT and delay only by .13% and .15% respectively. These modest responses are likely due to representing auto operating cost only in the mode choice model step. In theory, higher fuel prices would also suppress demand, cause shorter trips, and may even affect land use patterns.

Table 30. Toledo model sensitivity to fuel prices.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-.215%	-.00013
Delay	-.248%	-.00015
Transit Ridership	4.897%	.029
Walk/Bike Trips	.073%	.000043

3.3.2 Chattanooga Results

For the Chattanooga sensitivity test, auto operating cost was increased to $0.27 by changing a parameter in configuration file.[2]

Table 31 shows the change in tours by purpose from the fuel price increase. The overall number of tours decreases by a very small amount while each purpose increases or decreases by a modest amount. The increases in some purposes may be an artifact of simulation variation or in some cases (i.e., escort, meal, work-based) legitimate increases in trip-chaining behavior.

Table 31. Change in tours by purpose in fuel price increase scenario.

Purpose	Base Scenario	Fuel Price Increase Scenario	% Diff
Work	154,010	152,872	-0.74%
School	74,838	74,163	-0.90%
Escort	70,170	70,418	0.35%
Per. Bus	85,862	85,917	0.06%
Shop	53,021	52,736	-0.54%
Meal	30,162	30,433	0.90%
Soc/rec	53,122	52,854	-0.50%
Workbased	21,654	22,028	1.73%
Total	542,839	541,421	-0.26%

Table 32 shows the mode share changes. The mode shares generally increase or decrease in the expected direction. The model forecasts a shift from drive alone to shared ride and a general

[2] "Coefficients_BaseCostCoefficientPerMonetaryUnit," which has units of dollars per mile, was changed

shift to active transportation. While transit also increases, the 0.05% absolute increase in the transit share is notably less than the shared ride and active transportation increases.

Table 32. Change of tour mode share in fuel price increase scenario.

Mode	Base	Fuel Price Increase	Absolute Change	Relative Change
Drive Alone	41.28%	39.75%	-1.52%	-4%
Shared Ride 2	28.17%	28.29%	0.12%	0%
Shared Ride 3+	23.74%	24.77%	1.03%	4%
Drive-Transit	0.06%	0.06%	0.00%	8%
Walk-Transit	0.54%	0.59%	0.05%	9%
Bike	0.12%	0.15%	0.02%	18%
Walk	1.94%	2.20%	0.25%	13%
School Bus	4.15%	4.19%	0.04%	1%
Total	100.00%	100.00%	0.00%	0%

Table 34 shows the sensitivity of Chattanooga performance measures to the increase in fuel price. Transit riders increased by about 10% while walk/bike trips increased by about 13%. Despite a double-digit percent increase in transit ridership, the model sensitivity is still moderate as the .059 transit elasticity indicates that a doubling of fuel costs may result in only 6% more riders. VMT and delay decreased by a moderate amount in response to the price increase. Fewer total tours; shorter tours, and a shift to transit and active transportation may all contribute to the decrease in VMT and delay. While these changes are generally modest, they are substantially larger than those predicted by the Toledo model.

Table 33. Chattanooga model sensitivity to fuel prices.

Performance Measure	Percent Change in Test Scenario	Elasticity
VMT	-2.60%	-.01528
Delay	-6.35%	-.03733
Transit Ridership	10%	.059
Walk/Bike Trips	13%	.076

3.4 Transit Fares

Transit demand will generally decrease from a fare increase, although some “captive” riders may continue using transit particularly in the short-term. The case studies tested both a 50% reduction and a 100% increase in transit fare.

3.4.1 Toledo Results

For the Toledo sensitivity test, the 50% fare reduction and the 100% fare increase were implemented by halving and doubling the boarding time penalty.

Table 34 shows the change in transit share by purpose from the fare adjustments. Since home-based work (HBW) is the only purpose to include a boarding time penalty in its transit accessibility measure, the HBW share transit share is the only one to change in the sensitivity tests.

Table 34. Percent change in Toledo transit shares from fare adjustments.

PURPOSE	Base Scenario	50% Price Scenario	Change	200% Price Scenario	Change
Home based work trips	2.20%	2.49%	0.29%	1.72%	-0.48%
Home base shopping	0.43%	0.43%	0.00%	0.43%	0.00%
Home based school trips	0.33%	0.33%	0.00%	0.33%	0.00%
Home based high school trips	0.37%	0.37%	0.00%	0.37%	0.00%
Home based college trips	0.37%	0.37%	0.00%	0.37%	0.00%
Home based other trips	0.45%	0.45%	0.00%	0.45%	0.00%
Non-home based work trips	0.05%	0.05%	0.00%	0.05%	0.00%
Non-home based other trips	0.12%	0.12%	0.00%	0.12%	0.00%
TOTAL	0.64%	0.69%	0.05%	0.56%	-0.08%

Table 35 shows the sensitivity of the Toledo performance measures to the transit fare changes. As expected, transit ridership was the most sensitive of the measures. The estimated elasticity was between -.13 and -.15, suggesting that a 100% increase in transit fare might result in a 13% to 15% decrease in ridership. While elasticities between -.35 and -.65 are typically found in the literature, the overall Toledo elasticity was expected to be low since only work trips were (indirectly) sensitive to fare changes.

The VMT and delay changes were expected to be very small since the base transit share is much less than the automobile share and only a major change in transit ridership would materially affect these statistics. However, the VMT and delay totals do change in the expected direction.

Table 35. Toledo model sensitivity to transit fares.

Performance Measure	Percent Change in 50% Fare Scenario	Elasticity in 50% Fare Scenario	Percent Change in 200% Fare Scenario	Elasticity in 200% Fare Scenario
VMT	-.01%	-.0002	.07%	.0007
Delay	-.02%	.0004	.08%	.0008
Transit Ridership	7.7%	-.15	-12.6%	-.13
Walk/Bike Trips	-.02%	-.0004	.12%	.0012

3.4.2 Chattanooga Results

For the Chattanooga sensitivity test, the fare reduction and increase were implemented by editing the fare field in the transit route system file.

Table 36 shows the percent change in mode share from the transit fare adjustments. The walk-to-transit share increased by about 33% from the 50% reduction and decreased by about 35% from the 100% fare increase. The drive-to-transit response was a little weaker; the share increased by about 13% from the fare reduction and decreased by about 30% from the fare increase. Walk and bike trips increased by over 2% from the fare increase, likely because some transit riders may not own an automobile or may live in urbanized areas that are conducive to short trips.

Table 36. Percent change in Chattanooga transit share from fare adjustments.

Mode	Base Scenario	50% Fare Scenario	200% Fare Scenario	Percent Change in 50% Fare Scenario	Percent Change in 200% Fare Scenario
Drive Alone	41.3%	41.2%	41.3%	-0.14%	0.09%
Shared Ride 2	28.2%	28.1%	28.2%	-0.19%	0.23%
Shared Ride 3+	23.7%	23.7%	23.8%	-0.34%	0.20%
Drive-Transit	0.1%	0.1%	0.0%	13.17%	-29.72%
Walk-Transit	0.5%	0.7%	0.4%	33.44%	-34.55%
Bike	0.1%	0.1%	0.1%	0.81%	2.56%
Walk	1.9%	2.0%	2.0%	1.14%	2.14%
School Bus	4.1%	4.1%	4.2%	-0.49%	0.28%
Total	100.0%	100.0%	100.0%	0.00%	0.00%

Table 37. Absolute increase in Chattanooga transit trips from fare adjustments.

Mode	Base	Half	Double	Half absolute change	Double absolute change
Drive Alone	224,051	223,890	224,450	-161	399
Shared Ride 2	152,905	152,728	153,387	-177	482
Shared Ride 3+	128,854	128,512	129,228	-342	374
Drive-Transit	317	359	223	42	-94
Walk-Transit	2,931	3,914	1,920	983	-1,011
Bike	678	684	696	6	18
Walk	10,548	10,676	10,783	128	235
School Bus	22,525	22,431	22,609	-94	84
Total	542,809	543,194	543,296	385	487

Table 38 shows the sensitivity of the Chattanooga performance measures to the transit fare changes. As expected, transit ridership is the most sensitive measure. The Chattanooga fare elasticities are generally in line with those found in the literature. However, the fact that the elasticity estimated from the price reduction scenario, -.67, and the elasticity estimated from the price increase scenario, -.35, are notably different illustrates how elasticities are most appropriate for forecasting moderate changes to an input and must be used carefully when forecasting major changes to an input. For example, using the -.67 elasticity to forecast a 100% price increase would give a much different answer from using -.35 to forecast a 100% price increase.

Table 38. Chattanooga model sensitivity to transit fares.

Performance Measure	Percent Change in 50% Fare Scenario	Elasticity in 50% Fare Scenario	Percent Change in 200% Fare Scenario	Elasticity in 200% Fare Scenario
VMT	-.18%	.0037	.04%	.0004
Delay	-.48%	.0097	.25%	.0025
Transit Ridership	33.69%	-.67	-34.90%	-.35
Walk/Bike Trips	-.95%	-.0190	1.89%	.019

4.0 Response Surface Simulation

Response surface simulation can be viewed as the multivariate generalization of univariate "sensitivity testing." Even when input probability distributions are not defined, the analyst can apply the techniques discussed in sections 4.1 and 4.2 to quantify multivariate input-output relationships and apply the reduced form response surface model to specific scenarios to quickly generate a much wider range of scenario outcomes than could be produced by the travel model itself. When the input probability distributions are also defined, the analyst can apply a technique such as Monte Carlo simulation to estimate the full distribution of outcomes, as demonstrated in sections 4.3 and 4.4.

Section 4.1 describes the experimental design of the response surface simulation used in the case studies. Section 4.2 shows the reduced form models, or regression equations estimated to quantify multivariate input-output relationships. The simulated input distributions are presented in Section 4.3, and the results are discussed in Section 4.4.

4.1 Experimental Design

There are many factors in a regional travel demand model that affect travel forecasts. The objective of the experimental design for this quantitative uncertainty modeling is to support the development of a simplified model of the effects of factors whose future values are uncertain and that have a significant effect on travel demand. For this application, ten such factors were identified.

1. Auto Operating Cost
2. Telecommuting
3. Transit Service Cost
4. Parking Cost
5. Non-auto preference Mode Share
6. Urban Core Pop Growth
7. Urban Core Employment
8. Remainder/Halo Area Population Growth
9. Remainder/Halo Area Employment Growth
10. Boom City Growth and Employment

Plausible ranges of values for each of these factors were determined and associated with discrete levels of the factors. For six of the factors which were assumed to be of generally lower impact on travel demand and for which the effects on travel demand were assumed to be linear, only two discrete levels were specified. For the other four factors, four discrete levels were used which means that up to three coefficients could be estimated for each, allowing, for example, estimation of quadratic or cubic polynomials to represent effects on travel demand (see Table 39).

Table 39. Levels for each factor.

Factor	Level	Definition
Auto Operating Cost	1	Base
Auto Operating Cost	2	$0.27/mile
Telecommuting	1	Base
Telecommuting	2	10% increase
Transit Service Cost	1	0.5 * Base
Transit Service Cost	2	2 * Base
Parking Cost	1	Base
Parking Cost	2	$15/day

Factor	Level	Definition
Non-Automobile Preference	1	Base
Non-Automobile Preference	2	Base + 0.5%
Urban Core Pop Growth	1	Low (-2 SD)
Urban Core Pop Growth	2	Medium-Low (-2/3 SD)
Urban Core Pop Growth	3	Medium-High (+2/3 SD)
Urban Core Pop Growth	4	High (+2 SD)
Urban Core Emp Growth	1	Low (-2 SD)
Urban Core Emp Growth	2	Medium-Low (-2/3 SD)
Urban Core Emp Growth	3	Medium-High (+2/3 SD)
Urban Core Emp Growth	4	High (+2 SD)
Halo Pop Growth	1	Low (-2 SD)
Halo Pop Growth	2	Medium-Low (-2/3 SD)
Halo Pop Growth	3	Medium-High (+2/3 SD)
Halo Pop Growth	4	High (+2 SD)
Halo Emp Pop Growth	1	Low (-2 SD)
Halo Emp Pop Growth	2	Medium-Low (-2/3 SD)
Halo Emp Pop Growth	3	Medium-High (+2/3 SD)
Halo Emp Pop Growth	4	High (+2 SD)
Boom City Growth	1	Growth consistent with Halo
Boom City Growth	2	Boom city scenario for population and employment

With six two-level and four four-level factors, there are 2^6x4^4 = 16,384 different possible combinations of levels. In theory, one could run the travel demand model with each of these combinations and use the resulting data to statistically determine how each level of each factor affected travel demand. However, doing all those runs would require a very significant computational effort which would likely make the effort infeasible. The alternative is to choose a subset of the 16,384 possible combinations in a way that allows the individual effects of each factor to be estimated reasonably accurately and does so parsimoniously. The minimum number of combinations necessary to be run to estimate coefficients for each level beyond a “base” level of each factor can be determined by summing the number of levels minus one for each factor. So, in this case, we would need at least six combinations to estimate the effects of the six two-level factors plus 12 (4x3) for the four-level factors, resulting in 18 total. Note that this is the minimum number required assuming the resulting model includes 18 coefficients.

But just choosing to run any 18 combinations would not necessarily allow us to develop good statistical estimates of the effects of these factors on travel demand. For example, choosing them in a way that not all levels of all factors are represented would preclude estimation of coefficients for those levels. In addition, always showing certain levels of one factor with specific levels of another factor would confound the ability to separately estimate the effects of those two factors. There are many ways to construct good experimental designs that consist of a small subset of the total number of possible combinations. The subset is called a fractional factorial design and a design in which the levels of all factors are uncorrelated with the levels of each other factor is called an orthogonal fractional factorial design. These types of designs in general result in the best (lowest variance) estimates of effects of factors in linear models but can also be used to estimate nonlinear effects (though other designs would generally be more efficient in those cases).

There is a large literature on experimental design, including several good texts[3] and some easily accessible tools to develop different types of designs. A US DOT guide[4], available in various places online, includes a set of tables of orthogonal fractional factorial designs developed by early researchers in the field. These tables span a very wide range of designs for different numbers of factors with different numbers of levels. In addition, there are open-source r packages that can be used to generate both simple and more complex designs of various types.[5]

A 20-experiment orthogonal fractional factorial design was developed for this application, using the open-source r package AlgDesign. The resulting design table is shown below. The levels, numbered 1-4, correspond with the definitions shown in Table 39.

Table 40. Experiment orthogonal fractional factorial design.

Experiment	Auto Cost	Tele-Commute	Transit Cost	Parking Cost	Non-auto	Urban Pop	Urban Emp	Halo Pop	Halo Emp	Boom City
1	2	1	1	1	1	3	4	2	4	1
2	1	1	2	2	1	2	2	2	4	2
3	2	1	1	1	2	2	4	1	2	2
4	1	2	2	2	1	2	4	1	1	1
5	2	1	2	1	2	2	3	4	1	1
6	2	2	2	2	1	3	3	2	2	1
7	2	1	2	2	2	1	2	3	4	1
8	2	2	1	2	2	4	2	2	1	2
9	1	2	1	2	2	2	3	4	4	1
10	1	2	2	1	2	3	4	3	2	2
11	2	2	2	1	1	4	2	3	3	1
12	2	1	1	2	1	4	3	3	3	2
13	1	1	2	2	2	3	3	1	3	2
14	2	2	2	1	2	1	1	2	3	2
15	2	2	1	2	1	1	1	3	2	2
16	1	1	1	2	2	3	2	4	1	1
17	1	2	1	1	2	4	1	1	4	1
18	1	2	1	1	1	1	4	4	3	2
19	1	1	2	1	1	4	1	4	2	2
20	1	1	1	1	1	1	1	1	1	1

Because this is an orthogonal design, the bivariate correlations among all factors are all zero, as shown in Table 41.

[3] See, for example, Montgomery, Douglas (2013). *Design and analysis of experiments* (8th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

[4] Kocur, G., T. Adler, et al. (1982) *Guide to Forecasting Travel Demand with Direct Utility Assessment*, U.S. DOT, Urban Mass Transit Administration, Report UMTA-NH-11-0001-82-1, September 1982.

[5] For a current listing of available r resources for experimental design, see https://cran.r-project.org/web/views/ExperimentalDesign.html

Table 41. Bivariate correlations.

	Auto Cost	Tele-Commute	Transit Cost	Parking Cost	Non-auto	Urban Pop	Urban Emp	Halo Pop	Halo Emp	Boom City
Auto Cost	1	0	0	0	0	0	0	0	0	0
Telecommute	0	1	0	0	0	0	0	0	0	0
Transit Cost	0	0	1	0	0	0	0	0	0	0
Parking Cost	0	0	0	1	0	0	0	0	0	0
Non-Auto	0	0	0	0	1	0	0	0	0	0
Urban Pop	0	0	0	0	0	1	0	0	0	0
Urban Emp	0	0	0	0	0	0	1	0	0	0
Halo Pop	0	0	0	0	0	0	0	1	0	0
Halo Emp	0	0	0	0	0	0	0	0	1	0
Boom City	0	0	0	0	0	0	0	0	0	1

The travel demand model was run for each of the 20 experiments in this design and the resulting data were used to develop simple regression-based models that represent the effects of the design's ten factors on travel forecasts. The regression estimation results are presented in next section.

4.2 Reduced Form Model Estimation

Regression equations were estimated for each of five performance measures discussed in chapter 2:

- Vehicles miles traveled (VMT)
- Delay[6]
- Transit ridership
- Car emissions
- Truck emissions

The five regression equations have similar specifications and include at least one explanatory variable for each of the ten factors discussed in section 4.1. Some specifications have an additional term for the square of halo population, but otherwise the specifications are the same. Similar specifications were used to help compare the case study results, but much different specifications could be estimated for each performance measure in practice. The models could omit factors that are believed (or found) to be unrelated to the forecast variable. For example, transit fares may have an immaterial effect on truck emissions outside of certain metropolitan areas. The set of explanatory variables are shown below.

Table 42. Explanatory variables in regression equations for reduced form model estimation.

Explanatory Variable	Note
Constant	
Auto Cost (¢)	

[6] Measured as congested vehicle hours traveled minus free flow time vehicle hours

Explanatory Variable	Note
Telecommuting Share (% * 100)	
Transit Cost (ratio)	Scenario cost over base cost
Parking Cost ($)	
Non-auto Share (% * 100)	
Urban Core Pop Growth (%)	Growth from 2010 to 2045
Urban Core Emp Growth (%)	Growth from 2010 to 2045
Halo Pop Growth (%)	Growth from 2010 to 2045
Halo Emp Growth (%)	Growth from 2010 to 2045
Boom City (dummy)	Dummy variable
Halo Pop Growth Squared ($\%^2$)	Growth from 2010 to 2045

The estimated reduced form models are shown in Table 43 to Table 50. Coefficients and standard errors are reported for all the explanatory variables, even when the variables are not statistically significant. In practice, some of the statistically insignificant variables could be dropped.

While many of the estimated coefficients are intuitive, most of the models have one or more coefficients with unexpected signs or magnitudes. These results have been retained and presented here because they are practical challenge that others may well also encounter, and it helpful to illustrate their identification and discuss their treatment. These coefficients can sometimes be ignored because they are statistically insignificant and not believed to be related to the performance measure, and/or the model simply be re-estimated, dropping them from the specification. In other cases, the counter-intuitive coefficients have large t-statistics or are otherwise believed to be inappropriately related to the performance measure. In practice, the analyst could again consider a different functional form for these counter-intuitive terms, such as by squaring them or using a log transformation. The analyst could also alternatively improve the experimental design by adding variable levels that capture a wider range of the theoretical input distribution. For example, more auto cost levels could be used in this experiment, although adding levels increases the number of required runs. In practice, two other explanations for counter-intuitive results should always be considered. First, the user may have set up the model incorrectly. Second, the model may have a fundamental limitation or problem. While sometimes frustrating, identifying model limitations can be very useful for prioritizing future enhancements.

