Making a Model a Good Predictive Tool

November 2016

U.S. Department of Transportation
Federal Highway Administration

TMIP

Better Methods. Better Outcomes.

FHWA-HEP-17-048

U.S. Department of Transportation

Federal Highway Administration

TMIP

Better Methods. Better Outcomes.

<table>
<tr><td>1. Report No.
FHWA-HEP-17-048</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Making a Model a Good Predictive Tool</td><td colspan="2">5. Report Date
November 30, 2016</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Authors
Thomas F. Rossi, Cambridge Systematics, Inc.</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Cambridge Systematics, Inc.
100 CambridgePark Drive, Suite 400, Cambridge, MA 02140</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
DTFH61-13-F-00101</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
United States Department of Transportation
Federal Highway Administration
1200 New Jersey Ave. SE
Washington, DC 20590</td><td colspan="2">13. Type of Report and Period Covered
September 2013 to November 2016</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
HEPP-30</td></tr>
<tr><td colspan="4">15. Supplementary Notes
The project was managed by Task Manager for Federal Highway Administration, Sarah Sun, who provided technical directions. Special thanks to the Baltimore Metropolitan Council and the Ohio-Kentucky-Indiana Council; of Governments for providing the model files and data for the analyses in the study and for their assistance throughout the study.</td></tr>
<tr><td colspan="4">16. Abstract
A basic part of travel demand model validation is running the model for a "base year" and comparing the outputs to observed data. Sensitivity testing and temporal validation are also critical components of any model validation effort. In this project, two model versions for each of two metropolitan areas were run twice each: 1) the more recent version for its base year and a backcast scenario (for the base year of the earlier model version), and 2) the earlier version for its base year and a forecast scenario (for the base year of the later model version). Some lessons from the comparison of the model results were presented, demonstrating the difficulty that models can have in forecasting and reinforcing the need for accuracy checking in model input data.</td></tr>
<tr><td colspan="2">17. Key Words
Model validation, backcasting, forecasting</td><td colspan="2">18. Distribution Statement
No restrictions.</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
46</td><td>22. Price
N/A</td></tr>
</table>

This page is intentionally left blank

Making a Model a Good Predictive Tool

Original: **November 2016**
Final: **November 2016**

Prepared for:

Federal Highway Administration

This page is intentionally left blank

Table of Contents

List of Tables

1.0 Executive Summary

Travel models are used by planning agencies to estimate demand for short and long term planning scenarios and alternatives. While it is impossible to anticipate all of the change factors that may affect travel behavior over the period for which forecasts may be needed, it is important to ensure that the model is appropriately sensitive to those factors that can be expected to change over the range of values that might be expected.

A basic part of travel demand model validation is running the model for a “base year” and comparing the outputs to observed data such as traffic counts, travel times and speeds, transit ridership, and other measures of travel demand. While this type of check by itself is insufficient to declare a model validated, it provides a measure of confidence that the model reflects the amount of travel demand. Sensitivity testing and temporal validation are also critical components of any model validation effort.

The Federal Highway Administration (FHWA) undertook this project to provide information for agencies performing this part of the validation process, sometimes known as dynamic validation. This research was intended to produce useful data on which model components are most stable over time and their sensitivities to the factors affecting travel demand that vary over time. Cambridge Systematics, Inc. (CS) was retained to perform the work on this project.

1.1 Analysis Approach

The models for two U.S. metropolitan areas, Baltimore and Cincinnati, were chosen as case studies for this work. For each region, each of the current and previous model versions were run for the base year for that model, and for the base year of the other model. This means that the previous model was run for the base year and a forecast year, and the current model was run for the base year and a backcast year. The MPOs for the two regions, the Baltimore Metropolitan Council (BMC) and the Ohio-Kentucky-Indiana Regional Council of Governments (OKI), provided the study team with the files needed to run the models. The study team then performed the model runs for all scenarios, comparing the results to output files produced by the agencies and to documented results.

Questions examined during the evaluations included the following:

- How well does the model forecast/backcast for the scenario several years removed from the base year?
- Does the model perform appreciably better for forecasting or backcasting?
- Are there particular areas where the model performs better for forecasting/ backcasting? (These areas may be defined by geography, land use type, mode of travel, levels of congestion, time periods, and other segmentations available in the model.)
- Does the model show reasonable sensitivity to the factors that changes between scenarios?

1.2 Discussion of Model Results

A number of observations can be made regarding the results of the model runs performed in this study, concerning how the models performed in forecasting and backcasting:

- Some problems with travel demand forecasts in general have to do with inaccuracies in model inputs such as socioeconomic data.

- The OKI model underwent substantial changes between the development of their previous model and new model. The magnitude of the changes between the two models manifested itself in several ways in examining the results of the four scenarios (which also made it somewhat difficult to analyze the results).

- Not surprisingly, the results of the BMC model runs show that highway assignment results from base year model scenarios more accurately reflect observed traffic volumes than results from backcast or forecast year scenarios. Model fit depends on many factors, including model inputs, model structure, model parameters, and the specific mathematical formulations that form the model. The analyses show that differences in the features between model versions affect the model results. While there is no way to know the precise reasons for the underestimation of growth in travel demand in our analyses, some possible reasons might include the following:

    - Changes in travel demand for trips entering and leaving the model region are not adequately reflected.

    - The effects of model improvements made in the newer version may have compensated for calibration adjustments to the older model. This leads to an important hypothesis: It is likely that model calibration adjustments in the earlier model compensated for some of the deficiencies that the later model improvements addressed. This points out the risk involved in model calibration adjustments made to better reflect observed base year travel demand, which may not reflect the behavioral issues that caused the initial inaccuracies and may not carry forward in forecasting (or backward in backcasting).

    - The effects of calibration changes on model results are not well understood.

- The model results are very similar when using the same model for either version, for both analysis years.

- Model parameters that are constants may fail to reflect real changes related to the behavior being modeled by a particular component (for example, the use of fixed factors to convert daily travel to peak period travel).

- The BMC model results seem to hint that model components near the beginning of the model stream (e.g., trip generation) may be more stable in forecasting.

- Both BMC model versions had difficulty capturing trends in transit travel over the analysis period.

- It seems that, as a general rule, a model tends to produce similar and/or consistent results for different scenarios, but a new or updated model can produce substantially different results for a scenario from those for an earlier model version for the same scenario. It would seem to be a worthwhile effort when updating a model to consider the effects of the changes being made to the model.

1.3 Lessons and Recommendations for Modeling Practitioners

While it proved difficult to answer some of the questions posed above, some useful observations can be shared:

- Whenever a model is updated, the results will change, and this does not mean that either the original or updated model results should be considered incorrect.

- Undiscovered error in model inputs can have substantial effects on model results, reinforcing the need to thoroughly check model inputs such as network coding.

- When updating a travel model, it is natural for analysts to want to "fix" issues identified or suspected with previous model versions. All of the effects of the "fix" should be examined during the validation of the updated model.

- During model calibration, consider changes from estimated parameters only as needed to make the model a better forecasting tool. The effects of calibration changes in forecasting are unknown, and it is difficult to estimate these effects.

- It is important to recognize error propagation from upstream to downstream model components.

- When using fixed factors or constants in models, the sensitivity of model results to the factors that will vary between planning scenarios should be checked, as well as how the constants in models contribute to the lack of proper sensitivity.

The study produced the following recommendations for modelers:

1. Model validation should include temporal validation and sensitivity testing. Insights can be provided well beyond what can be learned only from comparisons of base year model results to observed data.

2. If possible, temporal model validation should include a backcast and/or a forecast year application.

3. Recognize that in a model update, changes in model procedures, assumptions, and input data can result in changes in model results that can go well beyond changes in travel behavior over time.

4. Model inputs, including networks and socioeconomic data, need to be thoroughly checked during model development and validation.

5. Whenever model parameters are changed or recalibrated during validation, the effects of these changes should be estimated if possible, and should be recognized in any case. Sensitivity tests can be structured to examine such effects.

6. Recognize the effects of error propagation from model components to subsequent components, and test for error propagation whenever possible. Understand that

downstream components may have more error associated with them than upstream components.

7. If it is necessary to use fixed factors or constants in models, recognize and test for the effects of model sensitivity of such factors. When using results of models that use fixed factors, recognize the limitations associated with insensitivity to factors that are not included in the models.

2.0 Introduction

2.1 Disclaimer

The views expressed in this document do not represent the opinions of FHWA and do not constitute an endorsement, recommendation or specification by FHWA. The document is based solely on the research conducted by Cambridge Systematics, Inc.