The Toledo and Chattanooga reduced form models for VMT are presented in Table 43 and Table 44, respectively. Diffuse population and employment growth were expected to contribute disproportionately to VMT; the models generally show positive correlation between VMT and growth in the halo areas.

Some of the other Toledo coefficients are not intuitive. For example, auto cost is positively correlated with VMT and has a high t-statistic. The Toledo result seems counter-intuitive since higher fuel prices would generally lead to fewer and shorter auto trips. However, as discussed in section 3.3, fuel prices only impact work trips in the Toledo model and only in the mode choice step and thus have a modest impact. Parking cost is positively but weakly correlated with VMT. This counter intuitive result may likewise be due to not considering parking cost until the mode choice step. For Chattanooga, auto cost shows a clear and strong negative relationship with VMT; parking cost also shows a negative relationship, but a much weaker, more tentative one.

Table 43. Regression results for VMT in Toledo (trip-based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	15,757,601.0	184,225.3	85.53	3.89E-13
Auto Cost	13,092.2	5,343.5	2.45	3.99E-02
Telecommute	-42,453.3	6,765.1	-6.28	2.39E-04
Transit Cost	35,241.2	45,100.9	0.78	4.57E-01
Parking Cost	2,793.4	6,132.7	0.46	6.61E-01
Non-Auto Preference	-308,132.0	105,704.9	-2.92	1.94E-02
Urban Population Growth	21,009.9	2,074.9	10.13	7.73E-06
Urban Employment Growth	-1,695.0	2,271.5	-0.75	4.77E-01
Halo Zones Population Growth	29,493.7	13,068.7	2.26	5.40E-02
Halo Zones Employment Growth	8,587.5	2,035.7	4.22	2.92E-03
Boom City	929,590.8	69,483.9	13.38	9.32E-07
Halo Population Growth Square	878.7	350.3	85.53	3.89E-13
Adjusted R Square				0.982

Table 44. Regression results for VMT in Chattanooga (activity-based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	13,860,181.3	554,104.2	25.014	.000
Auto Cost	-26,994.5	6,312.8	-4.276	.002
Telecommute	-5,861.3	15,507.4	-.378	.714
Transit Cost	38,232.2	71,316.1	.536	.605
Parking Cost	-612.0	7,188.1	-.085	.934
Non-Auto Preference	-363,941.9	118,890.0	-3.061	.014
Urban Population Growth	21,103.0	16,712.6	1.263	.238
Urban Employment Growth	-20,064.1	14,440.7	-1.389	.198
Halo Zones Population Growth	26,934.5	6,028.0	4.468	.002
Halo Zones Employment Growth	28,613.6	10,257.8	2.789	.021
Boom City	418,178.7	109,600.2	3.815	.004
Adjusted R Square				0.890

The Toledo and Chattanooga reduced form models for delay are presented in Table 45 and Table 46 respectively. Concentrated population and employment growth were expected to clearly contribute to delay since dense growth causes congestion. The results show that more growth, including the existence of the “Boom City”, generally increases delay. As expected, higher transit costs and lower non-auto shares also contribute to delay.

The Chattanooga model indicates that auto cost is negatively correlated with delay whereas the Toledo model indicates that auto cost is positively correlated with delay. As already discussed, this may be due to not considering auto cost until the mode choice step and only for work trips.

Table 45. Regression results for delay in Toledo (trip based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	44,572.1	6,723.8	6.63	0.00
Auto Cost (¢)	308.1	195.0	1.58	0.15
Telecommuting Share (% * 100)	-830.5	246.9	-3.36	0.01
Transit Cost (ratio)	1,260.6	1,646.1	0.77	0.47
Parking Cost ($)	-96.1	223.8	-0.43	0.68
Non-auto Share (% * 100)	-14,420.9	3,858.0	-3.74	0.01
Urban Core Pop Growth (%)	390.1	75.7	5.15	0.00
Urban Core Emp Growth (%)	-87.2	82.9	-1.05	0.32
Halo Pop Growth (%)	-345.2	477.0	-0.72	0.49
Halo Emp Growth (%)	215.7	74.3	2.90	0.02
Boom City (dummy)	21,101.3	2,536.0	8.32	0.00
Halo Pop Growth Squared ($\%^2$)	35.8	12.8	6.63	0.00
Adjusted R Square				0.927

Table 46. Regression results for delay in Chattanooga (activity based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	25,862.0	17,845.9	1.45	0.19
Auto Cost (¢)	-96.4	61.3	-1.57	0.15
Telecommuting Share (% * 100)	32.4	135.1	0.24	0.82
Transit Cost (ratio)	163.2	589.0	0.28	0.79
Parking Cost ($)	-62.1	91.3	-0.68	0.52
Non-auto Mode Share (% * 100)	-1,412.2	899.6	-1.57	0.16
Urban Core Pop Growth (%)	86.3	121.7	0.71	0.50
Urban Core Emp Growth (%)	100.3	279.1	0.36	0.73
Halo Pop Growth (%)	-9.7	43.1	-0.22	0.83
Halo Emp Growth (%)	-345.3	1,518.5	-0.23	0.83
Boom City (dummy)	1,692.1	822.6	2.06	0.07
Halo Pop Growth Squared ($\%^2$)	7.7	23.6	0.33	0.75
Adjusted R Square				0.869

The Toledo and Chattanooga reduced form models for transit ridership are presented in Table 47 and Table 48 respectively. Lower transit cost and higher non-auto preference were expected to increase transit use. Non-auto preference shows the expected positive relationship in both models; however, transit cost is only weakly correlated with ridership in the Toledo model and has virtually no correlation in the Chattanooga model. In practice, adding more transit cost levels to the experimental design may help the results. Higher parking and auto costs are also positively associated with transit use in both models as would be expected, but the auto cost correlation is weak in the Toledo model.

Dense population and employment growth is generally conducive to transit. While neither model showed a strong or consistent relationship between growth and transit ridership, the urban population growth coefficient was positive and significant for Toledo.

Table 47. Regression results for transit ridership in Toledo (trip based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	8,906.1	6,278.3	1.42	1.94E-01
Auto Cost	80.7	182.1	0.44	6.69E-01
Telecommute	44.3	230.6	0.19	8.53E-01
Transit Cost	-1,875.9	1,537.0	-1.22	2.57E-01
Parking Cost	565.7	209.0	2.71	2.68E-02
Non Auto Preference	22,576.0	3,602.4	6.27	2.41E-04
Urban Population Growth	187.5	70.7	2.65	2.92E-02
Urban Employment Growth	-4.5	77.4	-0.06	9.55E-01
Halo Zones Population Growth	137.0	445.4	0.31	7.66E-01
Halo Zones Population Growth	-9.6	69.4	-0.14	8.94E-01
Boom City	-2,097.6	2,368.0	-0.89	4.02E-01
Halo Population Growth Square	0.2	11.9	0.02	9.85E-01
Adjusted R Square				0.743

Table 48. Regression results for transit ridership in Chattanooga (activity based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	-615.5	8,233.6	-0.07	0.94
Auto Cost	143.2	68.6	2.09	0.07
Telecommute	0.3	124.9	0.00	1.00
Transit Cost	-6.1	568.5	-0.01	0.99
Parking Cost	179.8	55.7	3.23	0.01
Non Auto Preference	4,062.9	957.5	4.24	0.00
Urban Population Growth	71.4	128.9	0.55	0.60
Urban Employment Growth	-67.3	109.8	-0.61	0.56
Halo Zones Population Growth	-170.9	684.3	-0.25	0.81
Halo Zones Population Growth	-41.8	88.5	-0.47	0.65
Boom City	-1,425.6	908.2	-1.57	0.16
Halo Population Growth Square	3.1	11.5	0.27	0.79
Adjusted R Square	--	--	--	0.644

The Toledo and Chattanooga reduced form models for auto emissions are presented in Table 49 and Table 50 respectively. These results are analogous to the VMT and delay results, which was expected since more travel and frequent acceleration contribute to auto emissions. For Toledo, all types of population and employment growth are positively correlated with auto emissions. For Chattanooga, diffuse halo area growth is also positively correlated with auto emissions.

Table 49. Regression results for auto emissions in Toledo (trip based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	618,160.0	4,016.6	153.90	3.55E-15
Auto Cost	146.3	116.5	1.26	2.45E-01
Telecommute	-1,913.5	147.5	-12.97	1.18E-06
Transit Cost	713.2	983.3	0.73	4.89E-01
Parking Cost	33.5	133.7	0.25	8.09E-01
Non-Auto Preference	-8,114.4	2,304.6	-3.52	7.84E-03
Urban Population Growth	1,154.8	45.2	25.53	5.95E-09
Urban Employment Growth	149.9	49.5	3.03	1.64E-02
Halo Zones Population Growth	2,173.2	284.9	7.63	6.15E-05
Halo Zones Population Growth	148.0	44.4	3.33	1.03E-02
Boom City	39,032.1	1,514.9	25.77	5.52E-09
Halo Population Growth Square	21.7	7.6	2.85	2.16E-02
Adjusted R Square	--	--	--	0.996

Table 50. Regression results for auto emissions in Chattanooga (activity based) model.

Coefficients	Beta	Std. Error	T-Stat	Significance
Constant	440,141.5	19,792.5	22.238	.000
Auto Cost	-972.2	225.5	-4.312	.002
Telecommute	-319.2	553.9	-.576	.579
Transit Cost	1,607.7	2,547.4	.631	.544
Parking Cost	-33.7	256.8	-.131	.898
Non-Auto Preference	-12,508.9	4,246.7	-2.946	.016
Urban Population Growth	787.8	597.0	1.320	.220
Urban Employment Growth	-600.5	515.8	-1.164	.274
Halo Zones Population Growth	924.9	215.3	4.295	.002
Halo Zones Employment Growth	934.9	366.4	2.552	.031
Boom City	14,639.0	3,914.9	3.739	.005
Adjusted R Square	--	--	--	0.893

4.3 Monte Carlo Simulation

About 100,000 combinations of model input values were randomly drawn from the probability distributions discussed in Section 2.3. These sets of input values were entered into the reduced form models to simulate 2045 distributions for each performance measure. This section discusses the Monte Carlo simulation used to draw the input value combinations and presents a frequency distribution for each input. The next section discusses the simulated 2045 forecast distributions.

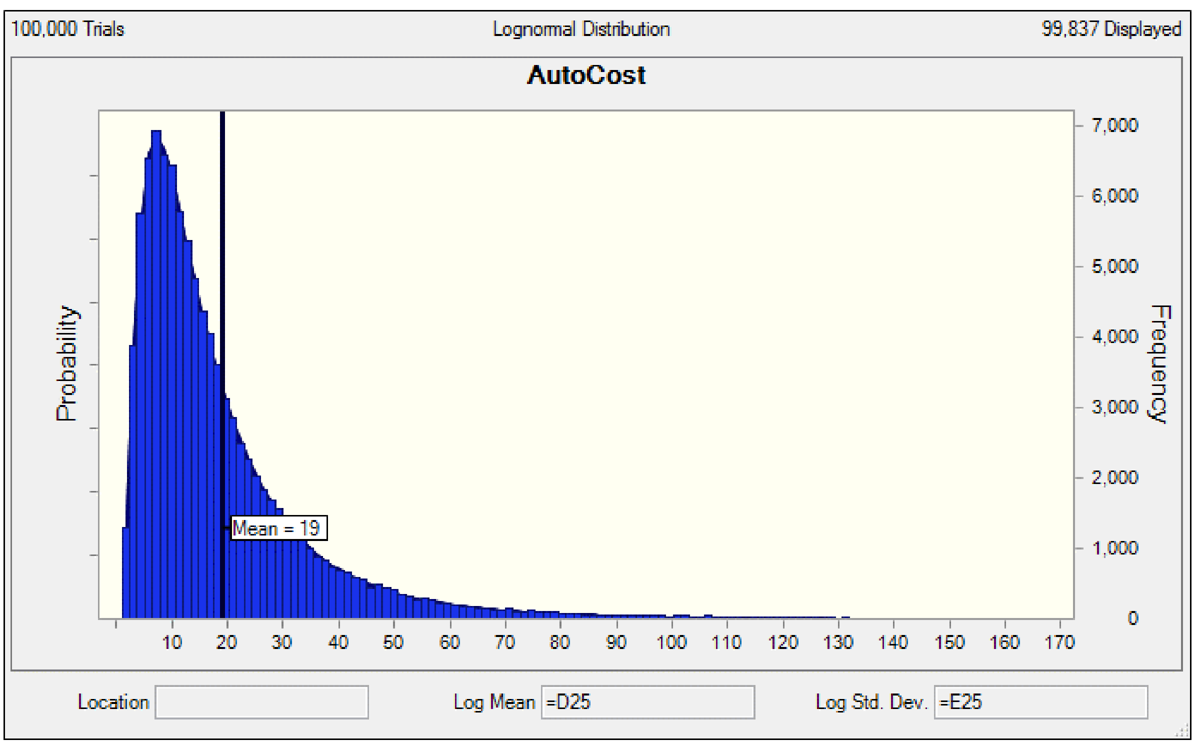
The Excel plug-in Crystal Ball was used to perform Monte Carlo simulation, which included about 100,000 random draws for each input variable. Figure 22 to Figure 35 show frequency distributions for each input value. The light blue bars show the frequency of each input value, and the dark blue background shape shows the theoretical distribution. The simulated and theoretical distributions aligned very well in most cases, suggesting that 100,000 draws were sufficient.

Population and employment growth percentages were simulated separately for Toledo and Chattanooga, but the models used the same frequency distributions for the other inputs.

As noted in section 2.3, the original 2045 non-auto preference distribution (Figure 21) had to be adjusted. The original distribution included a wide range of non-auto shares that extended well beyond the levels used in the experimental design; extrapolating to these high non-auto shares elicited overly sensitive model responses. Further, the original non-auto preference distribution was based on national data and may have included too much variation for auto-centric cities like Toledo and Chattanooga. A non-auto preference distribution with a lower mean and variance was ultimately asserted for the case studies; Figure 25 shows the asserted distributed of absolute growth in transit share relative to the base.

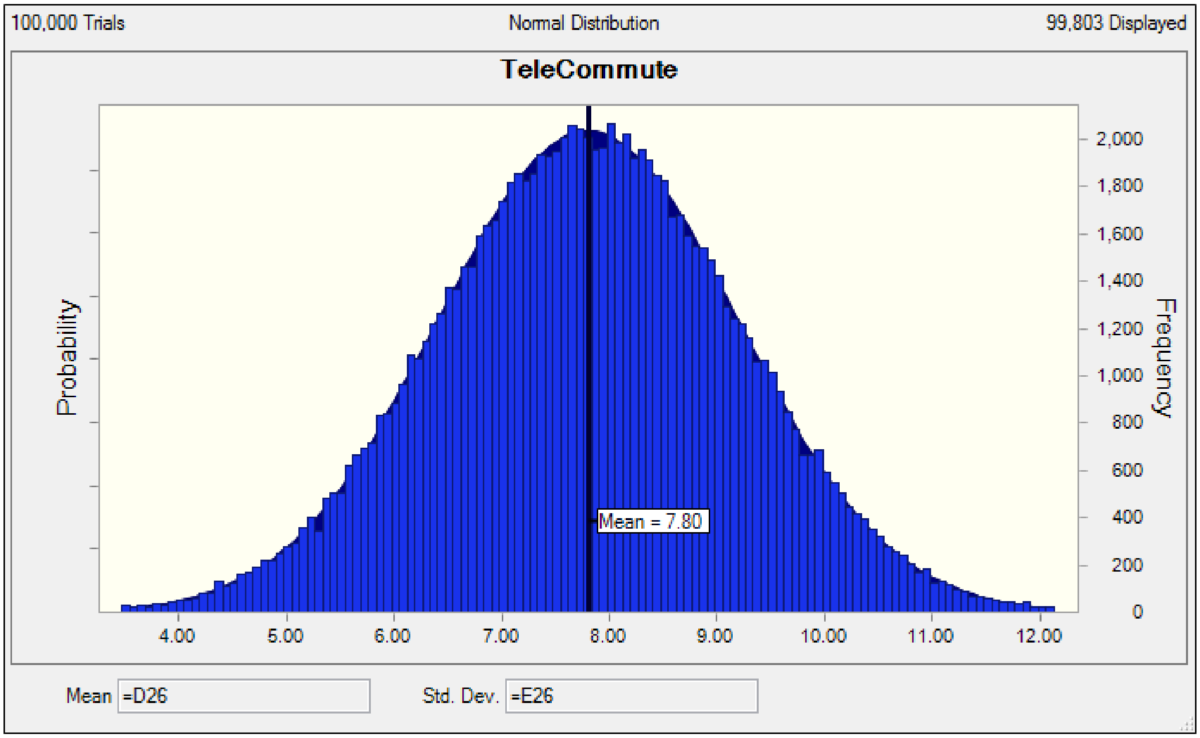
The presence of Boom City was treated as a dummy variable in the reduced form equations, with a value of 1 indicating that Boom City exists and a value of 0 indicating that Boom City does not exist. Although input frequency distribution for Boom City (Figure 35) is likewise bound between 0 and 1, the simulated distribution is continuous between 0 and 1, with “0” indicating that 0% of the Boom City growth occurs and “1” indicating that 100% of the growth occurs.

The simulations also included draws from the standard error of the regression distribution for each model. These error distributions are shown in Figure 36 through Figure 43. Generally speaking, drawing from the error distribution serves to increase the overall forecast variation, or uncertainty, by accounting for the known error in the estimated model.



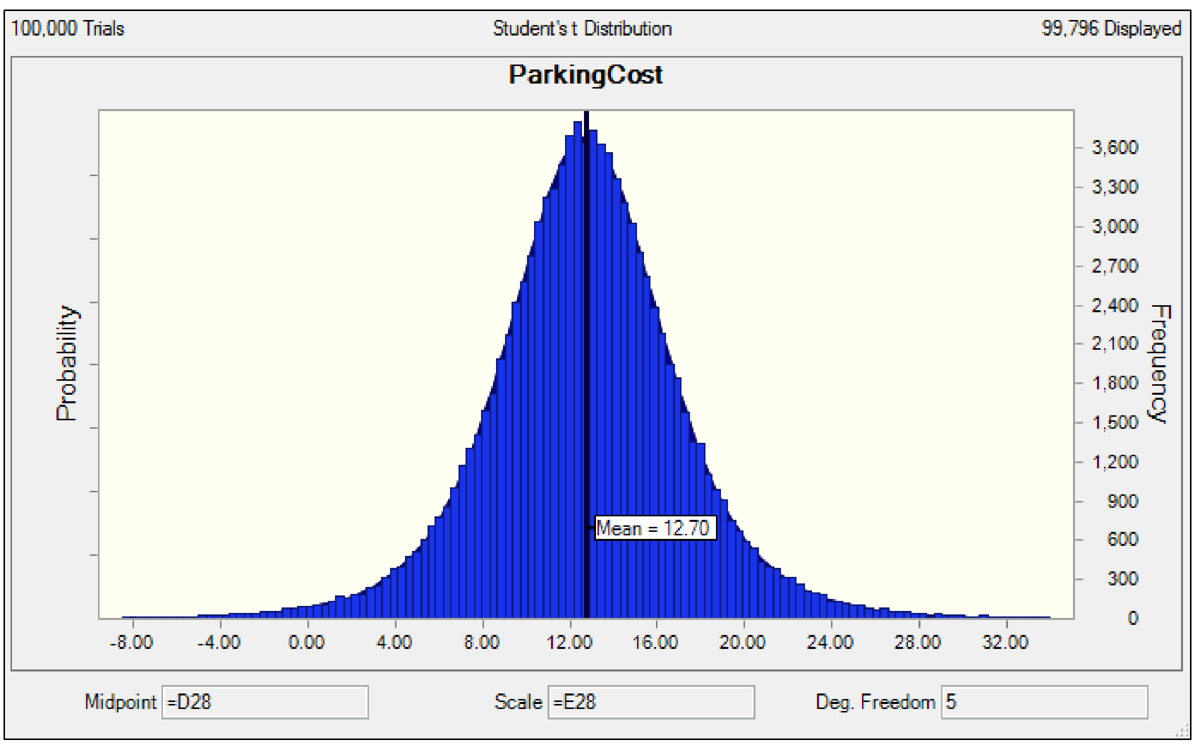
Source: FHWA

Figure 22. 2045 auto cost distribution (Toledo and Chattanooga).