2.2 Acknowledgments

The FHWA would like to acknowledge the assistance of two metropolitan planning organizations (MPOs) who generously agreed to share their models and provide some of their time for this study:

- **Baltimore Metropolitan Council (BMC)** – The MPO for the Baltimore region, BMC has recently developed an activity based model, but the models used during this study included the previous trip based model, validated to a base year of 2008, and the previous model, validated to a base year of 2000.
- **Ohio-Kentucky-Indiana Regional Council of Governments (OKI)** – The MPO for the Cincinnati region, OKI maintains a state-of-the-practice four-step travel model for the region and has staff experienced in model development, validation, calibration, and application. The previous model was validated to a base year of 2005, and the current model has a base year of 2010.

2.3 Study Approach

2.3.1 Background

Travel models are used by planning agencies to estimate demand for short and long term planning scenarios and alternatives. While it is impossible to anticipate all of the change factors that may affect travel behavior over the period for which forecasts may be needed, it is important to ensure that the model is appropriately sensitive to those factors that can be expected to change over the range of values that might be expected.

A basic part of travel demand model validation is running the model for a “base year” and comparing the outputs to observed data such as traffic counts, travel times and speeds, transit ridership, and other measures of travel demand. While this type of check by itself is insufficient to declare a model validated, it provides a measure of confidence that the model reflects the amount of travel demand. Sensitivity testing and temporal validation are also critical components of any model validation effort. As noted in the Travel Model Validation and Reasonableness Checking Manual (Cambridge Systematics, Inc., 2010):

> “Sensitivity testing is the application of the models and the model set using alternative input data or assumptions…sensitivity testing should also include the application of the entire model set using alternative assumptions regarding the input demographic data, socioeconomic data, or transportation system to determine if the model results are plausible and reasonable…

> "Temporal validation is an important aspect of model validation since…it implies comparing model results to data not used in model estimation.
>
> "Most travel models are based on "snapshot" data…The model relationships, parameters, and coefficients might be significant and accurately reproduce travel for the point in time represented by the model estimation data. However, the relationships may not hold true over time; the further one moves from the base year for validation, the more uncertain one should be regarding the veracity of the models. For this reason, good validation practice should include temporal validation for at least one year other than the base year for model estimation or calibration."

The Federal Highway Administration (FHWA) undertook this project to provide information for agencies performing this part of the validation process, sometimes known as dynamic validation. This research was intended to produce useful data on which model components are most stable over time and their sensitivities to the factors affecting travel demand that vary over time. Cambridge Systematics, Inc. (CS) was retained to perform the work on this project.

2.3.2 Backcasting and Forecasting as Part of Model Validation

Many agencies' model validation efforts include "backcasts," where the model is run for a year prior to the base year of the model and compared to observed data from that year. Running the model for two points in time that may have similarities does not constitute a complete sensitivity test for the entire range of factors that the model may be used to analyze. However, the second point of reference besides the base year provides additional confidence that the model is reasonably sensitive to changes in conditions that affect travel demand.

While it is good practice to perform a backcast in addition to the base year validation, it can be challenging to assemble the necessary information for the backcasting process. Often, the backcast year was the base year for a previous version of the travel model; however, the model inputs to the previous version may no longer be valid for the updated model, for several reasons. For one thing, the input variables used in the updated model will likely be different; new variables may be incorporated, variable definitions (for example, income ranges) may change, and the reference period for cost measures (for example, 2010 dollars) may be updated. Additionally, zone boundary definitions often change when a model is updated, in response to development patterns and the capability for more detailed analysis afforded by increases in computing power. In many growing areas, the model region is expanded when the model is updated.

It can also be a challenge to obtain appropriate observed data to use in the validation checks. The types of model changes mentioned above affect the assembly of the validation data set. Additionally, it may be impossible to retroactively fill past gaps in data.

A further limitation of backcasting is that it addresses changes in travel demand that are generally in the opposite directions, temporally, from those that will be evident in forecasting applications. In most regions, travel demand is higher in forecast years than in the base year while demand is lower in a backcast year. While models should exhibit appropriate sensitivities for changes in both directions, changes in demand may not be symmetric with regard to changes in input data. In cases where there is low to moderate congestion in the base year, there will likely be relatively little congestion in the backcast year. As a result, the sensitivity of the model to increased congestion may not be visible from a backcasting exercise.

In regions where the base year is far enough in the past, it may be possible to use the model to forecast conditions for a point in time in the recent past after the base year, and to compare the model results to observed data for this point. This "short-term forecasting" approach has the following advantages relative to backcasting:

- The model features, including the model area, zone structure, model components, and variables, are consistent between the base year and "forecast" year;
- Data gaps can be more readily filled by obtaining new data, such as traffic counts, transit ridership counts, and recent census data, since the "forecast" year is in the very recent past; and
- Changes from the base year are in the "same direction" temporally as changes for forecast scenarios for which the model would be run (although probably lesser in magnitude).

FHWA's concept for this project was to perform short-term forecasts to help identify "what works" (i.e., does the model have the appropriate levels of sensitivity to input variables?) and what may not work as well. The work plan included both backcasting and short term forecasting. This provided some understanding of the sensitivity of the models to changes in both directions temporally as well as how much the advantages of forecasting checks affect validation, relative to backcasting checks.

2.4 Analysis Approach

As discussed in Section 2.3, the models for two U.S. metropolitan areas were chosen as case studies for this work. The testing plans for the two models were similar but reflected the differences between the modeling contexts in the two regions. For each region, each of the current and previous models were run for the base year for that model, and for the base year of the other model. This means that the earlier model was run for the base year and a forecast year, and the later model was run for the base year and a backcast year.

Each agency provided the study team with all of the files needed to run its model (except the proprietary model software licensed to the agencies, for Cambridge Systematics already had a valid license) which for all four scenarios, including the model input files, scripts, and executables. The study team then performed the model runs for all scenarios, comparing the results to output files produced by the agencies and to documented results (for example, Allen, 2006 and Baltimore Metropolitan Council, 2007). Observed data used for model validation, such as traffic counts, were also provided by the agencies. The details of the model runs for each region are presented in Chapter 3.

The analysis for each region consisted of the following elements:

- **Review of the Model Input Data** – As noted in the *Travel Model Validation and Reasonableness Checking Manual*, "A major concern for validation of travel models is error inherent in the collection of input data or historical data used for validation," and

"The success or failure of the modeling process rests on the input data." The input data sets for the base year and forecast/backcast scenarios were reviewed. Checks were made to ensure not only the quality of the data, but also consistency between the data sets for the different scenarios. This was necessary to ensure that the differences in the model results between the scenarios are due to the changes in conditions reflected in the scenarios and not due to data inconsistencies.

- **Definition of the measures of effectiveness** (MOE) to be used in the evaluations. FHWA defined three core MOEs:
  - Regional and subregional vehicle miles traveled (VMT) by facility type
  - Travel measures for traffic bottlenecks, including volume/capacity ratio (V/C)
  - Travel times (and speeds) on key corridors
  - Average trip lengths
  - Mode shares
  - Transit ridership

  The study team worked with the metropolitan planning organizations (MPOs) for the two regions to develop procedures to produce these measures for all scenarios (to some extent these procedures already existed).

- **Documentation of the observed validation data** for each model year. CS, with assistance from the MPOs, documented the observed validation data for each year, with particular attention to differences between the two model years. These differences included changes in highway volumes and VMT, transit ridership, travel times and delays, and travel patterns.

Each MPO prepared model input data sets, including socioeconomic data, highway and transit networks, and other required inputs, for the base year for each model. These served as the input data sets for the testing scenarios.

The testing plan for each model used the validated models for base years. The following scenarios were tested:

- Earlier model with base year model inputs
- Earlier model with forecast year model inputs
- Later model with (later) base year model inputs
- Later model with backcast year model inputs

There are six possible pairwise comparisons of these four scenarios. Five of these comparisons were performed:

- **The two base year scenarios** – This comparison provided information on the differences between the model results for the two years and in the model validation results.
- **Forecast to base year and base year to backcast year using the same model** – These comparisons provided information on differences in model results between the two model years.

- **"Same year" comparisons using the two models** – These comparisons provided information on differences between the validated model results and the forecasted (or backcasted) results for the same year.

We also compared the model results for each of the four scenarios to the observed data for the corresponding model year, to the extent data were available. Questions examined during the evaluations included the following:

- How well does the model forecast/backcast for the scenario several years removed from the base year?
- Does the model perform appreciably better for forecasting or backcasting?
- Are there particular areas where the model performs better for forecasting/ backcasting? (These areas may be defined by geography, land use type, mode of travel, levels of congestion, time periods, and other segmentations available in the model.)
- Does the model show reasonable sensitivity to the factors that changes between scenarios?

3.0 Summary of Model Results

This chapter presents the results of the various base year, forecast year, and backcast year model runs for both the BMC and OKI models. Section 3.1 presents the results for the BMC model, and Section 3.2 presents the results for the OKI model. Section 3.3 discusses the implications of the model results.

3.1 BMC Model Runs

3.1.1 Model Versions and Scenarios

Two versions of the BMC model were used for analysis (labeled Model A and Model B for convenience). It should be noted that the BMC model region includes not only the BMC planning region, consisting of the city of Baltimore and five surrounding counties, but also the District of Columbia and three Maryland Counties that are part of the planning region of the Metropolitan Washington Council of Governments (MWCOG). The two model versions are:

- **Model A –** This model was estimated with a base year of 2000. As a starting point, BMC used model parameters transferred from another area and used for calibration purposes the Baltimore Region 2001 National Household Travel Survey (NHTS) add-on survey and (for the Washington region) a household survey conducted by MWCOG in 1994. Other data used in model calibration included a transit on-board survey conducted in 1996, traffic count data from 2000, and transit route summary boarding counts for 2000.
- **Model B** – This model has a validation base year of 2008. The model included some new components and market segments, used new input data, and parameters updated using the Baltimore/Washington Region 2008 Household Travel Survey and 2008 transit on-board survey data. Key changes included revised volume-delay functions for highway assignment and the incorporation of a new toll diversion model.

The following scenarios were run using the two BMC model versions:

- **Scenario A00** – Model A run for the base year of 2000
- **Scenario A08** – Model A run for the forecast year of 2008
- **Scenario B08** – Model B run for the base year of 2008
- **Scenario B00** – Model B run for the backcast year of 2000

3.1.2 Socioeconomic Data Inputs

The socioeconomic data inputs are shown in Table 3-1 and are the same for each model year (2000 and 2008) in both Models A and B. Therefore, any differences between the results of the two models for the same analysis year are not due to differences in the model input socioeconomic data. Note that BMC had estimated a slight decline in population from 2000 to 2008 although an increase in the number of households was forecast, indicating a decline in average household size. Also note that a significant increase in employment from 2000 to 2008 was estimated, indicating an increase in workers who live outside the region.

Table 3-1 BMC Model Socioeconomic Data Input Summary

Measure	2000	2008	Percentage Difference
Households	1,891,211	1,973,108	4.3%
Population	4,964,091	4,956,580	-0.2%
Employment	2,934,952	3,479,179	18.5%

3.1.3 Highway Assignment Validation

BMC focused the highway assignment validation in both models on the BMC portion of the model region and noted that the MWCOG portion of the model region is underassigned relative to the traffic counts. The overall statistics presented here show relatively low overall volume/count ratios, but the ratios are better when only the BMC portion of the model region is considered.

Tables 3-2 and 3-3 show the modeled volume/count ratios by facility type and area type for the two base year scenarios (A00 and B08) respectively. Note that there are many more counts for 2008, which makes some of the comparisons between scenarios difficult.

The traffic volumes in Model A match the traffic counts slightly better on average than those in Model B. However, the match is better in Model A on freeways/ expressways, meaning that the match is better on the other roadway types in Model B. (Note that the facility type definitions are not exactly the same for the two models.)

Table 3-4 shows the modeled volume/count ratios by facility type and area type for the Model A forecast scenario (A08). Since these results represent a forecast to the year 2008, they can be compared to those of the Model B base year scenario (B08). As a comparison of Tables 3-3 and 3-4 shows, the forecast of 2008 from Model A produces a worse match to the traffic counts than Model B does for the base year. This is not surprising since Model B was validated to that year, and the Model B development process incorporated improvements to model components such as volume-delay functions.

Table 3-5 shows the modeled volume/count ratios by facility type and area type for the Model B backcast scenario (B00). Since these results represent a backcast to the year 2000, they can be compared to those of the Model A base year scenario (A00). The backcast of the year 2000 from the 2008 model produces a noticeably better match (on average, for the entire model region) with traffic counts for the year 2000 than in the Model A base year. However, since the MWCOG region is known to be underassigned, this implies that the BMC region is overassigned in Scenario B00, and so Model A does a better job matching 2000 traffic counts. Again, this is expected since Model A was validated to the year 2000.

Table 3-2 Volume and Count Summary by Facility Type and Area Type, Scenario A00

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway/Expressway	23,955,843	24,996,659	0.96	508
Primary Arterial	8,688,806	9,948,706	0.87	650
Minor Arterial	3,008,853	3,641,154	0.83	545
Collector	1,034,205	1,289,412	0.80	372
Centroid Connector	1,532,624	1,514,000	1.01	84
Other	133,804	162,700	0.82	19
Area Type[1]	**Volume**	**Count**	**Volume/Count**	**# of Links**
Area Type 1	12,768,989	13,147,948	0.97	923
Area Type 2	13,057,987	14,341,286	0.91	546
Area Type 3	8,325,397	9,248,652	0.90	385
Area Type 4	1,704,268	1,890,430	0.90	103
Area Type 5	1,174,121	1,317,141	0.89	86
Area Type 6	42,269	52,100	0.81	8
Area Type 7	750,789	931,761	0.81	76
Area Type 8	121,850	183,121	0.67	25
Area Type 9	408,465	440,192	0.93	26
Total	**38,354,135**	**41,552,631**	**0.92**	**2178**

[1] The BMC model defines area types objectively through a combination of residential and employment density ranges. Area type 1 represents the least dense areas while area type 9 represents the densest areas.

Table 3-3 Volume and Count Summary by Facility Type and Area Type, Scenario B08

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway/Expressway	30,410,381	33,920,562	0.90	622
Primary Arterial	22,284,161	22,101,227	1.01	1462
Minor Arterial	6,855,860	7,720,468	0.89	1073
Collector	2,233,674	2,486,036	0.90	672
Centroid Connector	1,468,457	1,474,516	1.00	84
Other	139,495	189,596	0.74	25
Area Type	**Volume**	**Count**	**Volume/Count**	**# of Links**
Area Type 1	18,781,776	18,721,381	1.00	1492
Area Type 2	18,830,637	20,228,701	0.93	916
Area Type 3	16,311,550	18,577,835	0.88	832
Area Type 4	4,860,844	5,317,038	0.91	306
Area Type 5	2,019,335	2,264,724	0.89	170
Area Type 6	326,056	271,169	1.20	24
Area Type 7	1,049,569	1,148,007	0.91	105
Area Type 8	300,395	389,799	0.77	33
Area Type 9	911,866	973,751	0.94	60
Total	**63,392,028**	**67,892,405**	**0.93**	**3938**

Table 3-4 Volume and Count Summary by Facility Type and Area Type, Scenario A08

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway/Expressway	27,651,525	31,897,217	0.87	579
Primary Arterial	19,887,434	21,126,553	0.94	1389
Minor Arterial	5,762,336	7,225,138	0.80	1007
Collector	1,872,658	2,335,184	0.80	630
Centroid Connector	1,535,016	1,462,416	1.05	84
Other	118,976	142,304	0.84	23
Area Type	**Volume**	**Count**	**Volume/Count**	**# of Links**
Area Type 1	15,374,517	17,084,579	0.90	1379
Area Type 2	16,400,766	18,716,142	0.88	841
Area Type 3	15,933,720	18,151,617	0.88	813
Area Type 4	5,142,532	5,547,453	0.93	306
Area Type 5	1,674,538	1,973,068	0.85	162
Area Type 6	367,616	309,377	1.19	25
Area Type 7	882,552	1,095,707	0.81	95
Area Type 8	266,515	399,473	0.67	36
Area Type 9	785,189	911,396	0.86	55
Total	**56,827,945**	**64,188,812**	**0.89**	**3712**

Table 3-5 Volume and Count Summary by Facility Type and Area Type, Scenario B00

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway/Expressway	24,808,001	24,996,659	0.99	508
Primary Arterial	9,430,400	9,948,706	0.95	650
Minor Arterial	3,388,832	3,649,554	0.93	546
Collector	1,165,796	1,289,412	0.90	372
Centroid Connector	1,515,576	1,514,000	1.00	84
Other	149,128	162,700	0.92	19
Area Type	**Volume**	**Count**	**Volume/Count**	**# of Links**
Area Type 1	13,831,057	13,147,948	1.05	923
Area Type 2	13,897,304	14,341,286	0.97	546
Area Type 3	8,460,258	9,248,652	0.91	385
Area Type 4	1,755,463	1,890,430	0.93	103
Area Type 5	1,191,122	1,317,141	0.90	86
Area Type 6	40,663	52,100	0.78	8
Area Type 7	768,005	940,161	0.82	77
Area Type 8	132,888	183,121	0.73	25
Area Type 9	380,973	440,192	0.87	26
Total	**40,457,733**	**41,561,031**	**0.97**	**2179**

3.1.4 Vehicle Miles Traveled

Table 3-6 presents a summary of average weekday vehicle miles traveled (VMT) by facility type for the four scenarios. Both Models (A and B) show about a five percent increase in VMT between 2000 and 2008. However, the difference in VMT between the two validated base year scenarios (A00 and B08) is about 10 percent. This implies that **either model by itself underestimates growth in VMT by about half**. More information about this difference is provided through subsequent analyses discussed later.