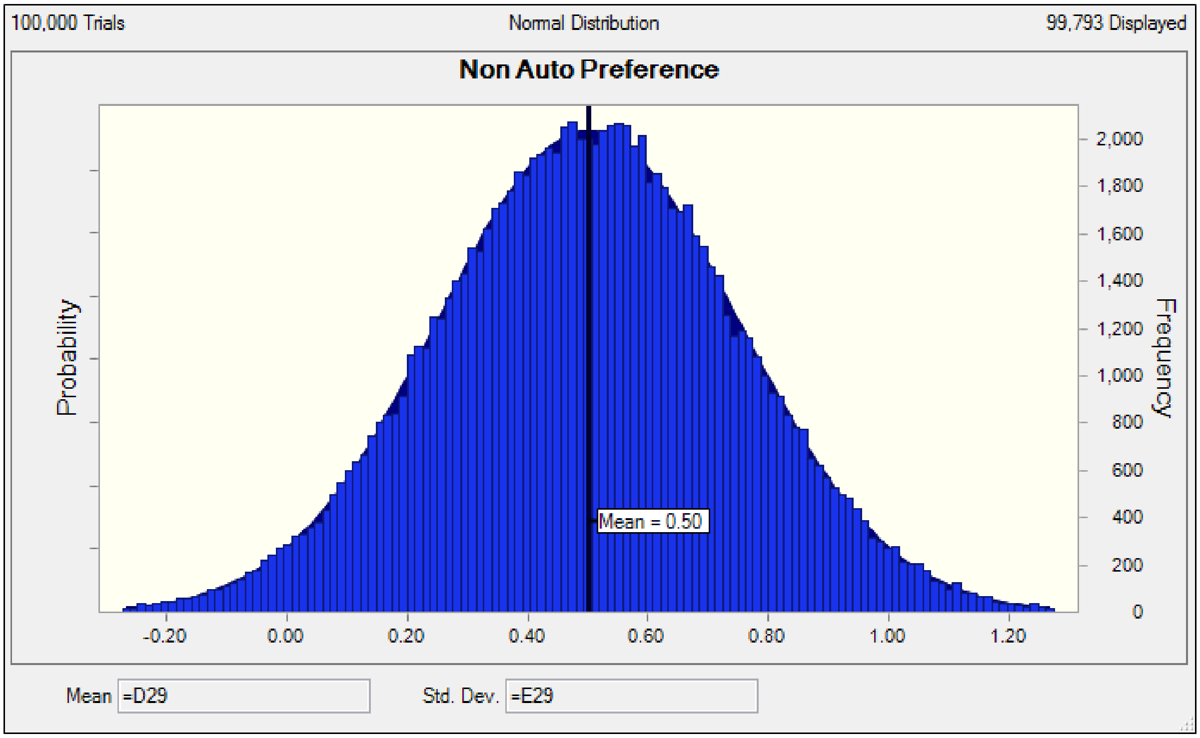
Source: FHWA

Figure 23. 2045 telecommuting share distribution (Toledo and Chattanooga).



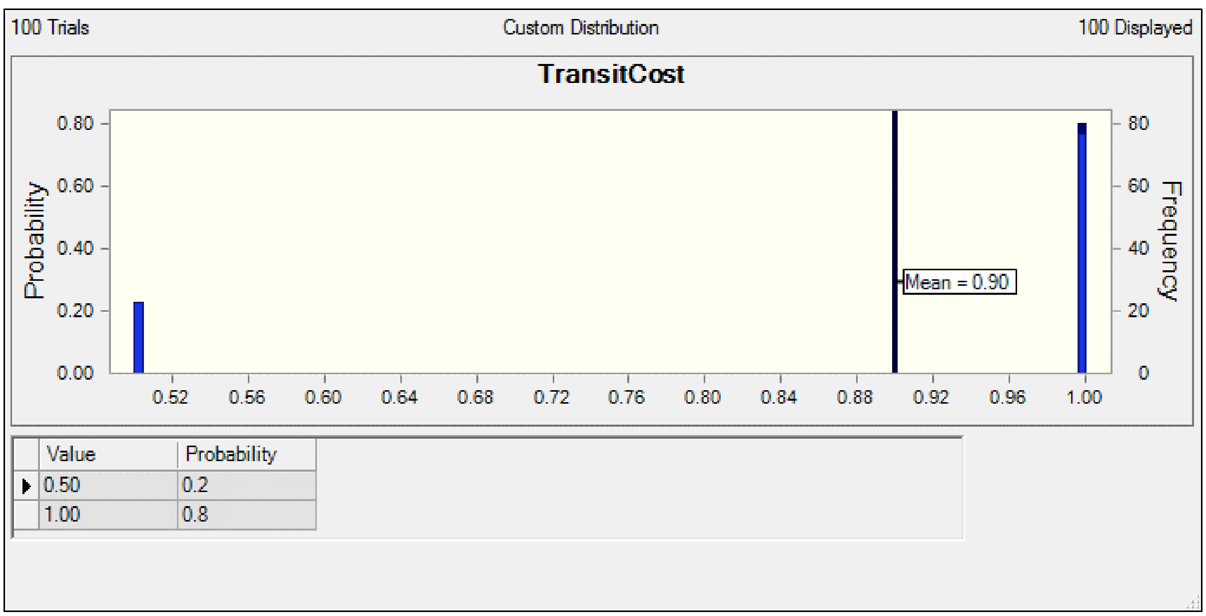
Source: FHWA

Figure 24. 2045 parking cost distribution (Toledo and Chattanooga).



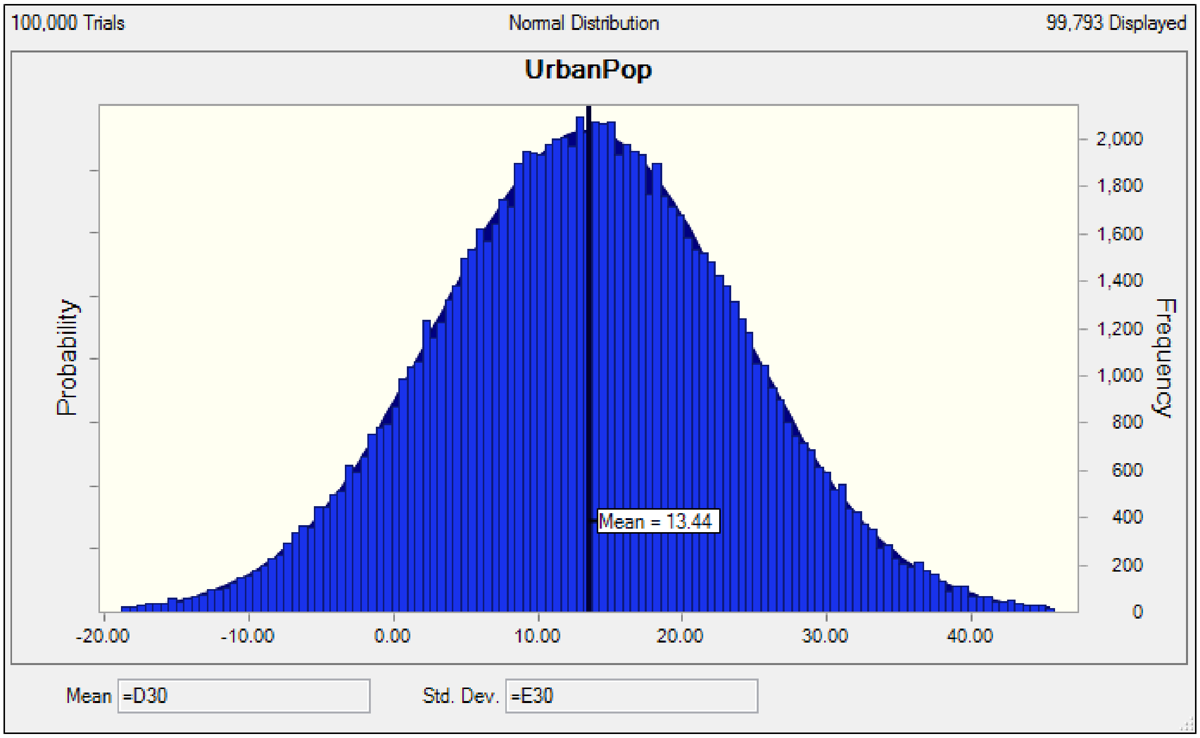
Source: FHWA

Figure 25. 2045 non-auto mode share distribution (Toledo and Chattanooga).



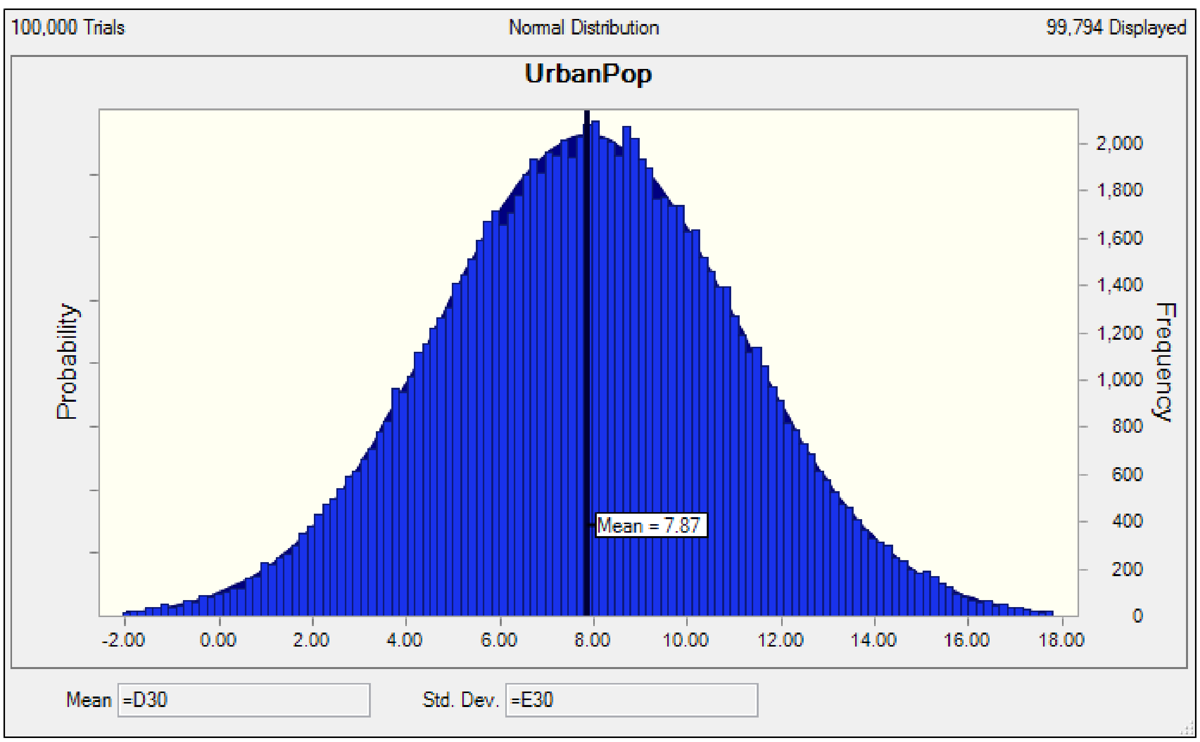
Source: FHWA

Figure 26. 2045 transit fare distribution (Toledo and Chattanooga).



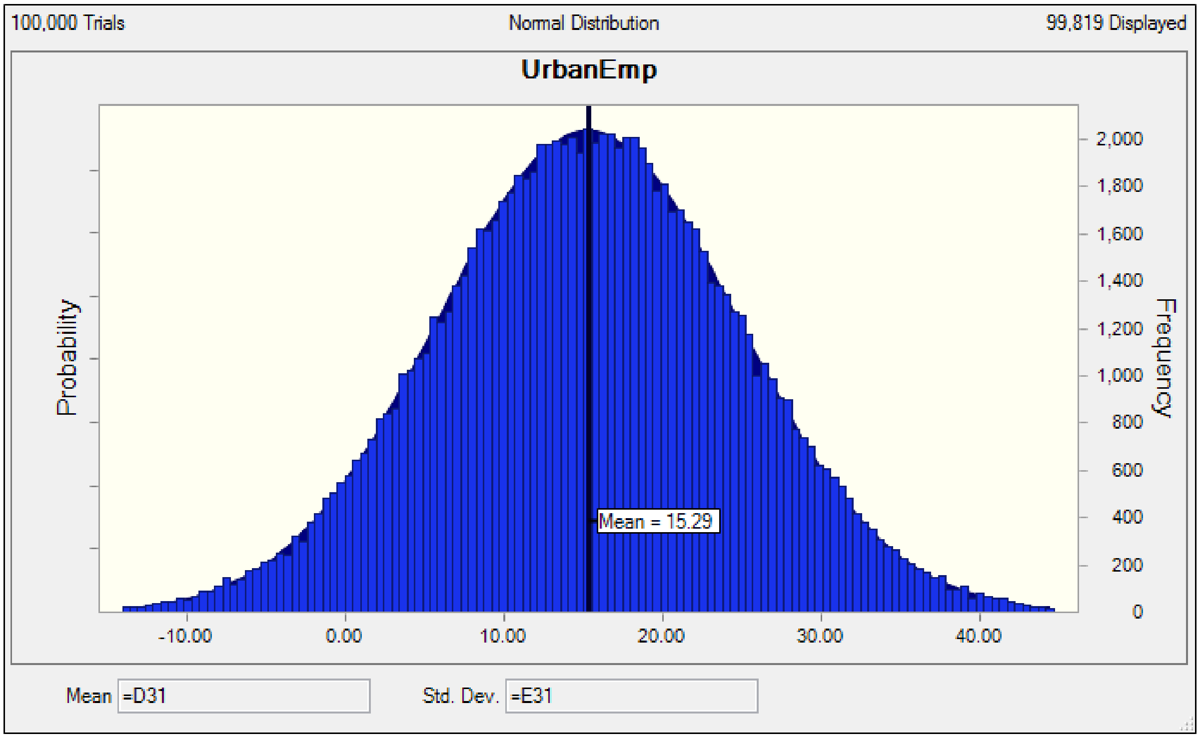
Source: FHWA

Figure 27. Toledo model 2045 urban core population growth rate distribution (2045 vs 2010).



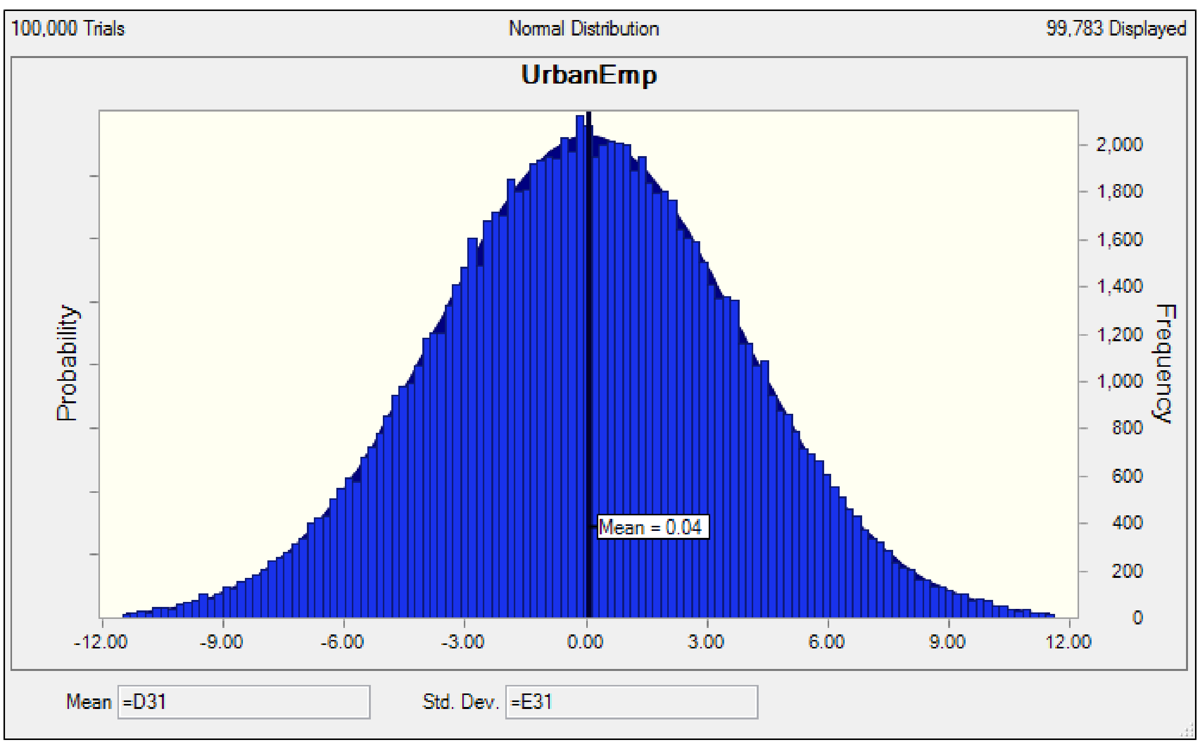
Source: FHWA

Figure 28. Chattanooga model 2045 urban core population growth rate distribution (2045 vs 2010).



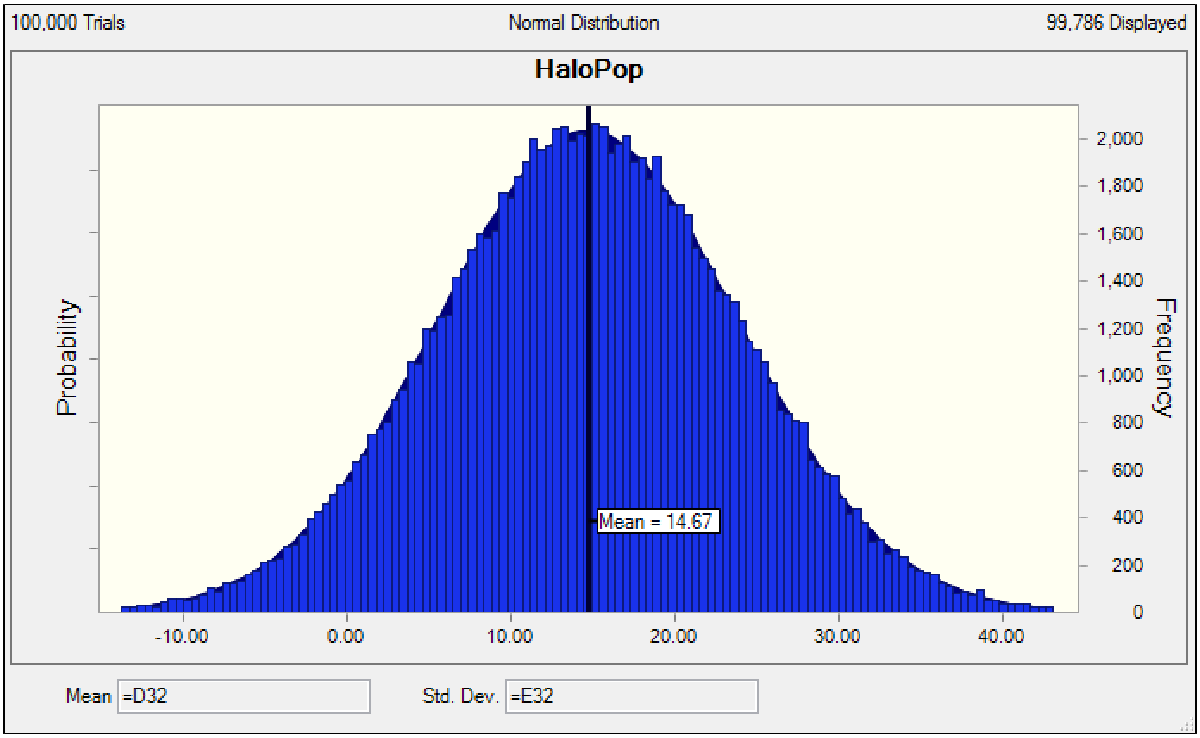
Source: FHWA

Figure 29. Toledo model 2045 urban core employment growth rate distribution (2045 vs 2010).