Table 3-6 Average Weekday VMT Summaries by Facility Type for All Scenarios (millions)

Facility Type	Scenario A00	Scenario A08	Scenario B00	Scenario B08
Freeway/Expressway	49.8	50.2	50.5	51.6
Primary Arterial	25.7	29.5	27.6	31.4
Minor Arterial	22.9	23.5	25.7	26.3
Collector	2.9	6.5	5.1	6.1
Centroid Connector	12.1	12.7	11.9	12.3
Other	3.0	3.5	3.1	3.0
Total	113.5	119.3	118.8	124.6

Model A's forecast of VMT for 2008 is nearly five percent lower than the validated 2008 base year model (Scenario B08) while Model B's backcast of year 2000 VMT is higher than the validated 2000 base year model (Scenario A00) by about the same amount. Since the two years' socioeconomic data inputs are the same, **it is very likely that these discrepancies result from differences in the two models' parameters and/or assumptions**.

Table 3-7 shows the VMT estimates from the four model scenarios by time period. Model B shows slightly higher amounts of travel in peak periods than Model A, regardless of which year is examined. This result likely reflects higher factors used to estimate peak period demand from daily demand. Both models show little change in peak travel percentages from 2000 to 2008, with very slight increases in peaking in both cases. The main takeaway from these results is that **models that use fixed factors to convert daily demand to peak period demand can show very little change in peaking over time**. It is unknown whether peaking of travel actually increased from 2000 to 2008, as the comparison of the two base year scenarios (A00 and B08) would indicate, because the differences are not significant given the relatively small samples of survey data on which the factors are based. But even if they are, the models do not seem able to capture these differences.

Table 3-7 VMT Summaries by Time Period for All Scenarios (millions)

	Scenario A00		Scenario A08		Scenario B00		Scenario B08	
Time Period	VMT	% of daily	VMT	% of daily	VMT	% of daily	VMT	% of daily
A.M. Peak	24.8	21.9%	26.2	22.0%	27.6	23.2%	29.0	23.3%
Mid-Day	35.2	31.0%	36.8	30.8%	35.8	30.1%	37.5	30.1%
P.M. Peak	31.3	27.6%	33.0	27.7%	33.3	28.0%	35.1	28.2%
Night	22.1	19.5%	23.3	19.5%	22.0	18.5%	23.0	18.5%
Total	113.5		119.3		118.8		124.6	

3.1.5 Average Speeds

Table 3-8 presents a summary of average speeds by facility type for all four scenarios. (In this case, average speeds are computed by dividing VMT by vehicle hours traveled (VHT), and so the effects of traffic congestion are considered.) There are some notable observations:

- Average speeds are about five percent lower in Model B although this difference is almost entirely attributable to significantly lower speeds on freeways/expressways. The lookup tables that provide the free flow speed inputs for the two models have lower freeway/ expressway speeds in Model B than in Model A, which explains the differences. BMC has reported that Model A had overestimated volumes on freeways/expressways, and as a result the input speeds are lower in the newer Model B.
- Both models show small increases in speeds from 2000 to 2008 (i.e., comparing the results of Scenarios A00 and A08, and Scenarios B00 and B08), which is inconsistent with the five percent increase in VMT in both models. The speed increases are a little larger for freeways/expressways. The source of this inconsistency is unknown.

Table 3-8 Average Speed Summaries by Facility Type for All Scenarios (mph)

Facility Type	Scenario A00	Scenario A08	Scenario B00	Scenario B08
Freeway/Expressway	52.3	53.7	45.9	46.9
Primary Arterial	32.4	33.6	31.9	32.3
Minor Arterial	30.4	30.5	30.4	30.4
Collector	25.8	23.7	27.1	26.1
Centroid Connector	26.1	26.1	26.2	26.3
Other	25.9	28.2	27.5	28.0
Total	36.9	37.4	35.2	35.5

3.1.6 Traffic Congestion

Table 3-9 displays for each scenario the percentage of VMT within various ranges of volume/capacity ratio for the peak periods (a.m. and p.m. combined). Also shown are the average speeds for highway links within each v/c ratio range. Model B shows more congestion than Model A for both 2000 and 2008. About eight percent of links have v/c ratios above 1.0 in the peak periods in Model A; this rises to 13 to 15 percent in Model B. There is a slight increase in congestion going from 2000 to 2008 in both models, consistent with an increase in VMT.

Speeds on congested links (those with v/c greater than 1.0) are a bit higher in 2008 than in 2000 in both models. Speeds on congested links are generally higher in Model B than Model A, in contrast to average speeds on all links, which are lower in Model B.

Table 3-9 Percentage of VMT and Average Speeds by Volume/Capacity Ratio for All Scenarios (Peak Periods)

Volume/ Capacity Ratio	Scenario A00		Scenario A08		Scenario B00		Scenario B08	
	% VMT	Average Speed	% VMT	Average Speed	% VMT	Average Speed	% VMT	Average Speed
0.00-0.25	17.5%	29.7	17.0%	29.6	15.1%	29.1	14.4%	28.8
0.25-0.50	17.9%	39.4	17.6%	39.4	14.8%	36.9	14.2%	36.8
0.50-0.75	26.8%	44.6	26.6%	43.5	23.4%	40.1	23.0%	39.6
0.75-1.00	30.3%	40.8	31.0%	40.0	33.3%	37.7	33.8%	36.9
Over 1.00	7.5%	12.3	7.8%	15.7	13.5%	15.9	14.5%	18.4
Total		33.3		34.4		30.9		31.5

3.1.7 Trip Generation and Trip Lengths

Table 3-10 summarizes the trip generation results for the four scenarios. These numbers represent motorized person trips. Both Model A and Model B show about a three percent increase in total trips from 2000 to 2008 and a slight decrease in the number of trips per household. These figures are consistent with the four percent increase in the number of households from 2000 to 2008 along with the decrease in household size. In both models, more than half of the increase in total trips from 2000 to 2008 is in home based work trips; home based school trips decrease slightly while home based non-work trips decrease a little. Non-home based trips increase in both models. Model A produces about two percent more trips than Model B although Model B produces more home based school and non-home based trips. All of these differences can be attributed to changes in trip generation rates, which for each model were based on the latest available household survey data.

Table 3-10 Trip Generation Summary for All Scenarios

	Scenario A00			Scenario A08		
Trip Purpose	Trips	Trips/ Household	% Trips	Trips	Trips/ Household	% Trips
HBW	2,844,412	1.5	19%	3,107,253	1.6	20%
HBSc	942,617	0.5	6%	902,515	0.5	6%
HBSh	2,542,957	1.3	17%	2,605,133	1.3	17%
HBO	4,418,732	2.3	30%	4,449,273	2.3	29%
NHB	3,993,140	2.1	27%	4,174,094	2.1	27%
Total	14,741,858	7.8		15,238,268	7.7	
	Scenario B00			Scenario B08		
Trip Purpose	Trips	Trips/ Household	% Trips	Trips	Trips/ Household	% Trips
HBW	2,551,942	1.3	18%	2,812,507	1.4	19%
HBSc	1,115,558	0.6	8%	1,070,455	0.5	7%
HBSh	2,179,029	1.2	15%	2,238,189	1.1	15%
HBO	4,544,240	2.4	31%	4,564,698	2.3	31%
NHB	4,086,304	2.2	28%	4,260,245	2.2	29%
Total	14,477,073	7.7		14,946,094	7.6	

Trip purpose abbreviations: HBW – home based work, HBSc – home based school, HBSh – home based shopping, HBO – home based other, NHB – non-home based

Table 3-11 presents the average trip lengths in minutes by trip purpose for all four scenarios. In both Model A and Model B, average trip lengths increase slightly from 2000 to 2008, consistent with a small increase in traffic congestion. The main difference between the two models is the higher home based work trip lengths in Model B (about four minutes, or 20 percent). This is consistent with higher levels of peak period congestion in Model B. Note that BMC feels that it is difficult to determine whether actual work trip lengths changed from 2000 to 2008. Typically, U.S. Census Journey to Work (JTW) data[1] would be used for this type of comparison; however, BMC has noted that JTW data has apparent geocoding issues regarding Baltimore City and Baltimore County, and so they rely more on travel survey data for these types of analyses. However, there is no household survey for 2000 for Baltimore.