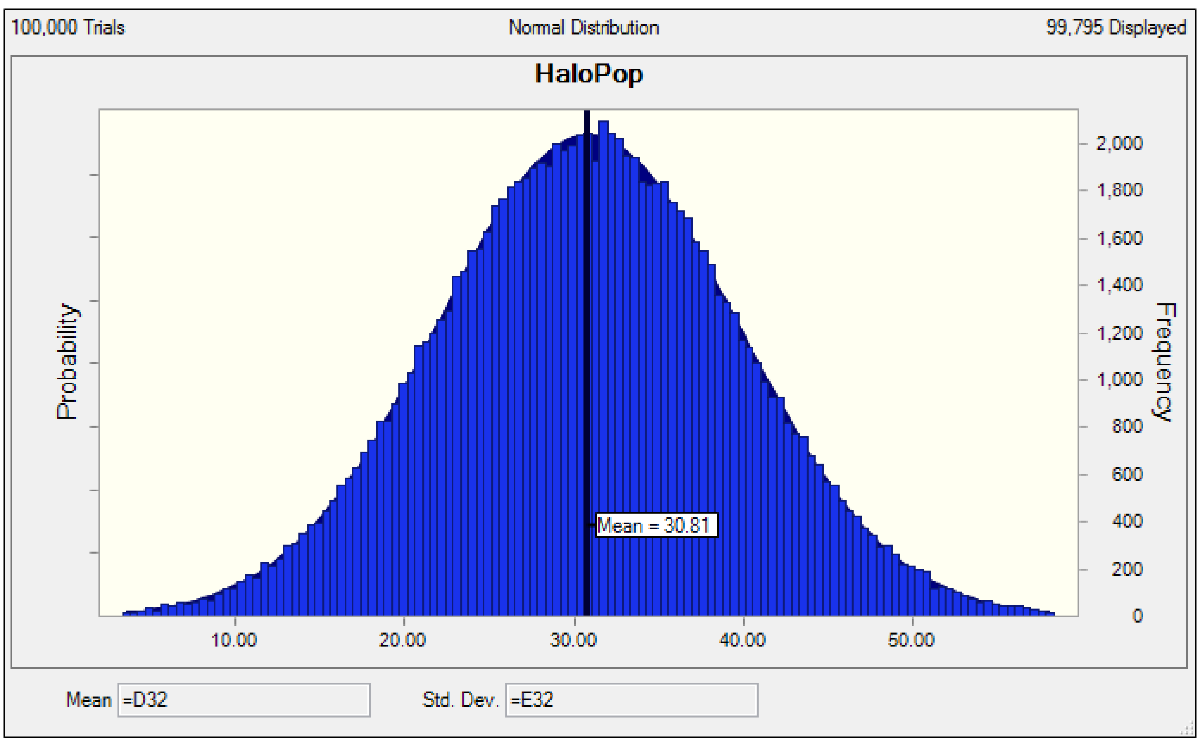
Source: FHWA

Figure 30. Chattanooga model 2045 urban core employment growth rate distribution (2045 vs 2010).



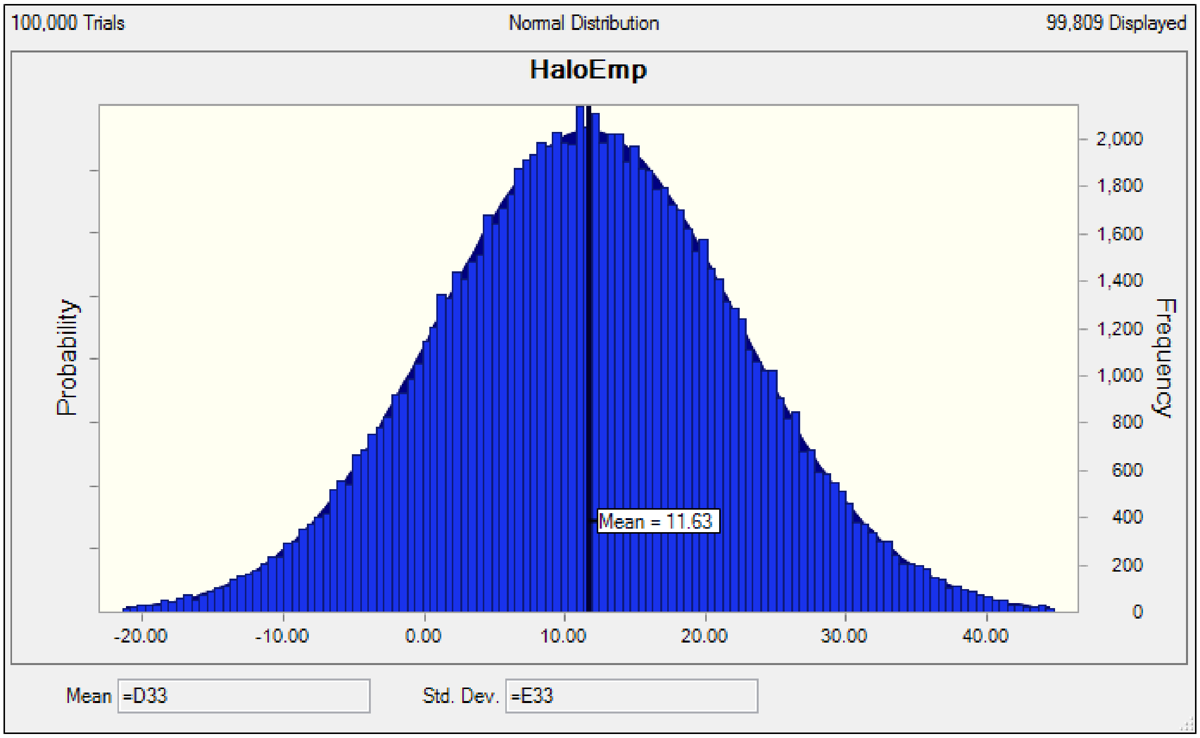
Source: FHWA

Figure 31. Toledo model 2045 halo population growth rate distribution (2045 vs 2010).



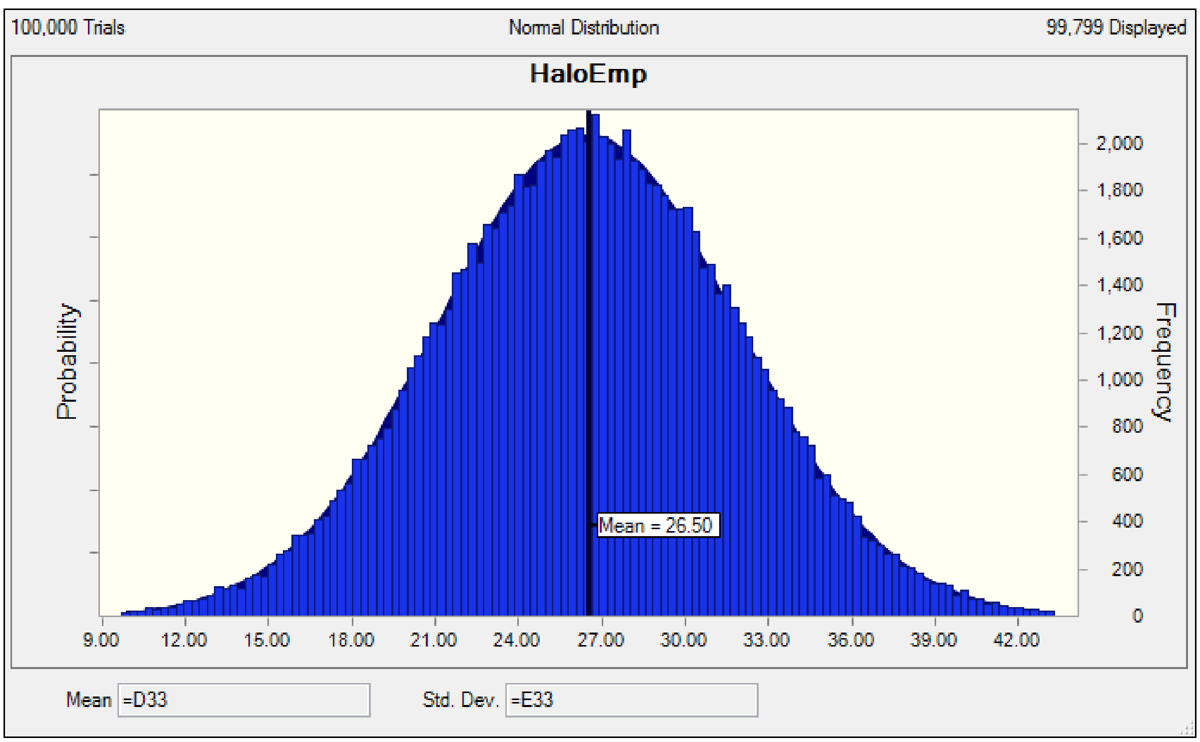
Source: FHWA

Figure 32. Chattanooga model 2045 rest of area population growth rate distribution (2045 vs 2010).



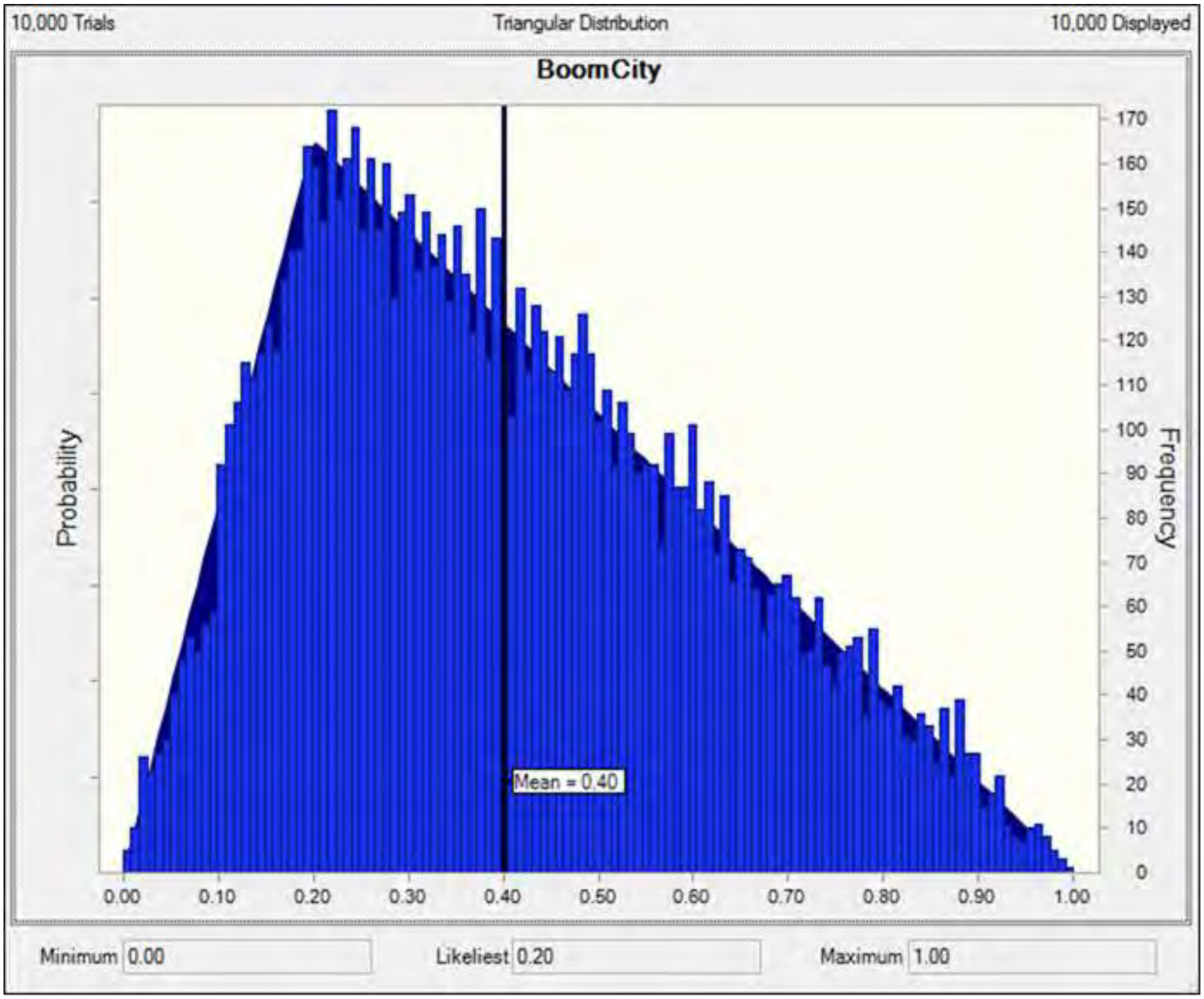
Source: FHWA

Figure 33. Toledo model 2045 halo employment growth rate distribution (2045 vs 2010).



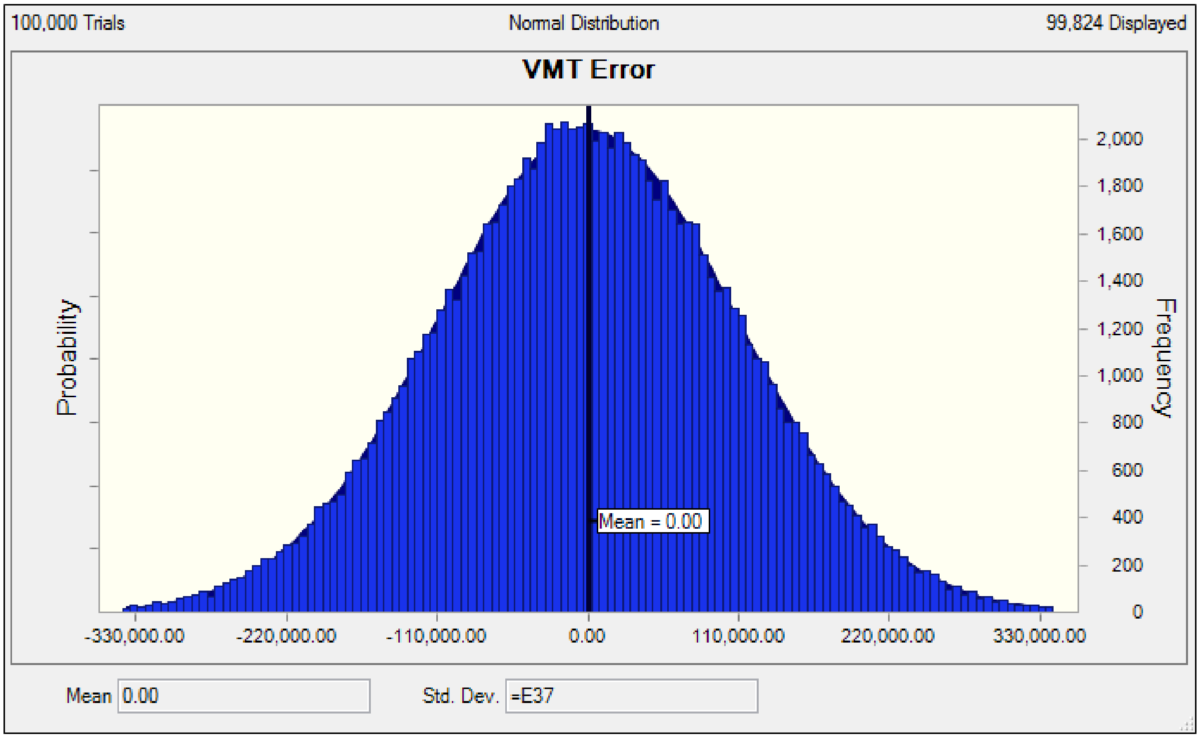
Source: FHWA

Figure 34. Chattanooga model 2045 halo employment growth rate distribution (2045 vs 2010).



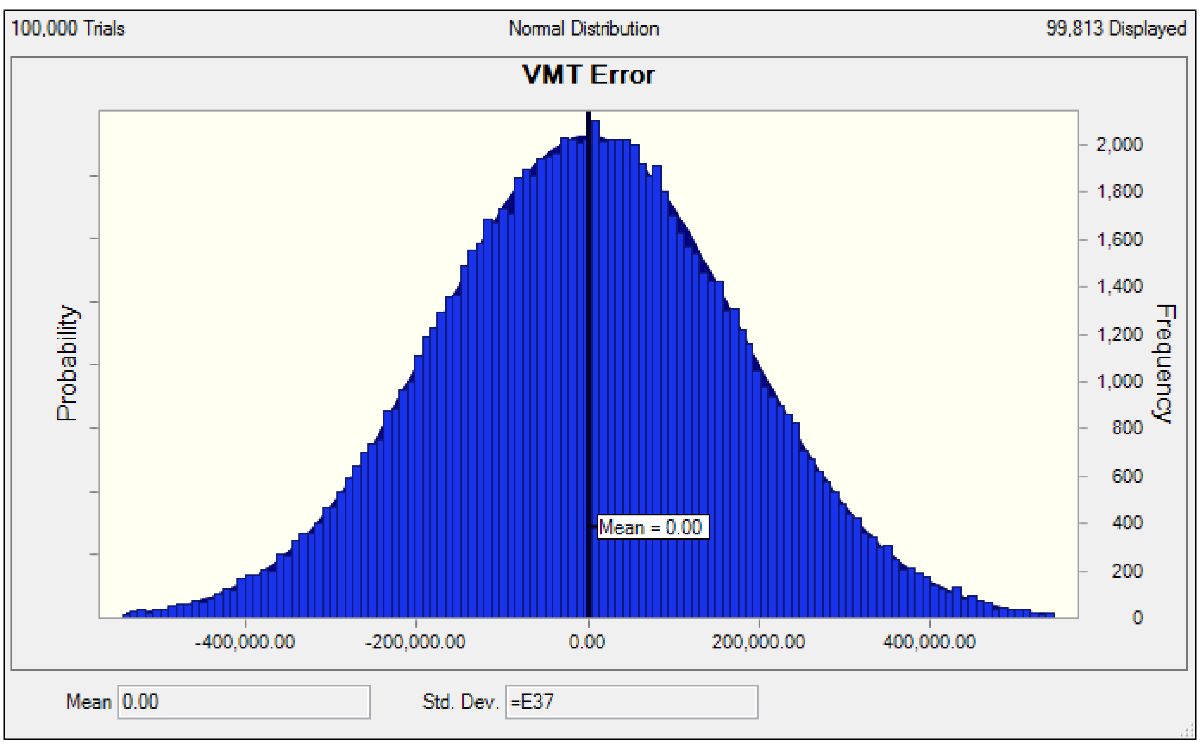
Source: FHWA

Figure 35. 2045 percentage of Boom City (Toledo and Chattanooga).