[1] For 2000, JTW data would be developed from the 2000 U.S. Census "long form"; for 2008, data would be developed from the American Community Survey.

Table 3-11 Average Trip Lengths in Minutes for All Scenarios

Trip Purpose	Scenario A00	Scenario A08	Scenario B00	Scenario B08
HBW	21.1	21.4	25.3	25.7
HBSc	9.3	9.8	10.1	10.4
HBSh	10.5	10.5	10.7	10.8
HBO	12.5	13.0	12.5	12.8
NHB	17.3	17.6	17.5	17.7
Total	13.5	14.0	14.5	15.1

Trip purpose abbreviations: HBW – home based work, HBSc – home based school, HBSh – home based shopping, HBO – home based other, NHB – non-home based

3.1.8 Mode Shares and Transit Boardings

Table 3-12 presents a summary of the modeled mode shares for the four scenarios. Transit shares are similar in the two base year scenarios (A00 and B08), about four percent in each (slightly lower in 2008). The transit shares for home based work trips are in the 12 to 13 percent range. However, Model A shows an increase in transit share from 2000 to 2008, for total trips and home based work trips, with most of the increase being in transit with auto access. Model B also shows a transit share increase when comparing the backcast scenario (B00) to the base year scenario (B08) though in Model B most of the increase is in transit trips with walk access. Model A generally predicts higher transit shares than Model B (for home based work trips, about 13 percent versus 11 percent for 2000 and about 14 percent versus 12 percent for 2008; for all trips, about 4.1 percent versus 3.4 percent for 2000 and about 4.6 percent versus 3.9 percent for 2008).

Table 3-12 Mode Shares by Trip Purpose for All Scenarios

Trip Purpose	Scenario A00 SOV	Scenario A00 HOV	Scenario A00 Tr-W	Scenario A00 Tr-A	Scenario A08 SOV	Scenario A08 HOV	Scenario A08 Tr-W	Scenario A08 Tr-A
HBW	73.5%	14.0%	11.7%	0.8%	72.3%	13.7%	11.7%	2.3%
HBSc	3.8%	93.9%	2.2%	0.0%	3.7%	94.0%	2.2%	0.1%
HBSh	45.1%	53.7%	1.2%	0.0%	45.7%	53.3%	1.0%	0.0%
HBO	35.9%	61.5%	2.5%	0.0%	35.6%	61.6%	2.8%	0.0%
NHB	50.3%	47.8%	1.9%	0.0%	53.2%	49.1%	2.2%	0.0%
Total	46.3%	49.7%	3.9%	0.2%	46.8%	48.6%	4.1%	0.5%

Trip Purpose	Scenario B00 SOV	Scenario B00 HOV	Scenario B00 Tr-W	Scenario B00 Tr-A	Scenario B08 SOV	Scenario B08 HOV	Scenario B08 Tr-W	Scenario B08 Tr-A
HBW	76.0%	13.0%	9.7%	1.4%	74.6%	12.9%	11.0%	1.4%
HBSc	3.6%	94.9%	1.5%	0.0%	3.6%	95.0%	1.4%	0.1%
HBSh	44.7%	53.7%	1.5%	0.0%	45.0%	53.7%	1.3%	0.0%
HBO	39.9%	58.0%	2.0%	0.1%	39.8%	57.7%	2.4%	0.1%
NHB	54.6%	43.7%	1.6%	0.0%	57.4%	44.9%	2.0%	0.0%
Total	48.4%	48.2%	3.1%	0.3%	48.9%	47.2%	3.6%	0.3%

Trip purpose abbreviations: HBW – home based work, HBSc – home based school, HBSh – home based shopping, HBO – home based other, NHB – non-home based

Mode abbreviations: SOV – single occupant vehicle (drive alone), HOV – high occupancy vehicle (auto carpool), Tr-W – transit with walk access, Tr-A – transit with auto access

The percentage of transit trips that have auto access is relatively low in both models. However, for the year 2000, the auto access share of transit trips is about twice as high in Model B (nine percent versus four percent in Model A). For the year 2008, the auto access share of all transit trips is 11 percent in Model A but only eight percent in Model B. This means that Model A predicts a significant increase in the auto access share from 2000 to 2008, but Model B predicts a slight decrease. There is no evidence to suggest why a significant increase in transit with auto access from 2000 to 2008 should have occurred.

Table 3-13 summarizes the transit boardings and transfer rates for the four scenarios. The higher estimated transit shares in Model A are evident in these results. Both models show an increase in transit linked trips from 2000 to 2008; the percentage increase is higher in Model A. However, the transit boardings decrease from 2000 to 2008 in Model A while they increase by more than 10 percent in Model B, as the transfer rate drops significantly from 2000 to 2008 in Model A.

Table 3-13 Transit Boarding Summaries for All Scenarios

	Scenario A00	Scenario A08	Scenario B00	Scenario B08
Transit boardings	345,659	341,520	271,045	303,213
Transit linked trips	193,286	231,005	168,295	186,628
Transfer rate (boardings per trip)	1.79	1.48	1.61	1.62

Data on observed transit ridership indicates that actual year 2000 ridership was about 333,000 and actual year 2008 ridership was about 296,000. The base year models (A00 and B08) therefore reflect observed transit ridership fairly accurately. There is no available observed data to determine whether the transfer rate may have changed substantially between 2000 and 2008. BMC has noted that there were only minor transit system changes between 2000 and 2008 (a few new routes, combining of some routes, and some schedule changes) though there was a significant fare increase in 2003. While there is not enough information to say with certainty, it seems improbable that the transfer rate dropped by 17 percent from 2000 to 2008, as estimated by Model A. On the other hand, actual transit ridership decreased about 10 percent from 2000 to 2008, as contrasted with the 12 percent increase predicted by Model B. **Neither model seemed to be able to capture the differences in travel behavior related to transit that are observed between 2000 and 2008 very well**.

3.2 OKI Model Runs

3.2.1 Model Versions and Scenarios

Two versions of the OKI model were used for analysis (labeled Model C and Model D for convenience). Note that the OKI model region includes the Miami Valley region around Dayton as well as the OKI planning region around Cincinnati. Because Dayton has its own MPO, the OKI staff focuses more on the Cincinnati portion of the model in terms of validation.

The two model versions are:

- **Model C** – This model was estimated with a base year of 2005. OKI has stated that their later analysis of the 2005 model validation indicated indicates that much of the traffic count data associated with Model C was suspect, and they discovered other issues with the model that they attempted to address in the development of Model D.
- **Model D** – This model has a validation base year of 2010. The OKI modeling staff had significant turnover since the validation of Model C, and the new staff determined to correct some of the issues with that model. Therefore, there are some major changes in Model D that make some comparisons to Model C difficult. It should be noted that the validation of Model D was not 100 percent complete at the time of the analyses presented here.

The TAZ system was revised when Model D was developed, and so the TAZ systems are not the same for the two models. The following scenarios were run using the two OKI model versions:

- **Scenario C05** – Model C run for the base year of 2005
- **Scenario C10** – Model C run for the forecast year of 2010
- **Scenario D10** – Model D run for the base year of 2010
- **Scenario D05** – Model D run for the backcast year of 2005

In developing Model D, OKI made substantial improvements over Model C. The extent of these changes made some comparisons difficult, and some tasks that we wished to perform could not be done due to resource limitations. For example, the network changed substantially between the two model versions, and the project resources were insufficient to allow for the attachment of the traffic counts from each model's base year to the forecast and backcast scenarios (e.g., attaching 2010 counts to the network for Scenario C10).

3.2.2 Socioeconomic Data Inputs

The socioeconomic data inputs to Models C and D are shown in Table 3-14 and are, in the aggregate, the same for each model year (2005 and 2010) in both models. (There are localized differences because of the changes in the TAZ structure between the two models.) Therefore, differences between the results of the two models for the same analysis year are not due to differences in the model input socioeconomic data. Note that OKI had estimated nearly flat employment growth from 2005 to 2010.