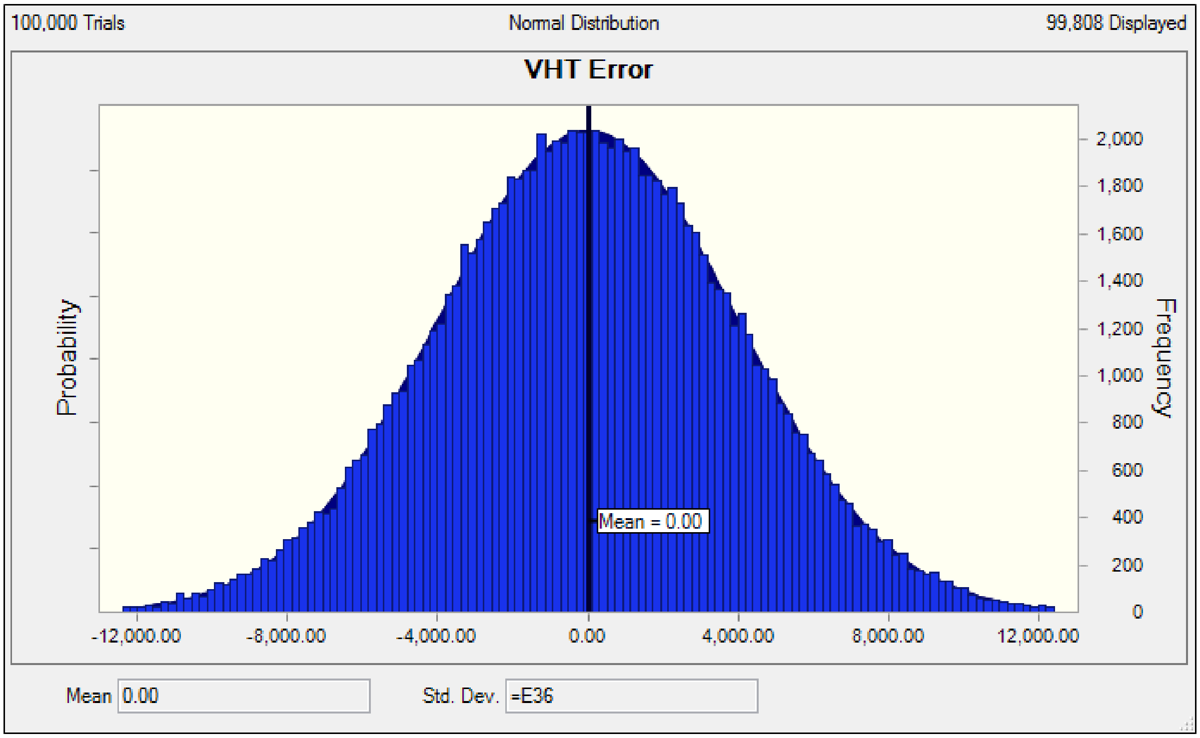
Source: FHWA

Figure 36. Toledo model squared error of prediction for VMT.



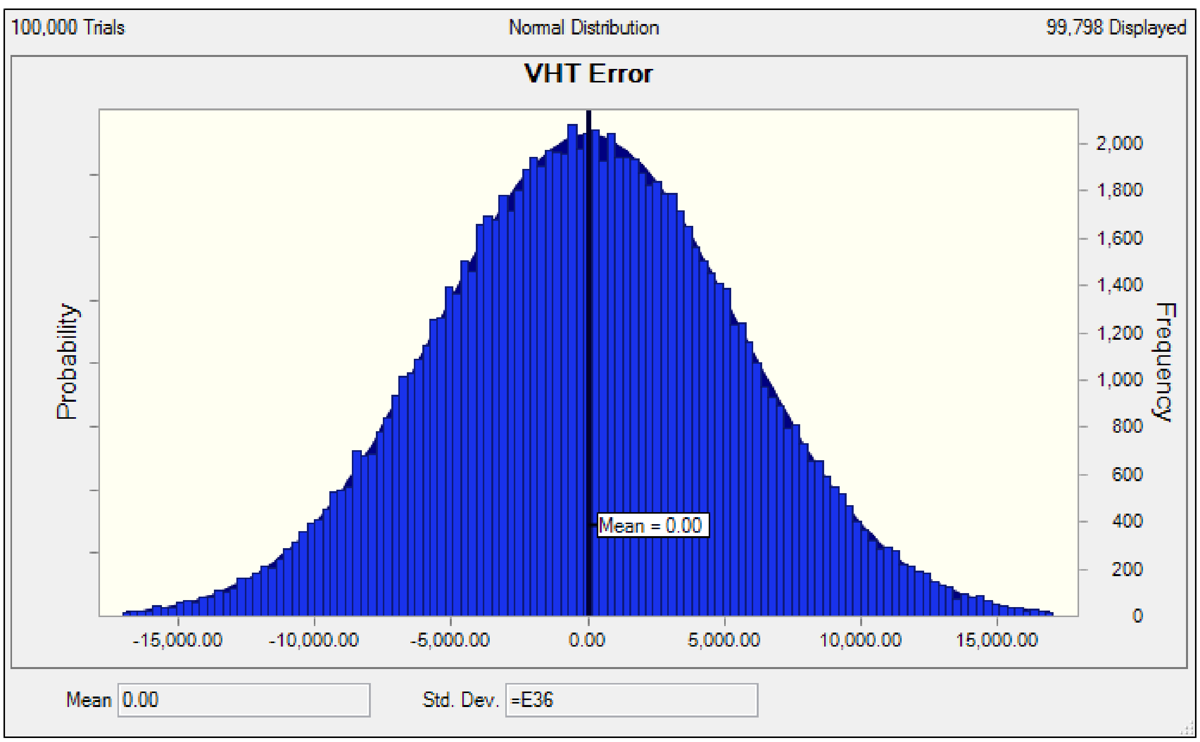
Source: FHWA

Figure 37. Chattanooga model squared error of prediction for VMT.



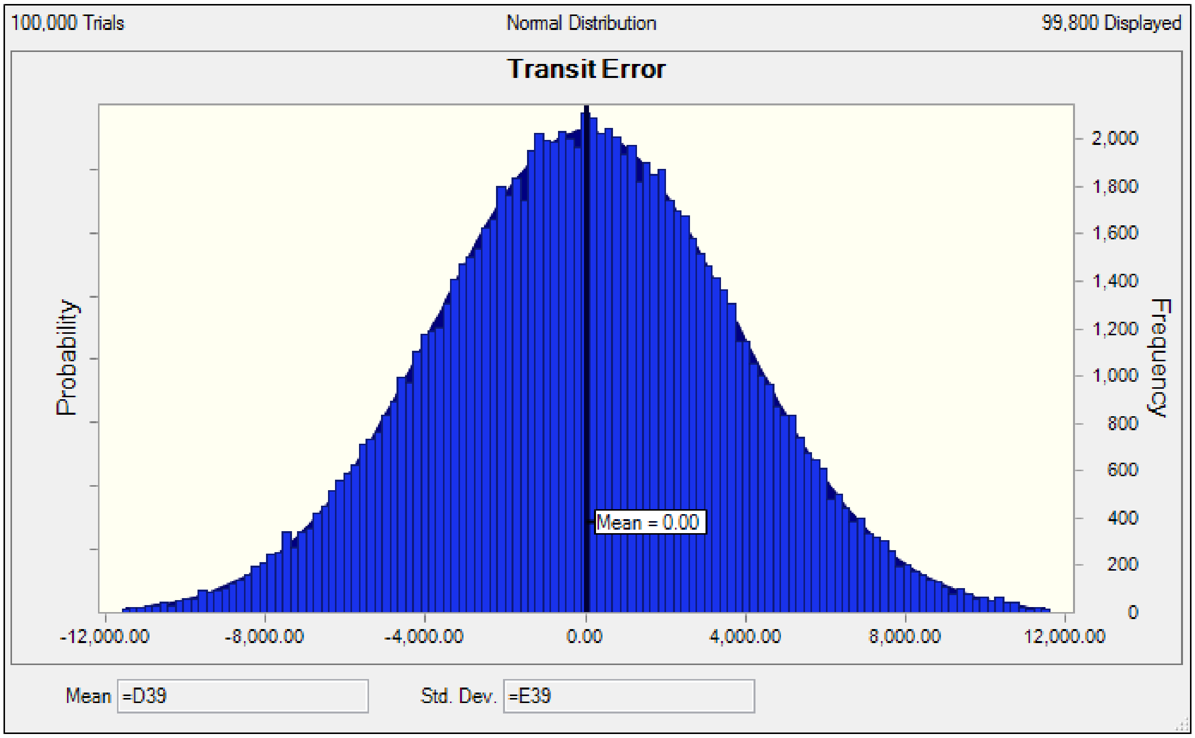
Source: FHWA

Figure 38. Toledo model squared error of prediction for delay.



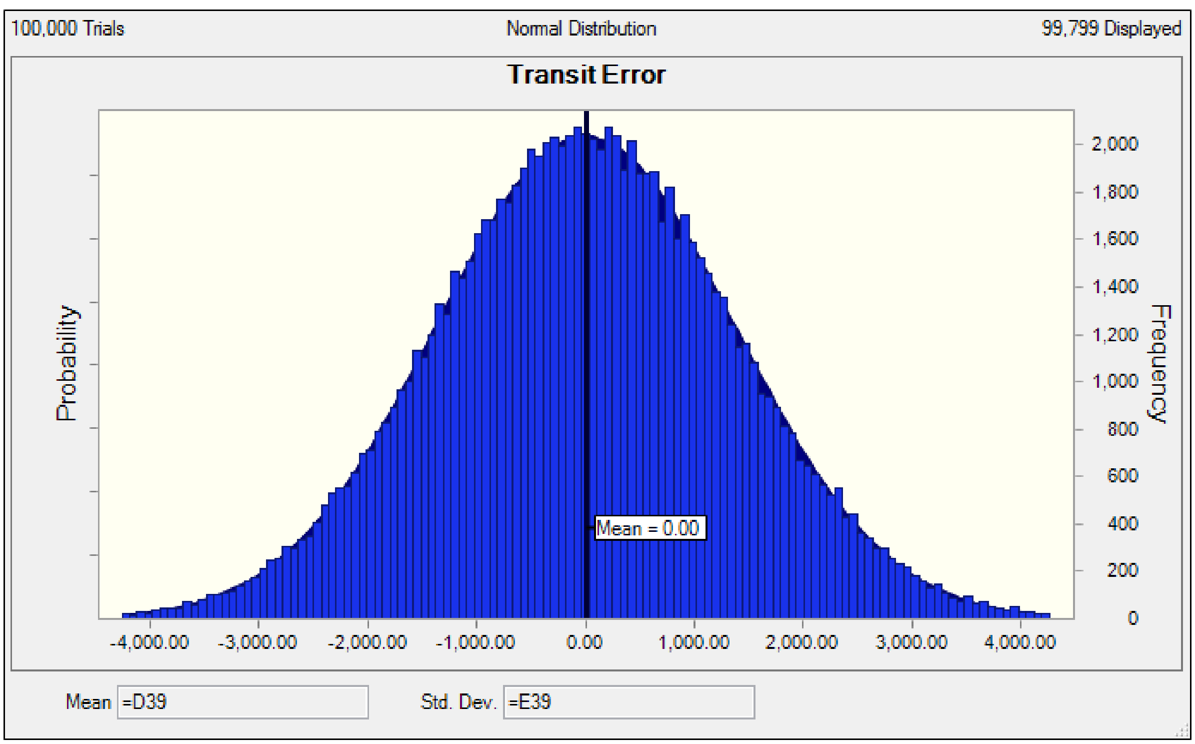
Source: FHWA

Figure 39. Chattanooga model squared error of prediction for delay.



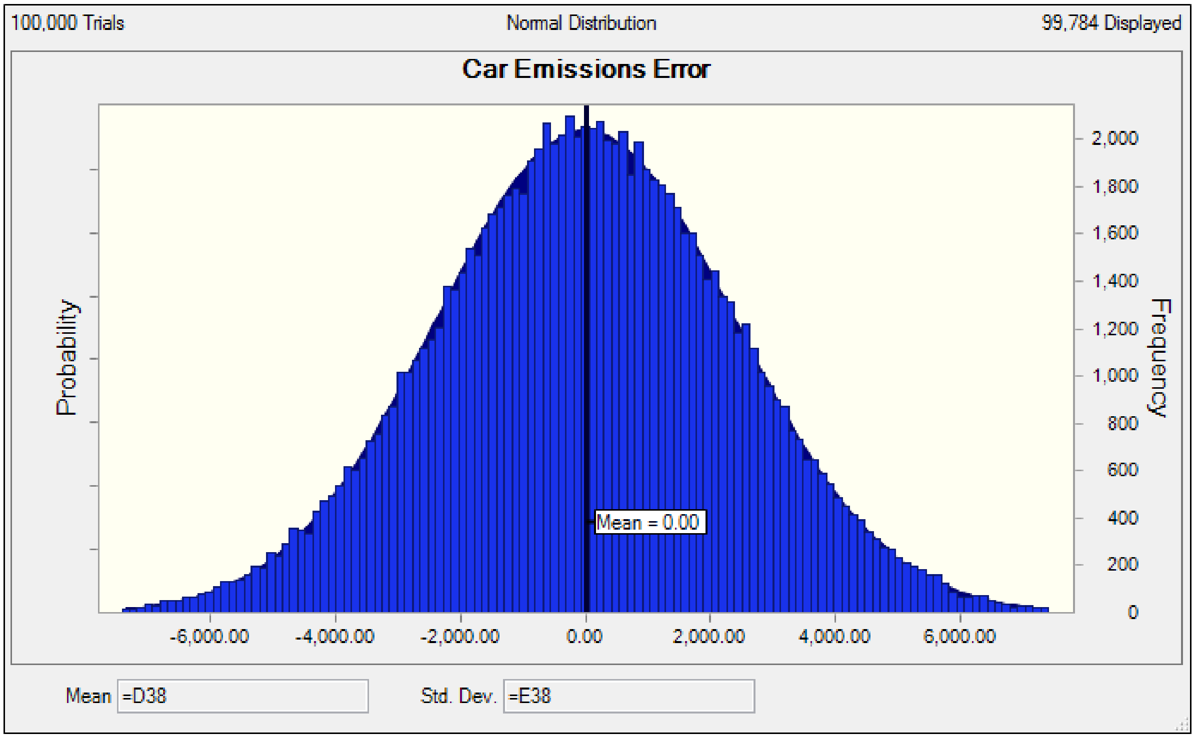
Source: FHWA

Figure 40. Toledo model squared error of prediction for transit ridership.



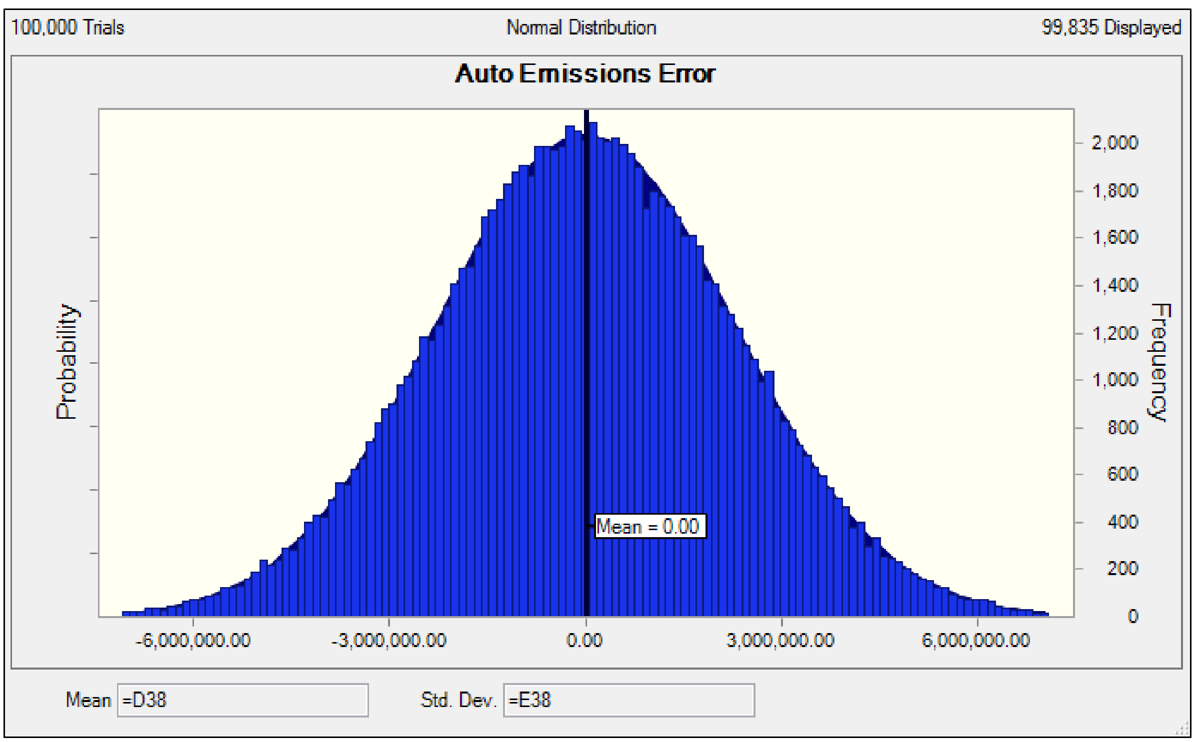
Source: FHWA

Figure 41. Chattanooga model squared error of prediction for transit ridership.



Source: FHWA

Figure 42. Toledo model squared error of prediction for auto emissions.



Source: FHWA

Figure 43. Chattanooga model squared error of prediction for auto emissions.

4.4 Results

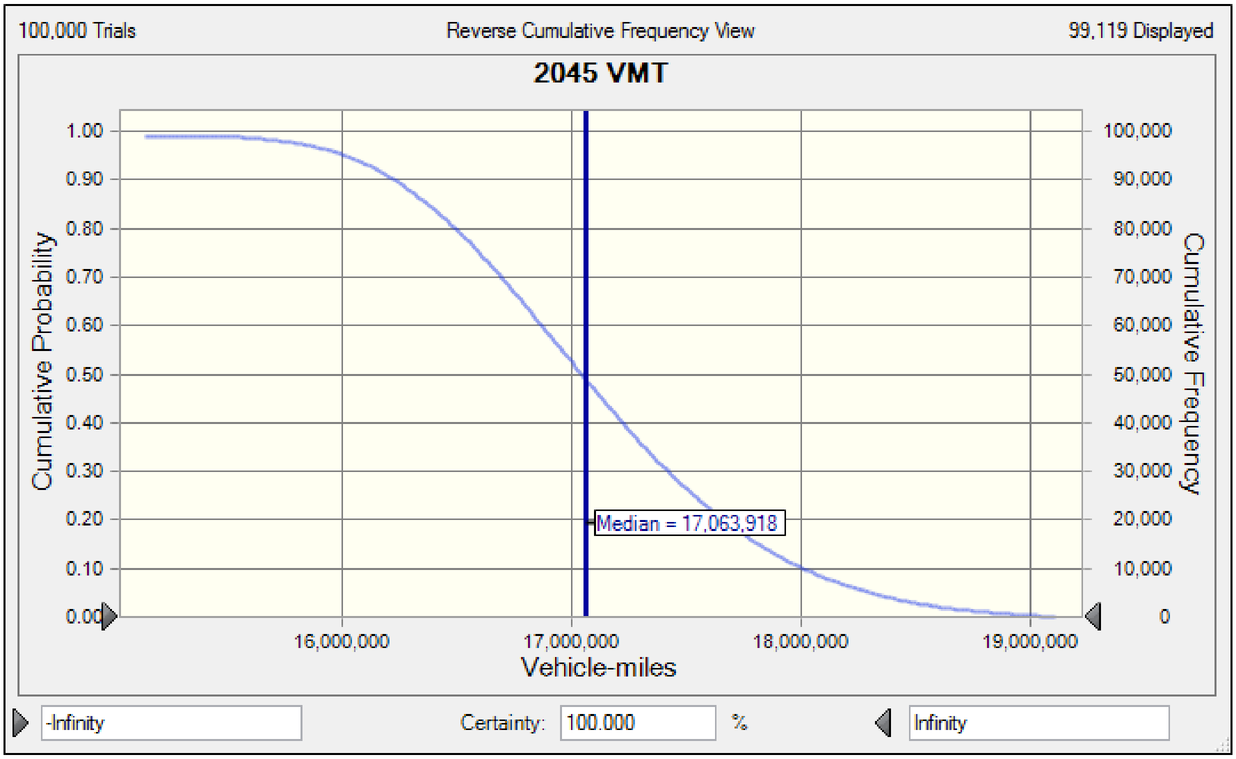
This section reviews the forecast results. A cumulative probability distribution and a multivariate sensitivity analysis, or variation decomposition, are presented for each forecast value

The cumulative probability distribution indicates the likelihood of the performance measure exceeding each value in its domain. For example, when the curve has a y-value of .5, there is a 50% chance the performance measure will exceed the associated x-value. The distribution can also be used to impute the probability of the x-value falling within a certain range.

The sensitivity diagrams, or variation decompositions, indicate the percentage of the forecast variation that can be attributed to each input or to the forecast error. The absolute value of each bar indicates its approximate contribution to the forecast variation, with negative (positive) bars implying the input is negatively (positively) associated with the forecast variable. Thus, each forecast variable is most sensitive to the longest bar. The forecast variable may be sensitive to a given input either because the input is inherently uncertain or because changes elicit strong model responses. An input is likely to be a dominant source of forecast variation when it is relatively uncertain and changes in its value elicit a sharp response. The diagrams also indicate the contribution of the error term to the forecast variation. The error term generally has a lower contribution when the model fit is better.

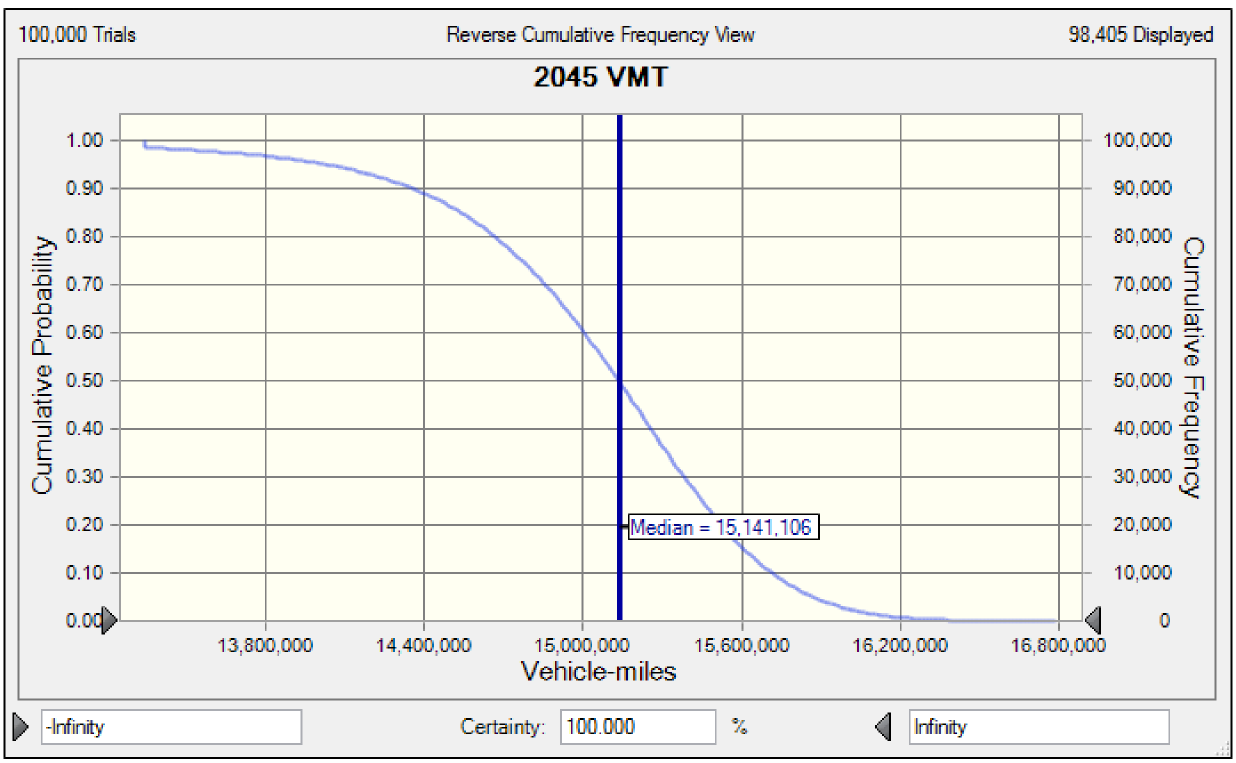
The Toledo and Chattanooga VMT cumulative probability distributions are shown in Figure 44 and Figure 45 respectively. The Toledo distribution suggests there is an 85% chance that 2045 VMT will between about 16,000,000 and 18,000,000, while the Chattanooga distribution suggests there is an 85% chance that 2045 VMT will between about 14,000,000 and 15,700,000. The Toledo forecast distribution is slightly more uncertain than the Chattanooga forecast distribution both in relative and in absolute terms.

There are several possible explanations for the greater Toledo forecast variation: the Toledo inputs, such as the 2045 population and employment growth, may be inherently more uncertain than the Chattanooga inputs; the Toledo model is more sensitive to its uncertain inputs; or due to inconsistency in the specification of either reduced form model. Lending credence to the third explanation is the fact that a large percentage of the variation in 2045 Toledo VMT is driven by the uncertainty in halo population (Figure 46) and the Toledo model, but not the Chattanooga model, includes a halo population squared term.



Source: FHWA

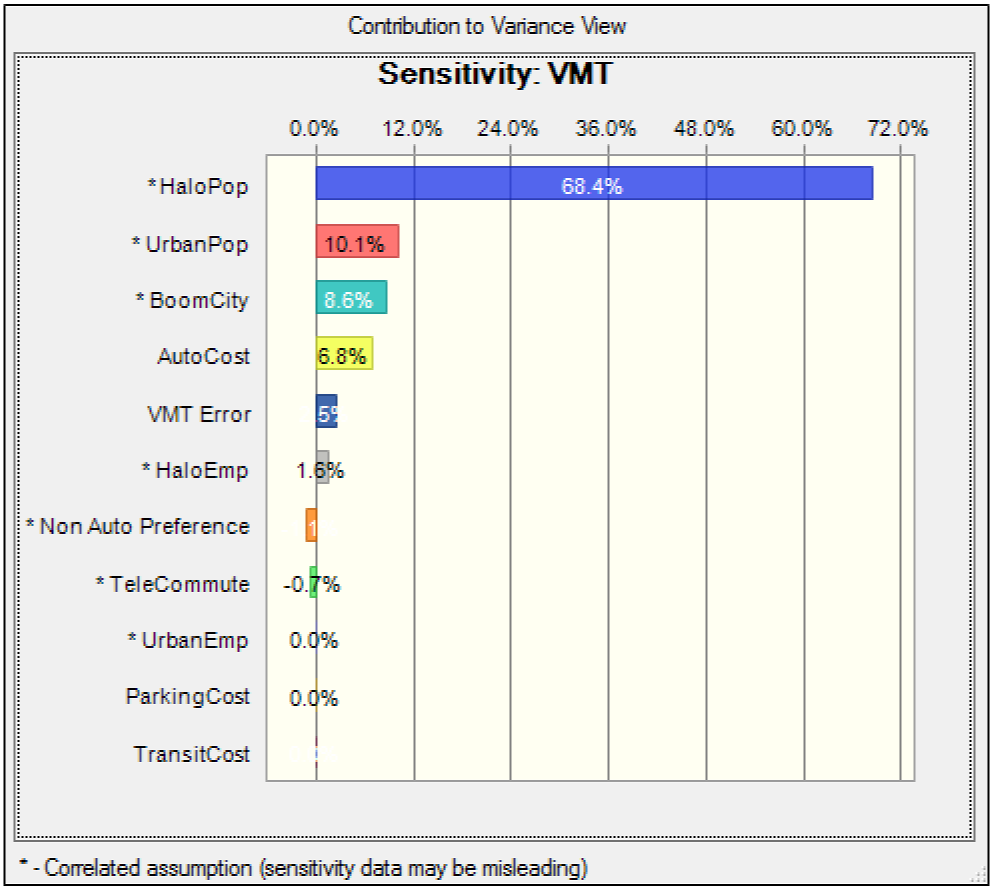
Figure 44. Cumulative probability distribution of 2045 VMT for Toledo.



Source: FHWA

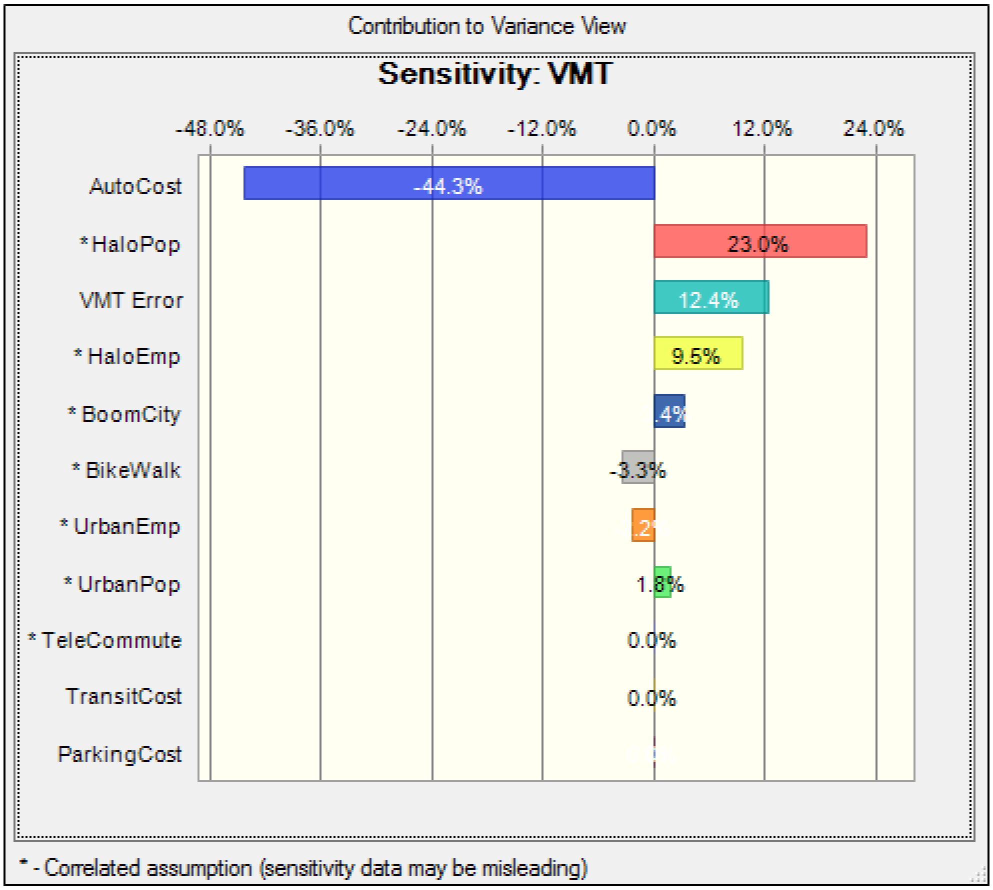
Figure 45. Cumulative probability distribution of 2045 VMT for Chattanooga.

Figure 46 and Figure 47 show the relative contribution of each model input to the variation, or uncertainty, in the 2045 VMT forecasts for Toledo and Chattanooga. For Toledo, much of the forecast variation is due exclusively to the halo population input, whereas for Chattanooga, fuel prices and halo population are each responsible for significant variation. Fuel cost uncertainty may be much less important for Toledo because this variable is only used in the mode choice step for work trips.



Source: FHWA

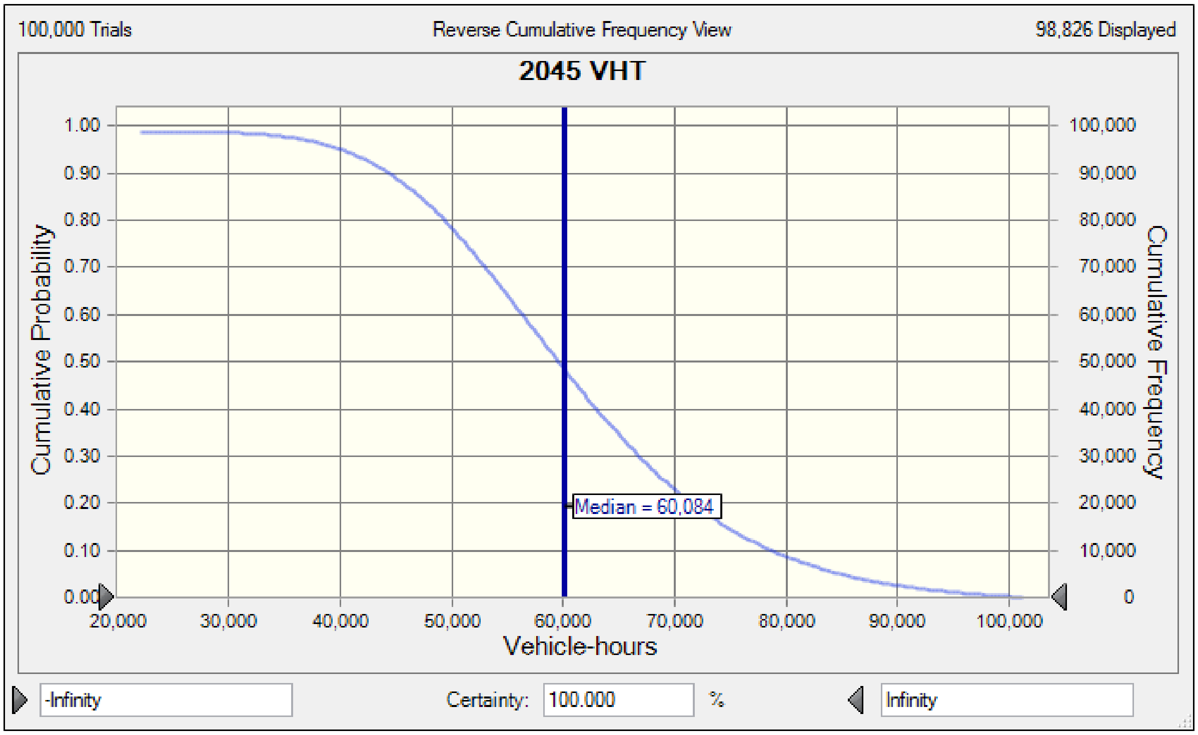
Figure 46. VMT sensitivity for Toledo.



Source: FHWA

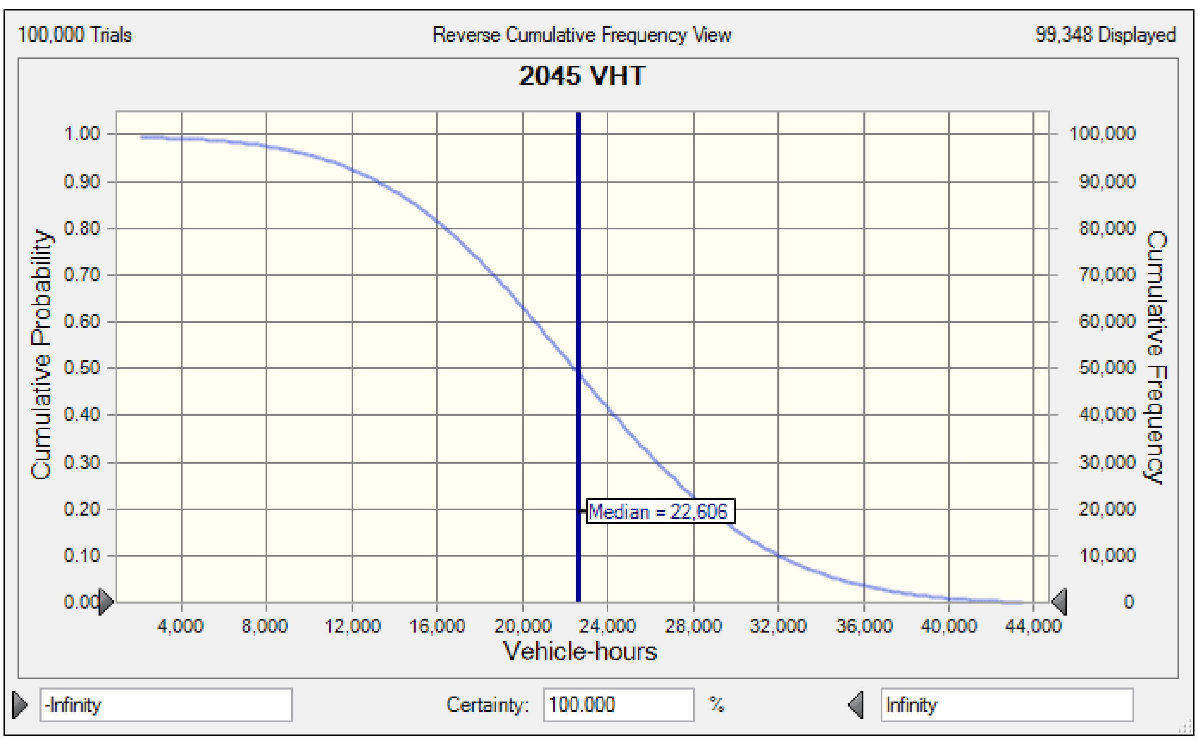
Figure 47. VMT sensitivity for Chattanooga.

The Toledo and Chattanooga delay cumulative probability distributions are shown in Figure 48 and Figure 49 respectively. The Toledo distribution indicates there is an 85% chance the 2045 delay will between about 40,000 and 80,000, while the Chattanooga distribution indicates there is about an 85% the 2045 delay will between about 10,000 and 32,000. This result suggests the 2045 Toledo distribution is more uncertain in absolute terms, while the Chattanooga distribution is more uncertain in relative terms.



Source: FHWA

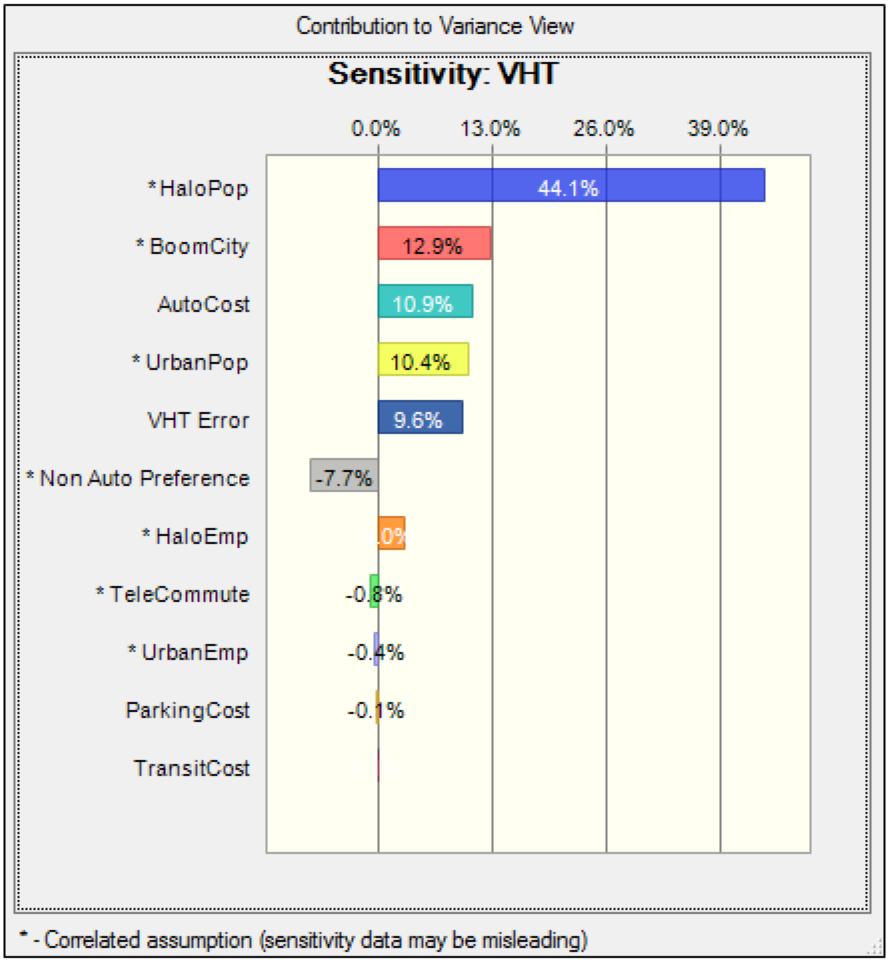
Figure 48. Cumulative probability distribution of 2045 delay for Toledo.