Table 3-14 OKI Model Socioeconomic Data Input Summary

Measure	2000	2008	Percentage Difference
Households	1,891,211	1,973,108	4.3%
Population	4,964,091	4,956,580	-0.2%
Employment	2,934,952	3,479,179	18.5%

3.2.3 Highway Assignment Validation

Tables 3-15 and 3-16 show the modeled volume/count ratios by facility type and area type for the two base year scenarios (C05 and D10) respectively. Note that there are many more counts for 2005, which makes some comparisons difficult. In addition, OKI believes that the data for many of the 2005 count locations are suspect.

Model C appears to overestimate highway travel (though, again, the count data are suspect). The traffic volumes in Model D match the traffic counts better on average than those in Model C. The modeled volumes on collectors are a bit lower (about 11 percent) than the counts in Model D.

Traffic counts that were provided by OKI were coded to the base year networks only. Because the networks are quite different between Models C and D, and the project resources were not sufficient to provide for the traffic counts for a particular year to be attached to the backcast and forecast year scenarios, no comparisons of the modeled volumes for the backcast and forecast year scenarios to the traffic counts were made.

Table 3-15 Volume and Count Summary by Facility Type and Area Type, Scenario C05

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway	33,215,836	32,523,216	1.02	738
Expressway	6,793,179	5,630,484	1.21	401
Ramp	7,685,111	7,511,937	1.02	1,100
Arterial	64,334,553	58,059,239	1.11	8,886
Collector/Local	10,661,460	11,288,380	0.94	4,762
Area Type	**Volume**	**Count**	**Volume/Count**	**# of Links**
CBD	3,955,208	3,780,201	1.05	501
Urban	40,429,128	37,537,613	1.08	4,689
Suburban	70,221,938	66,685,560	1.05	7,795
Rural	8,083,865	7,009,882	1.15	2,902
Total	**122,690,139**	**115,013,256**	**1.07**	**15,887**

Table 3-16 Volume and Count Summary by Facility Type and Area Type, Scenario D10

Facility Type	Volume	Count	Volume/Count	# of Links
Freeway	33,215,836	32,523,216	1.02	738
Expressway	6,793,179	5,630,484	1.21	401
Ramp	7,685,111	7,511,937	1.02	1,100
Arterial	64,334,553	58,059,239	1.11	8,886
Collector/Local	10,661,460	11,288,380	0.94	4,762
Area Type	**Volume**	**Count**	**Volume/Count**	**# of Links**
CBD	3,955,208	3,780,201	1.05	501
Urban	40,429,128	37,537,613	1.08	4,689
Suburban	70,221,938	66,685,560	1.05	7,795
Rural	8,083,865	7,009,882	1.15	2,902
Total	**122,690,139**	**115,013,256**	**1.07**	**15,887**

3.2.4 Vehicle Miles Traveled

Table 3-17 presents a summary of average weekday VMT by facility type for the four scenarios. (Note that there were some changes in facility types for some roadways between Models C and D, but OKI has indicated that these changes are not extensive.) Both Models (C and D) show an increase in VMT between 2005 and 2010, but Model D shows a much smaller percentage increase. However, the modeled VMT in the 2005 base year scenario (C05) is much higher than that in the validated 2010 base year scenario (D10). Given the lower level of confidence in the 2005 traffic counts and the 2005 base year validation, and their knowledge of growth in their region, OKI feels that the small VMT increase from 2005 to 2010 implied by Model D (as demonstrated by the difference between the VMT estimates for Scenarios D05 and D10) is more reasonable.

Table 3-17 Average Weekday VMT Summaries by Facility Type for All Scenarios (millions)

Facility Type	Scenario C05	Scenario C10	Scenario D05	Scenario D10
Freeway	25.5	26.7	24.4	25.0
Expressway	4.8	5.0	3.4	3.5
Ramp	2.2	2.3	2.6	2.6
Arterial	28.8	30.1	22.1	22.0
Collector/Local	11.1	11.7	13.4	13.7
Total	72.4	75.8	65.9	66.7

3.2.5 Average Speeds

Table 3-18 presents a summary of average speeds by facility type for all four scenarios. (Again, average speeds are computed by dividing VMT by VHT, and so the effects of traffic congestion are considered.) There are some notable observations:

- Average speeds in Model D are substantially higher than those in Model C, by about 10 mph. The speed differences between the two models are much greater on roadways with lower functional classifications. Some of the differences could be due to changes made to the input free flow speeds in the newer Model D.
- It is notable that when the summaries for Model C were initially created, the average speeds were even lower, especially for ramps. Examination of several specific network links indicated incorrectly coded capacities, which caused modeled speeds to be very low. These capacities were corrected, and Model C was rerun, resulting in the summaries provided in this report. However, an exhaustive check of all link capacities was not done, and so it is possible that there are additional links with incorrectly coded capacities that might be contributing to the lower speeds indicated in Model C.
- Model C showed about a 10 percent decrease in average speed from 2005 to 2010 while Model D showed small speed increases (i.e., when comparing the results of Scenarios C05 and C10, and Scenarios D05 and D10). This finding is consistent with the increase in VMT in Model C from 2005 to 2010 and the relatively small change in VMT for Model D. As noted previously, the VMT increase estimated by Model D is felt to be more reasonable.

Table 3-18 Average Speed Summaries by Facility Type for All Scenarios (mph)

Facility Type	Scenario C05	Scenario C10	Scenario D05	Scenario D10
Freeway	55.3	51.1	53.9	54.5
Expressway	50.0	48.5	50.6	51.0
Ramp	11.9	9.2	39.4	42.3
Arterial	25.9	25.1	34.2	34.5
Collector/Local	18.7	15.6	29.9	29.7
Total	29.6	26.9	39.2	39.6

Table 3-19 presents the modeled travel times and speeds for the four scenarios for five important corridors to downtown Cincinnati in the a.m. peak period. The overall higher speeds in Model D are consistent with the aggregate results presented in Table 3-18. Model D shows consistency between the 2005 and 2010 travel time and speed estimates, but Model C shows substantial reductions in speeds over the five year period. While observed data are not available, OKI staff feels that the substantial speed changes indicated by the Model C results do not reflect real changes in these corridors.

Table 3-19 Travel Times and Speeds to Downtown Cincinnati for Major Corridors for All Scenarios, A.M. Peak

			Scenario C05		Scenario C10	
From	Distance (miles)	Free Flow Time (min)	Model Time (min)	Model Speed (mph)	Model Time (min)	Model Speed (mph)
CVG[1]	13.4	18.2	33.0	24.0	53.7	14.8
Eastgate	16.2	20.0	44.9	21.0	58.6	15.9
NKU[2]	7.8	12.3	20.9	19.3	30.9	13.0
Kings Island	23.3	27.1	50.7	27.7	63.3	22.1
Sharonville	15.1	20.0	28.0	30.5	28.7	29.8
			Scenario D05		**Scenario D10**	
From	**Distance (miles)**	**Free Flow Time (min)**	**Model Time (min)**	**Model Speed (mph)**	**Model Time (min)**	**Model Speed (mph)**
CVG[1]	13.4	18.2	24.5	32.8	24.3	33.0
Eastgate	16.2	20.0	26.3	36.9	26.7	36.4
NKU[2]	7.8	12.3	15.2	30.8	15.2	30.8
Kings Island	23.3	27.1	33.2	42.1	34.5	40.5
Sharonville	15.1	20.0	26.4	34.3	27.2	33.3

1. CVG represents Cincinnati/Northern Kentucky International Airport.
2. NKU represents Northern Kentucky University in Highland Heights, Kentucky

3.2.6 Traffic Congestion

Table 3-20 displays for each scenario the percentage of VMT within various ranges of volume/capacity ratio for an average weekday. Also shown are the average speeds for highway links within each v/c ratio range. Model C shows more congestion than Model D for both 2005 and 2010. About nine percent of links have v/c ratios above 1.0 in the peak periods in Model C; this decreases to about two percent in Model D. Speeds on congested links (those with v/c greater than 1.0) are higher in Model D. In both models, the amount of congestion, as measured by the percentage of links with v/c ratios greater than 1.0, is about the same in both 2005 and 2010.

Generally, these results are consistent with previously presented results. As noted earlier, overall speeds are higher in Model D. Table 3-20 shows that there is more congestion in Model C, and the speeds on congested links are lower in Model C. Consistent with earlier observations, Model C shows a decrease in speed from 2005 to 2010 while Model D does not.

Table 3-20 Percentage of VMT and Average Speeds by Volume/Capacity Ratio for All Scenarios (Peak Periods)

Volume/ Capacity Ratio	Scenario C05 % VMT	Scenario C05 Average Speed	Scenario C10 % VMT	Scenario C10 Average Speed	Scenario D05 % VMT	Scenario D05 Average Speed	Scenario D10 % VMT	Scenario D10 Average Speed
0.00-0.25	10.5%	37.8	16.1%	27.1	22.7%	34.7	22.6%	34.5
0.25-0.50	28.3%	42.1	26.3%	42.2	30.9%	46.2	31.8%	46.6
0.50-0.75	32.4%	40.4	30.4%	40.8	30.7%	43.9	30.1%	44.6
0.75-1.00	19.5%	30.3	17.9%	29.1	13.3%	33.8	13.1%	34.1
Over 1.00	9.3%	10.8	9.3%	8.1	2.4%	17.8	2.4%	18.4
Total		30.7		26.9		39.2		39.6

3.2.7 Trip Generation and Trip Lengths

Table 3-21 summarizes the trip generation results for the four scenarios. These numbers represent motorized person trips. The trip rates per household are more than 10 percent higher in Model C than in Model D, and Model C has a higher percentage of trips that are home based work. The higher trip rates in Model C are consistent with that model's higher VMT estimates.

There is no information available to suggest that one model's rates are more "correct" than the other. Generally, the trip rates are consistent between the two scenarios for each model, in terms of both the number of trips per household and the percentages of trips by purpose.

Table 3-21 Trip Generation Summary for All Scenarios

Trip Purpose	Scenario C05 Trips	Scenario C05 Trips/ Household	Scenario C05 % Trips	Scenario C10 Trips	Scenario C10 Trips/ Household	Scenario C10 % Trips
HBW	1,782,633	1.6	19%	1,832,895	1.7	19%
HBNW	4,996,935	4.6	52%	5,110,826	4.6	52%
NHB	2,821,279	2.6	29%	2,888,266	2.6	29%
Total	**9,600,847**	8.7		**9,831,987**	8.9	

Trip Purpose	Scenario D05 Trips	Scenario D05 Trips/ Household	Scenario D05 % Trips	Scenario D10 Trips	Scenario D10 Trips/ Household	Scenario D10 % Trips
HBW	1,323,615	1.2	15%	1,267,228	1.1	15%
HBNW	4,873,090	4.4	56%	4,989,553	4.5	57%
NHB	2,487,994	2.3	29%	2,448,128	2.2	28%
Total	**8,684,699**	7.9		**8,704,909**	7.8	

Trip purpose abbreviations: HBW – home based work, HBNW – home based nonwork, NHB – non-home based

Table 3-22 presents the average trip lengths in minutes by trip purpose for all four scenarios. In both Model C and Model D, average trip lengths increase slightly from 2005 to 2010. The main difference between the two models is that Model C has higher trip lengths (by over 50 percent) than Model D, especially for home based trips. This result corresponds to the lower speeds in Model C noted earlier and indicates that trip lengths in terms of distance are similar between the two models.

Table 3-22 Average Trip Lengths in Minutes for All Scenarios

Trip Purpose	Scenario C05	Scenario C10	Scenario D05	Scenario D10
HBW	20.0	20.7	13.0	13.1
HBNW	12.2	12.5	8.1	8.2
NHB	10.0	10.1	8.0	8.0
Total	**13.0**	**13.3**	**8.8**	**8.9**

Trip purpose abbreviations: HBW – home based work, HBNW – home based nonwork, NHB – non-home based

3.2.8 Mode Shares and Transit Boardings

Table 3-23 presents a summary of the modeled mode shares for the four scenarios while Table 3-24 compares transit boardings among the scenarios. Transit shares are substantially higher in Model C than in Model D. OKI has indicated that the indicated that the actual transit share is closer to the Model D results although the Model D transit shares are too low. The Model D results may be off because the validation of transit demand in the Dayton part of the model had not been completed at the time of the conclusion of this study. The large differences between the two models makes transit demand and ridership comparisons between the models uninformative, regardless of the model years.

Table 3-23 Mode Shares by Trip Purpose for All Scenarios

	SOV	HOV	Transit
C05	51.8%	39.9%	8.3%
C10	51.8%	41.1%	7.1%
D05	45.8%	53.6%	0.6%
D10	45.4%	54.0%	0.6%

Mode abbreviations: SOV – single occupant vehicle (drive alone), HOV – high occupancy vehicle (auto carpool)

Table 3-24 Transit Boarding Summaries for All Scenarios

	Scenario C05	Scenario C10	Scenario D05	Scenario D10
Transit boardings	1,110,126	953,085	78,373	83,358
Transit linked trips	833,410	724,192	63,297	59,431
Transfer rate (boardings per trip)	1.33	1.32	1.24	1.40

3.3 Discussion

The analysis presented in this chapter shows a variety of results from the four scenarios that were run using the two model versions for each region. A number of observations can be made regarding these results concerning how the models performed in forecasting and backcasting:

- It is obvious that some problems with travel demand forecasts in general have to do with inaccuracies in model inputs such as socioeconomic data. This is unavoidable, of course, since future conditions such as land use and demographics are never known with certainty, and even transportation system conditions, which generally come from policy decisions, can turn out differently in the future than what is assumed in the present. For the analyses using both the BMC and OKI models, the socioeconomic data inputs were the same for each model year regardless of which model version was used. Transportation network inputs were kept consistent for both model years as well, to the extent possible. This means that the differences in model results can be attributed to differences between the model versions' structures, assumptions, and parameters.
- The OKI model underwent substantial changes between the development of Model C and Model D. The extent of the changes seems to have been influenced by changes in the key OKI modeling staff between the development periods of the two models and recognitions of some of the key shortcomings of Model C. The magnitude of the changes between the two models manifests itself in several ways in examining the results of the four scenarios. The most notable differences include the following:
    - Higher modeled speeds in Model D
    - Much higher transit demand estimates in Model C
    - Higher trip rates in Model C

  The OKI staff feels that while the validation of Model D was not 100 percent complete at the time of the analyses presented here, that model reflects travel behavior in their region more accurately than Model C. The major differences between the two models cause substantial differences in the results for the same model year (i.e., the Model C 2005 base year compared to the Model D 2005 backcast, and the Model D 2010 base year compared to the Model C 2010 forecast). Normally, we would expect the base year results for a particular model year to be more accurate than a forecast or backcast to that year. However, particularly because of the known issues with Model C, <u>it is not clear that the 2005 base year is a good basis for comparison for the Model D backcast, and the 2010 forecast from Model C is questionable as well</u>.

- Not surprisingly, the results of the BMC model runs show that highway assignment results from base year model scenarios more accurately reflect observed traffic volumes than results from backcast or forecast year scenarios. Model validation and calibration is done considering observed data for base year conditions, and analysts attempt to achieve good fits between model results and the observed data. But model fit depends on many factors, including model inputs, model structure (the specific set of model components and how they work together), model parameters, and the specific mathematical formulations that form the model (such as volume-delay functions or logit model formulations).

The analyses described in Section 3.1 show that differences in the features between model versions affect the model results. Specifically for the BMC model, both versions show only about half the observed percentage increase in VMT between the two analysis years, even though they are only eight years apart. This does not mean, of course, that all models underestimate traffic volume growth. In the case of the BMC model, the increase in VMT outpaced the growth in model input variables, such as the number of households in the model region, that drive travel demand. While there is no way to know the precise reasons for the underestimation of growth in travel demand in our analyses, some possible reasons might include the following:

- **Changes in travel demand for trips entering and leaving the model region are not adequately reflected**. It is noted that since regional employment growth in the analysis period was significant (10 percent over eight years) while population growth was stagnant, commuting patterns in and out of the region must have changed, and these changes may not have been captured in the model update.
- **The effects of model improvements made in the newer version may have compensated for calibration adjustments to the older model**. BMC had targeted specific areas in the earlier model that they wished to improve, resulting in changes in model components (e.g., adding a toll diversion model), model formulations (e.g., changes in volume-delay functions), model parameters (e.g., trip rates and time of day factors), and model inputs (e.g., free flow speed inputs). This leads to an important hypothesis: *It is likely that model calibration adjustments in the earlier model compensated for some of the deficiencies that the later model improvements addressed*. This points out the risk involved in model calibration adjustments made to better reflect observed base year travel demand, which may not reflect the behavioral issues that caused the initial inaccuracies and may not carry forward in forecasting (or backward in backcasting).
- **The effects of calibration changes on model results are not well understood**. At the time they were developed, both versions of the BMC model were validated to reflect base year conditions, with calibration changes made to model parameters as needed. However, no backcasts were performed when the models were originally validated. Without a second data point beside the base year, the effects of these calibration changes on model results over time are unknown. It is possible that the effects of some calibrated parameters could have been to dampen changes in demand due to various factors.