Source: FHWA

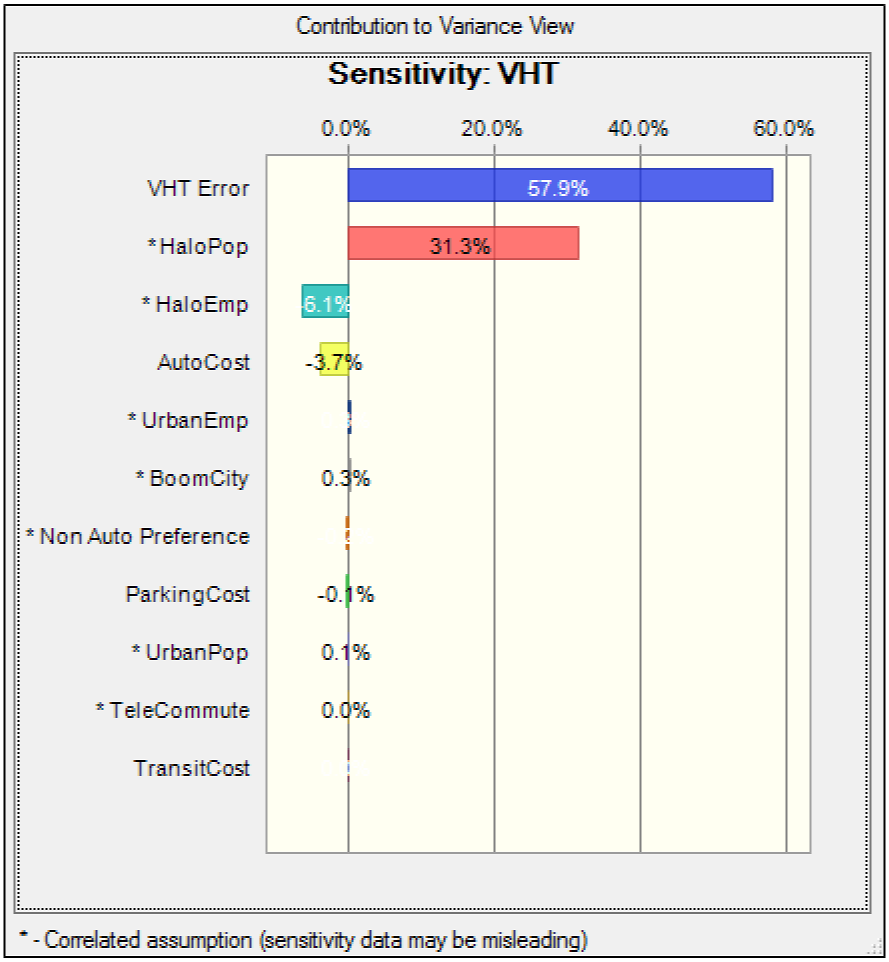
Figure 49. Cumulative probability distribution of 2045 delay for Chattanooga.

Figure 50 and Figure 51 show the relative contribution of each model input to the variation, or uncertainty, in the 2045 delay forecasts for Toledo and Chattanooga. For Toledo, much of the forecast variation is due to the population inputs and the existence (or absence) of Boom City. The 2045 Toledo delay is also sensitive to fuel costs and non-automobile preferences. For Chattanooga, the biggest source of delay variation, other than forecasting error, is the halo population growth rate. Adjusting the experimental design to include more factor levels or experiments may reduce the contribution of the error term for Chattanooga.



Source: FHWA

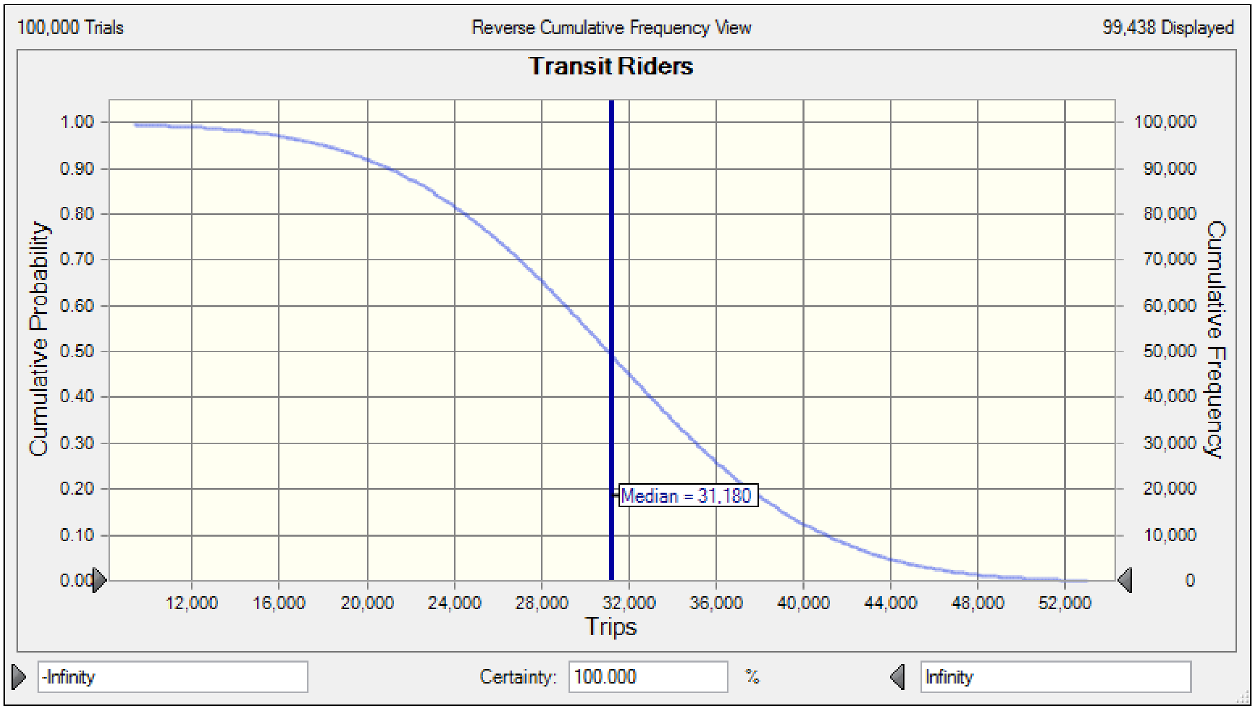
Figure 50. Delay sensitivity for Toledo.



Source: FHWA

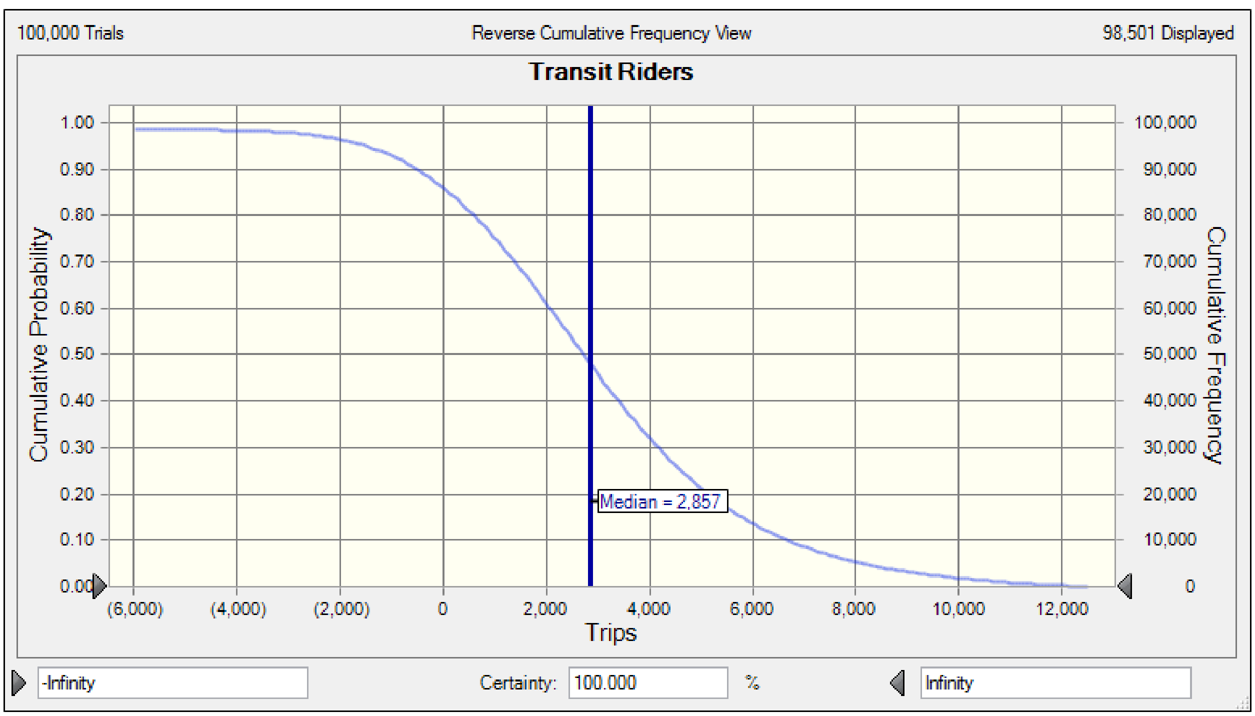
Figure 51. Delay sensitivity for Chattanooga.

The Toledo and Chattanooga transit ridership cumulative probability distributions are shown in Figure 52 and Figure 53 respectively. The Toledo distribution indicates there is an 80% chance the 2045 ridership will be between about 20,000 and 40,000 compared to actual base year ridership of 11,418. This distribution may be relatively wide since the Toledo transit share is small and even a one-tenth percentage point increase in the model share would yield an appreciable change in ridership. The Toledo model may have been overly sensitive to the assumptions of generational modal preferences and/or these assumed preferences may have interacted more strongly than anticipated with other scenario assumptions, which may or may not be realistic. The initial simulation indicated that there is an 80% chance the 2045 ridership will be between about 0 and 8,000, which is roughly the current ridership. It also indicated a 13% of negative ridership, which is clearly impossible. In practice, when facing this type of counter-intuitive result, an analyst may respond in a number of ways including modifying the experimental design to include more factor levels and experiments or by adjusting the reduced form model, such as by including a log transformation on key factors. However, for this example, the issue was addressed by running a second simulation with a different distributional assumption regarding generational modal preferences (refer to Section 2.3.6 for discussion of this assumption), assuming that they would be more in keeping with the national trend. The resulting second simulation shows an 80% probability that the transit ridership in Chattanooga is between 7,000 and 17,000, which seems far more plausible. This example both illustrates the importance of these distributional assumptions, but also the ease of changing them once a reduced form model is in place.



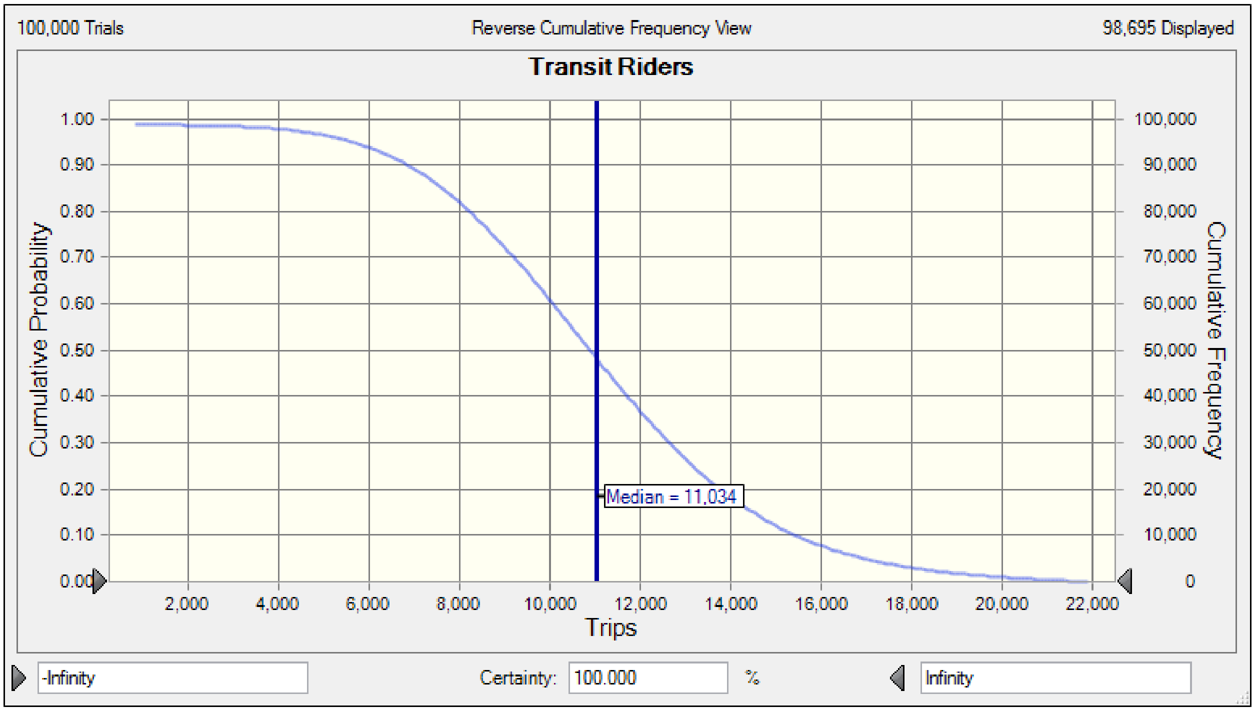
Source: FHWA

Figure 52. Cumulative probability distribution of 2045 transit ridership for Toledo.



Source: FHWA

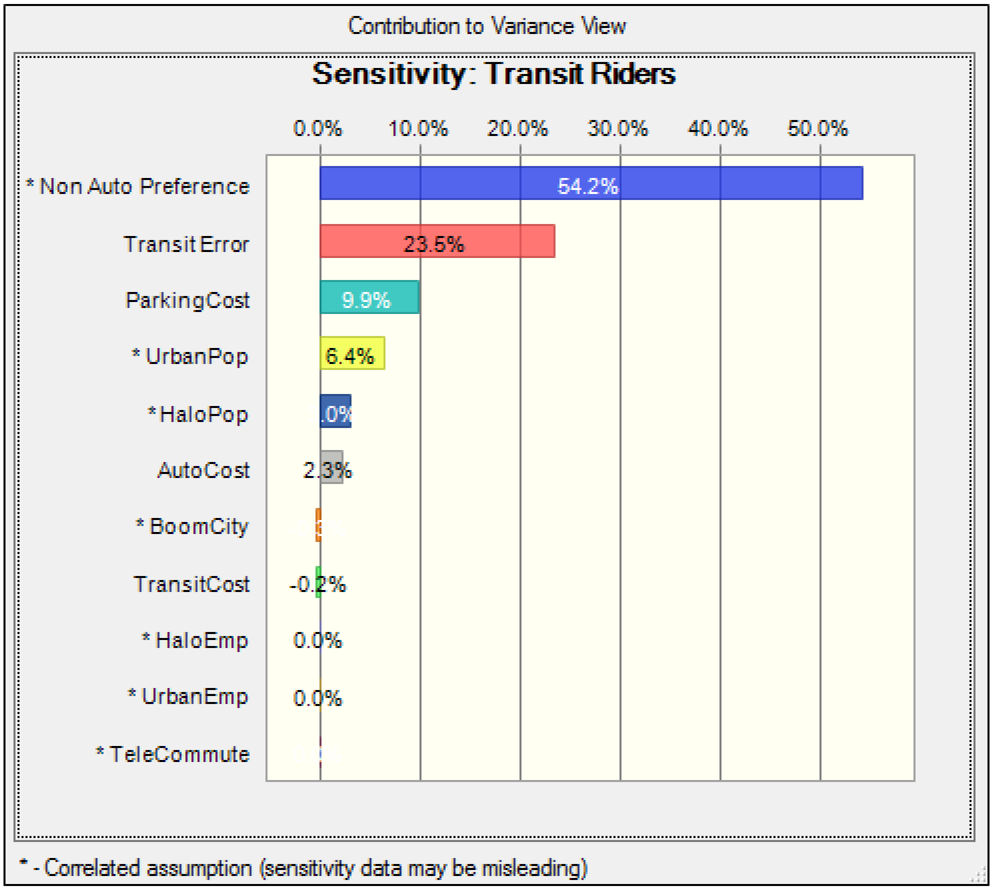
Figure 53. Cumulative probability distribution of 2045 transit ridership for Chattanooga (original).



Source: FHWA

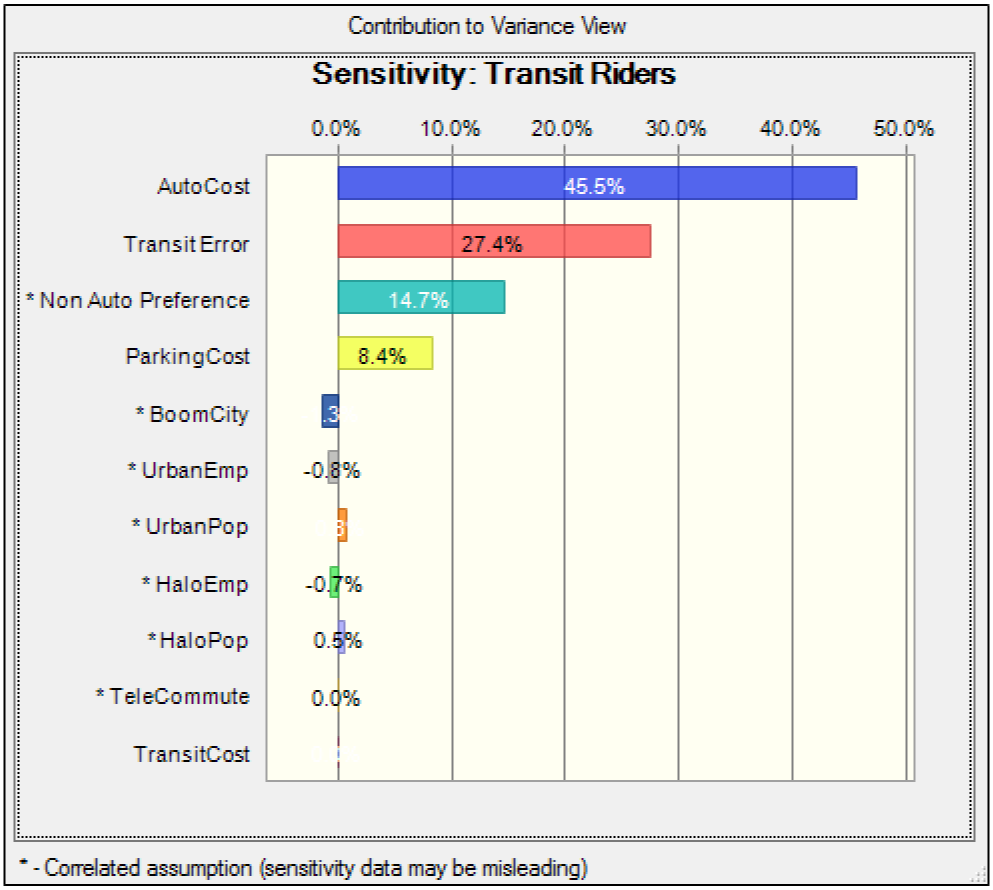
Figure 54. Cumulative probability distribution of 2045 transit ridership for Chattanooga (revised).

Figure 55 and Figure 56 show the relative contribution of each model input to the variation, or uncertainty, in the 2045 transit ridership forecasts for Toledo and Chattanooga. As expected, much of the variation in Toledo ridership is due to non-auto preference. Parking cost also contributes significantly to the Toledo forecast variation. For Chattanooga, auto cost contributes the most to transit ridership uncertainty. A potential explanation is that the right tail of the 2045 fuel cost distribution includes high prices that elicit strong substitution of transit for auto (see Figure 22).



Source: FHWA

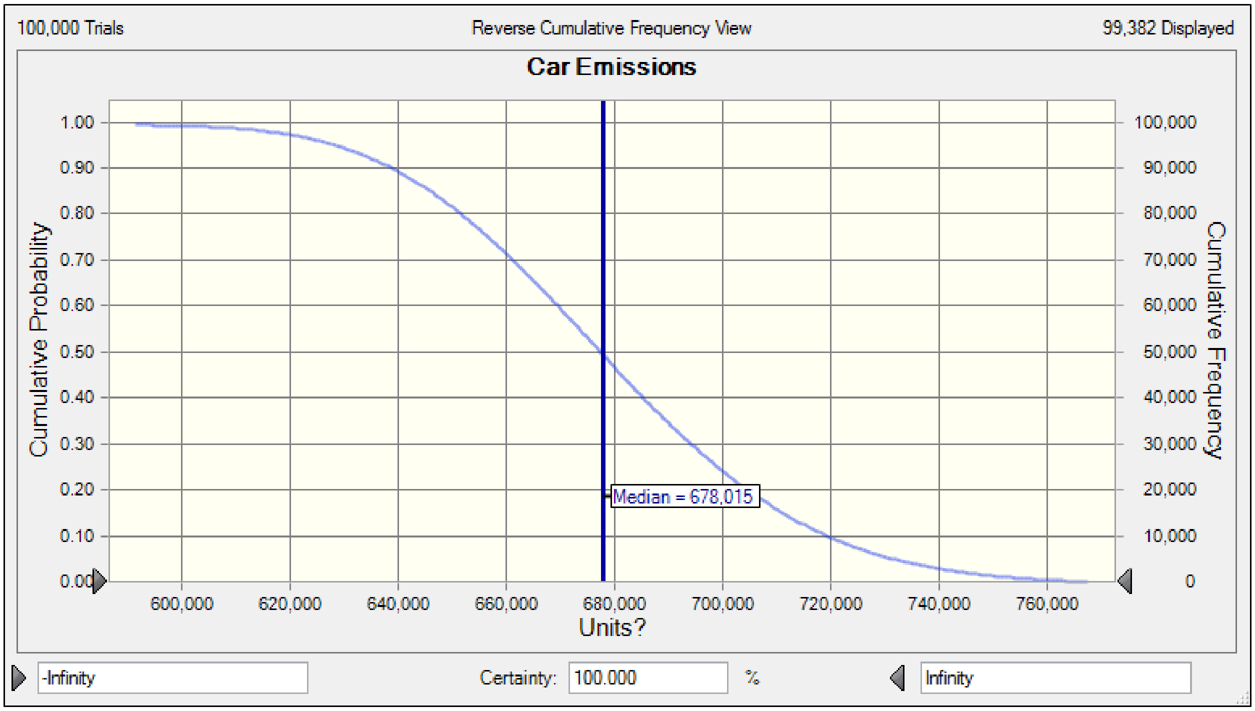
Figure 55. Transit ridership sensitivity for Toledo.



Source: FHWA

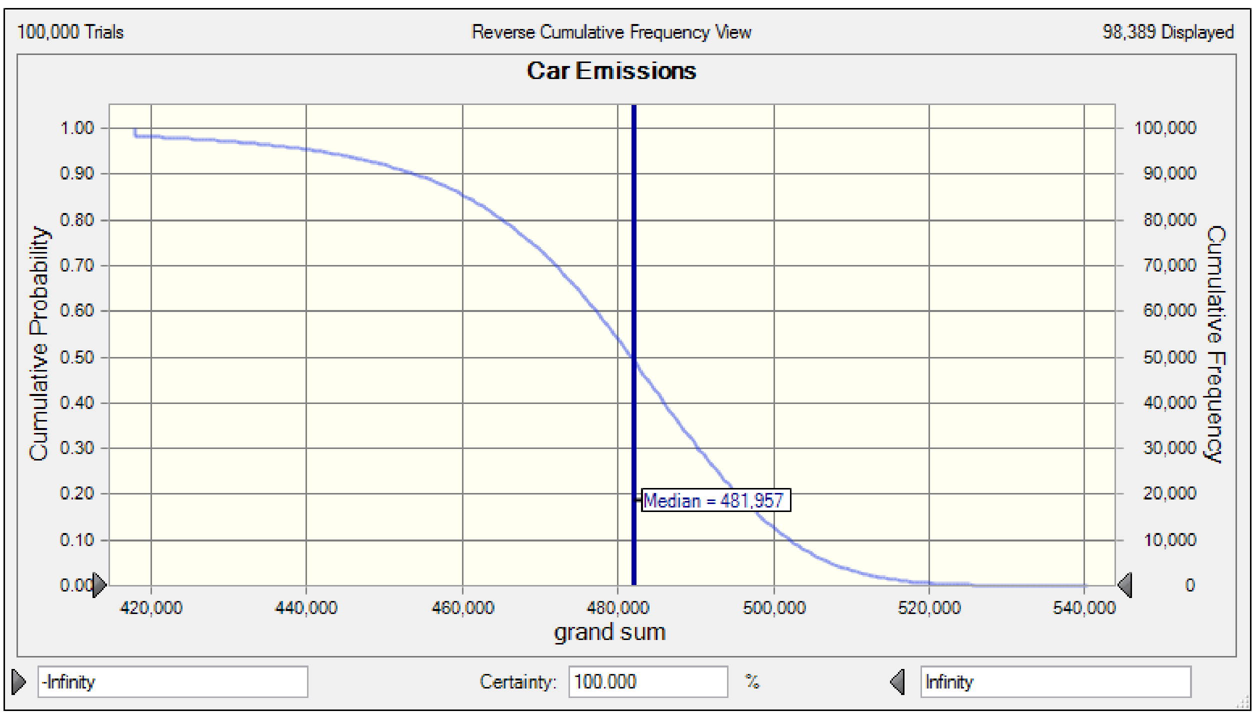
Figure 56. Transit ridership sensitivity for Chattanooga.

The Toledo and Chattanooga car emissions cumulative probability distributions are shown in Figure 57 and Figure 58 respectively. The Toledo distribution indicates there is an 80% chance of having between about 640,000 and 720,000 tons of car pollutants, while the Chattanooga distribution indicates there is 80% chance of having between about 440,000 and 500,000 tons of car pollutants. The relative amount of uncertainty is comparable between the two models.



Source: FHWA

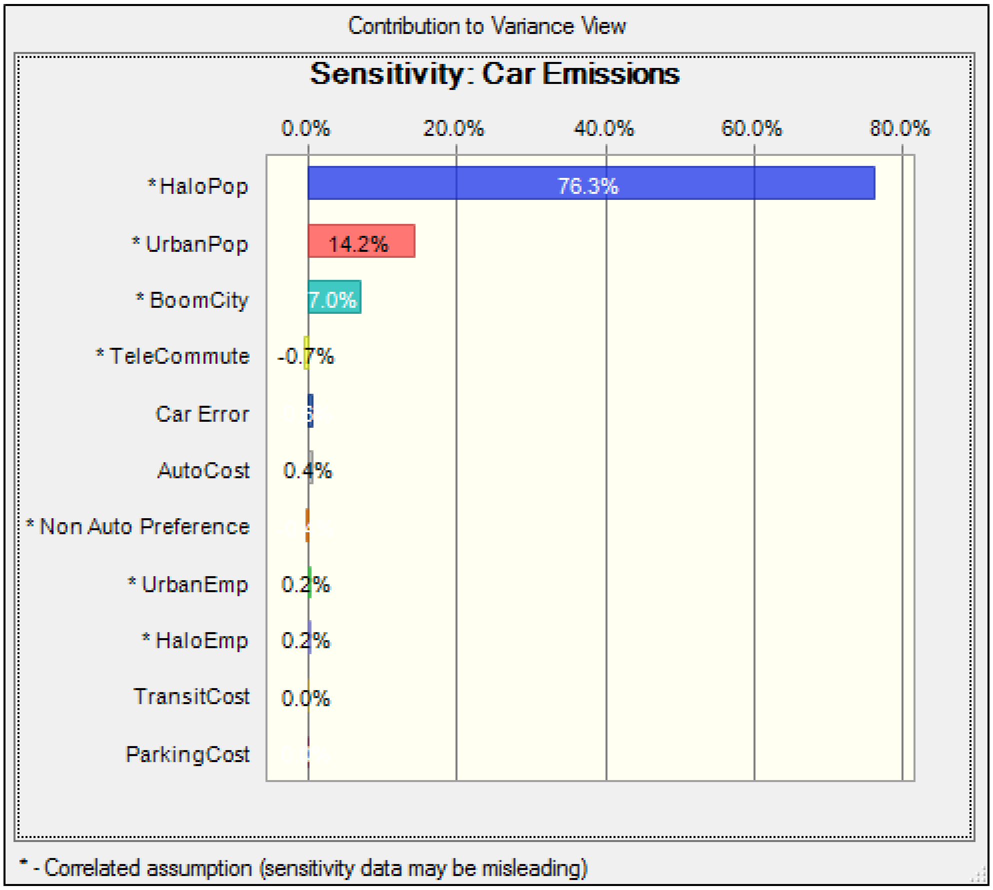
Figure 57. Cumulative probability distribution of 2045 car emissions for Toledo.



Source: FHWA

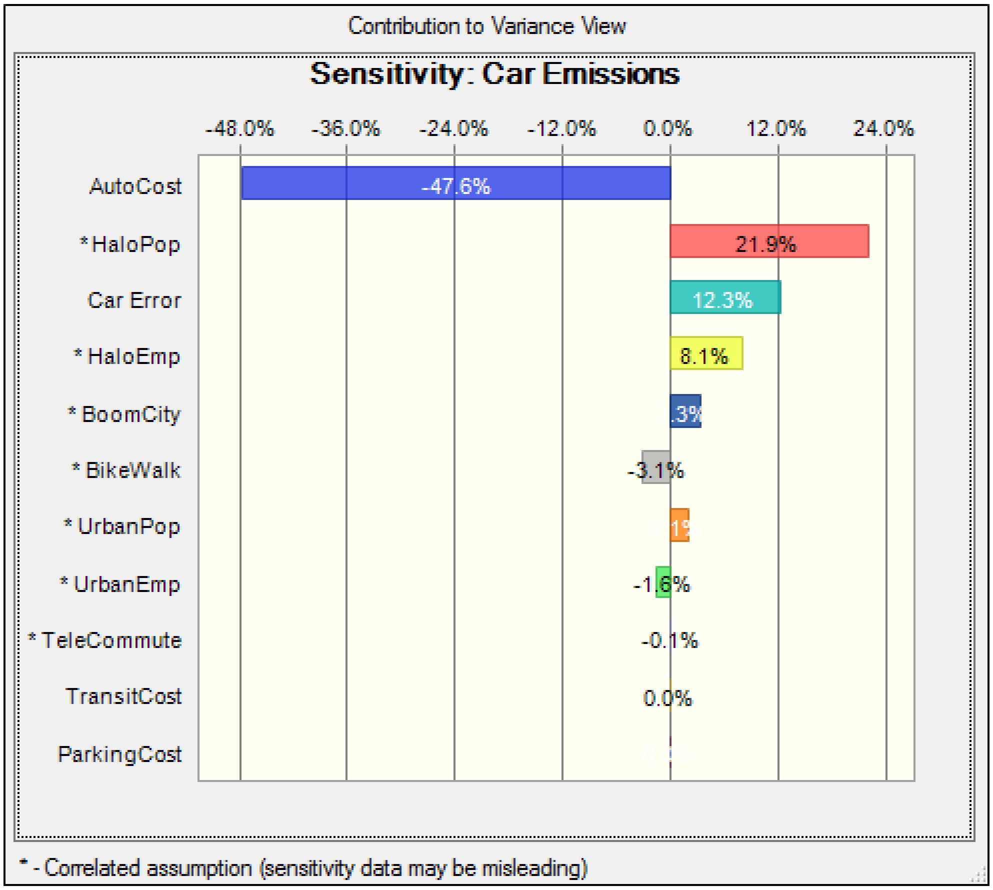
Figure 58. Cumulative probability distribution of 2045 car emissions for Chattanooga.

Figure 59 and Figure 60 show the relative contribution of each model input to the variation, or uncertainty, in the 2045 car emissions forecasts for Toledo and Chattanooga. Population growth rates and the existence (or absence) of Boom City causes much of the variation in the Toledo forecasts. Fuel price is the most significant source of variation in the Chattanooga model, which may be due to testing high 2045 fuel prices.



Source: FHWA

Figure 59. Car emissions sensitivity for Toledo.



Source: FHWA

Figure 60. Car emissions sensitivity for Chattanooga.

5.0 Discussion

This how-to guide demonstrates techniques for quantifying forecast uncertainty through two case studies. The case studies illustrated methods for defining input uncertainty distributions and estimating input-output relationships; these methods can be carefully integrated to yield forecast probability distributions. A 2045 trip-based model for Toledo, Ohio, was applied in the first case study, and a 2045 activity-based model for Chattanooga, Tennessee, was applied in the second case study.

This chapter summarizes the main lessons from the studies. It discusses the merit of the forecast methods and provides general observations on input-output relationships in travel models. Section 5.1 reviews the performance measures from the case studies, and section 5.2 compares the univariate and response surface methods.

5.1 Discussion of Example Results

Despite some of the limitations and simplifications of the examples presented here, they still help suggest some potentially interesting and useful insights into forecasting uncertainty in key performance measures.

The analyses suggest that there may be relatively little uncertainty in regional VMT, with an 80% chance of being within +/- 6% of a mean forecast. Auto emissions forecasts may be similarly reasonably precise. It is worth noting, that while total VMT may be relatively precise, if instead normalized by considering the VMT growth versus the base condition, the VMT growth would exhibit much higher uncertainty. Regions with higher growth therefore, could expect higher uncertainty in their total VMT as well.

In contrast, the analysis suggests there is much greater uncertainty in delay and transit ridership. The relatively greater uncertainty in delay relative to VMT may be in part due to the fact that these travel models – like the vast majority of travel models – were validated primarily against observed traffic volumes, not speeds. The relative uncertainty in transit ridership may in part be simply due to the small nature of the phenomenon in areas like Toledo and Chattanooga, but may also be in part due to uncertainty in modal preferences in younger generations and possibly also uncertainty in fuel costs.

The uncertainty in VMT and delay and auto emissions all are more associated with uncertainty in the amount of dispersed growth in suburban and rural areas. Chattanooga's activity-based model also indicates the importance of the uncertainty in fuel prices, while the same is not true of Toledo's trip-based model. It is important to consider that Toledo's model may understate the uncertainty in this regard, while the Chattanooga model overstate it (due to the limitations of the experimental design discussed).

It is also important to keep in mind the general limitations of the study in not considering factors such as uncertainty in freight growth and the impacts of automated vehicles. Even so, the results of the analysis can be useful for planning.

5.2 Conclusions

Univariates "sensitivity" analyses can be quick and effective, describing basic input and output relationships without requiring the analyst to estimate input probability distributions or conduct many model runs. Univariate testing should be considered for most model development and application work.

These tests generally do not require significant preparation and the basic input-output relationship can be estimated through simple or aggregate analyses. Calendar time is often one of the main obstacles to conducting sensitivity analyses since these tests are sometimes started late in the project cycle. Although trip-based models generally have a major run time advantage over activity-based models and support more sensitivity analysis, this was only marginally true for the Chattanooga model, which could be completed in two to three hours.

Beyond describing basic input-output relationships, sensitivity testing is an effective way to expose model errors and limitations. If an output is assumed to be closely related to an input but changing the input's value elicits an unexpected response, then this may be a symptom of a user error. For example, changing the Chattanooga parking cost initially elicited a small response, but then it was discovered that cost had been mistakenly coded in terms of cents instead of dollars. An unexpected response may also be a symptom of a model design limitation. While this type of finding can be frustrating, it is useful for identifying and prioritizing future model enhancements.

A key disadvantage of univariate sensitivity tests is that they cannot determine how the results would vary with simultaneous changes to multiple inputs. Response surface methods can quantify these types of multivariate relationships. As shown in Chapter 4, a multivariate reduced form equation can be estimated without conducting many more runs than were conducted for the sensitivity tests in Chapter 3. The experimental design for the reduced form model even included many of the same runs for the sensitivity tests; in general, this means that univariate and response surface methods should not be viewed as an either/or decision since the univariate work can often recycled for the multivariate work, if the modeling team feels more advanced analyses are needed. Aside from conducting more model runs, the main challenge to producing a reduced form model may be efficiently designing the experiment, which was discussed in section 4.1.

While reduced form models can be productively used to examine multivariate input-output relationships, their full potential is unlocked when the analyst can also define probability distributions for the uncertain inputs. Once these distributions are defined, the analyst can quantify the likelihood that the performance measures, or forecast variables, assume each value in their respective ranges. Unfortunately, quantifying input probability distributions is generally neither quick nor easy. The analyst must first identify the set of key uncertain inputs for the study and will need to evaluate whether it is time or cost prohibitive to include certain inputs. In some cases, the uncertain input distribution could be derived from historical data. In other cases, expert judgment may be needed to define the distribution.

It should be mentioned that defining input probability distributions can be a useful exercise even if the distributions are not ultimately used in a response surface method. The historical evidence suggests optimism biases are not uncommon in forecasting and having to carefully trace through the derivation of future year assumptions may lead modelers to consider a wider range of outcomes or simply different default assumptions for forecasting.

With both defined input probability distributions and reduced form equations for the forecast variables, the analyst can simulate the forecast probability distribution. This distribution may be very valuable to policy makers as, for example, it could be used to define the likelihood of collecting a certain amount of toll or fare box revenue. But even if the forecast probability distributions are not immediately useful to policy makers, they may still be very useful to modelers and planners. The analyst may see that distribution of certain performance measure is unexpectedly wide or narrow or that the variation in the performance measures depends heavily on an unexpected input. This type of result may expose a model limitation or error, or it may compel the analyst or planner to reconsider prior beliefs.

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United State Government assumes no liability for its contents or use thereof.

The United States Government does not endorse manufacturers or products. Trade names appear in the document only because they are essential to the content of the report.

The opinions expressed in this report belong to the authors and do not constitute an endorsement or recommendation by FHWA.

This report is being distributed through the Travel Model Improvement Program (TMIP).

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

April 2018

FHWA-HEP-20-021

U.S. Department of Transportation
Federal Highway Administration