• The results presented here show very similar results for either version of the OKI model for both 2005 and 2010 when using the same model, especially for the more robust Model D. While these similarities are expected due to the relatively low growth between

2005 and 2010 in the OKI model region, the consistency across scenarios echoes the results of the BMC model scenarios.

- Model parameters that are constants may fail to reflect real changes related to the behavior being modeled by a particular component. An example here is the use of fixed factors to convert daily travel to peak period travel. There are small differences in peak period travel percentages between 2000 and 2008 in the validated base year BMC models, but the forecasts and backcasts do not reflect these differences, instead producing peak period travel proportions that reflect their validation base years.
- The BMC model results seem to hint that model components near the beginning of the model stream (e.g., trip generation) may be more stable in forecasting; however, this observation is tempered by the fact that actual trip rates appeared to be relatively stable between 2000 and 2008 in the BMC model region. Still, both model versions captured a good portion of the overall decline in aggregate trip rates from 2000 to 2008. The results of the next step in the model stream, trip distribution, as demonstrated by average trip lengths, are also reasonable although data limitations make drawing specific conclusions difficult.
- Both BMC model versions had difficulty capturing trends in transit travel over the analysis period. Observed ridership data indicated about a 10 percent decline in transit boardings, and each validated base year model accurately portrayed that year's overall ridership. However, neither model captured the decline in transit travel from 2000 to 2008. Model B showed an increase in transit boardings, rather than a decrease, and while Model A showed a very small decrease, it was achieved through a substantial decline in the transfer rate that was unsupported by observed data. There may be other contributing factors, such as the higher error associated with the relatively small number of transit trips among all regional trips, but it seems likely that some of the factors that affect decisions on whether to use transit are not considered well in models.
- It seems that, as a general rule, a model tends to produce similar and/or consistent results for different scenarios, but a new or updated model can produce substantially different results for a scenario from those for an earlier model version for the same scenario. This is not necessarily a new revelation; analysts have always been concerned that using an updated model could change the results of previously done analyses. But it would seem to be a worthwhile effort when updating a model to consider the effects of the changes being made to the model. This is not to say that significant model improvements should not be made; it is undoubtedly true that the changes BMC made on the transit component of Model B were improvements compared to the previous version. But perhaps additional analysis, such as a backcast, might have led to further improvements so that transit demand was more accurately represented in both 2008 and 2000. (It is noteworthy that BMC is performing a backcast as part of the validation of its new activity based model.)

4.0 Lessons and Recommendations for Modeling Practitioners

At the outset of this project, it had been hoped that the comparisons of the various scenarios, including forecasts and backcasts, would provide a variety of specific lessons regarding the accuracy of model results for different types of scenarios. In particular, it was hoped that some conclusions could be drawn about which components in the models did better jobs of forecasting, and where in the models specific improvements were needed to produce more accurate forecasts.

It proved more difficult to reach these conclusions than anticipated. While it is not surprising that agencies who maintain models can make significant changes and improvements when updating a model, the differences between consecutive versions of the same model, for the two regions examined in this project, were so large in some cases that comparisons of forecasts and backcasts to observed data were difficult and, in some cases, problematic. It was therefore impossible to identify all of the specific components and features of the models where the most effort to improve them should be spent.

However, the project did provide some useful observations that can be shared. Furthermore, the unanticipated differences between consecutive model versions provided the opportunity to produce some guidance for analysts who are updating and validating models. This chapter describes the lessons learned along these lines and provides a set of recommendations for those developing and updating travel models.

This chapter refers to the findings from the runs of the BMC and OKI models for the various scenarios documented in Chapter 3. Some of the differences between the results of different scenarios, or between the model results and observed data, can be described as forecasting (or backcasting) "errors." In some cases, there is not enough information to determine why these "errors" occurred; in other cases, we can speculate on the reasons for them, attributing them to various features of the model or the input data. In such cases, we attempt to form lessons for modelers to help them avoid these "errors."

4.1 Lessons Learned

4.1.1 Challenges Associated with Backcasting

Backcasting is an appealing method for temporal forecasting since, like the base year, the backcast year has already occurred, and observed travel data are available, and often the backcast year has served as the base year for a previous model version, which means that some data may already have been processed for use in model application and validation. However, it was found that backcasting may be more challenging than might be expected. Specific challenges encountered included the following:

4.1.2 Changes in Model Parameters in Model Updates

Model updates can serve multiple purposes. For example:

- Changes in travel behavior since the last model update can be accounted for through the use of new or more recent data for model estimation and validation.

  - Examples: A new household travel survey can be used to reestimate model parameters, or GTFS data can be used to code the transit network.
- Errors discovered since the finalization of the previous model update can be corrected.
  - Example: Network coding errors can be corrected.
- Improved modeling methods that may have been too speculative to implement during the previous update, or for which resources were insufficient to implement, can be incorporated.
  - Example: The model structure could be changed from a trip based to an activity based model.
- Increases in computing power, hardware, or software that have become available since the last update can provide the opportunity for more powerful and robust methods to be incorporated.
  - Examples: A finer level of geographic detail can be incorporated to take advantage of faster processors, or the model can be updated from older legacy software implementations, such as FORTRAN programs.
- The model's analytical capabilities can be improved to address planning needs that have emerged since the last update.
  - Example: Active transportation and non-motorized travel can be more explicitly considered in the model.

The compound effects of all of these potential changes can be substantial, even if the base years of the previous and updated models are not that far apart (in the case of the OKI model, there was only a five year difference between the base years of the two model versions examined in this study). The newer model may produce different results for a particular analysis than the old model did, which can pose problems for planners using model results for regular planning functions (e.g., long range plans) and for studies that began but are not completed before the model update. While one might expect that, if the model update is done correctly, the results from the newer model are "better," having different sets of model results can pose problems for planners who had presented the results of the previous model as valid.

In this study, some of the model parameter changes produced noticeable changes in model results, but there was no way to determine whether the updated parameters were "more correct" than the previous parameters (or whether they were both "correct" and reflected real change in travel behavior over time). For example, in the BMC model, the trip generation results changed from Model A to Model B (in the terminology of Chapter 3), due to differences in the trip rates in the latest household survey results. There was no indication that the trip rates in either model were "incorrect."

The lesson for planners is that whenever a model is updated, the results will change, and this does not mean that either the original or updated model results should be considered incorrect. Planners need to be prepared to accept these differences and to acknowledge them when presenting model results or the results of analyses based on model outputs.

4.1.3 Accuracy of Data Inputs

It is obvious that inaccuracies in model input data, including socioeconomic data and transportation network data, can adversely affect model results. While this is hardly a new finding of this project, it is interesting to observe how network errors and their effects become evident in the forecasting/backcasting process.

The effects of inaccuracy in model input data pertains not only to the inputs to model application, but also to the data used for model development, including model estimation. Data sources such as household surveys also have errors associated with them, including sampling error, respondent errors, and data processing errors. These errors can affect model estimation, and therefore estimated model parameters.

In the OKI model, there were a number of ramps whose capacities were incorrectly coded in the base year network in their "old" model (Model C). When examining aggregate highway assignment results, the coding errors were not apparent, even when examining segmented results (such as vehicle miles traveled or average speeds by facility type). It was the comparison of the base year and forecast year results for Model C where the issue became apparent. The average modeled (congested) speeds for all ramps showed a substantial increase from the base year of 2005 to the forecast year of 2010. Since overall demand was relatively stable, and there were no highway improvements in the forecast year scenario that could have accounted for the speed increases, this finding prompted us to conduct a thorough review of the network coding for ramps in the highway network, and the coding errors on several ramps were discovered.

In the original concept of this project, it was not anticipated that the forecasting process would prove to be a means of identifying network coding errors. The comparison of speeds by facility type between the base and forecast years was intended as part of the comparison of highway assignment results, to show how these changed between base and forecast (and base and backcast scenarios) and how they differed between model versions. In this case, the comparison showed substantial differences in ramp speeds, which revealed the coding errors on some ramps, effecting the overall average speeds.

The ramp coding errors had been undiscovered prior to these comparisons, but there were other ways in which they could have been discovered during the original model validation. The lesson for modelers here is that this type of <u>undiscovered error in model inputs can have substantial effects on model results, reinforcing the need to thoroughly check model inputs such as network coding</u>.

4.1.4 Changes in Assumptions in Model Updates

Sometimes, changes made during model updates are not related only to differences between the data sources used to develop the previous and updated models. When updating a travel model, it is natural for analysts to want to "fix" issues identified or suspected with previous model versions. Often, these issues are discovered only after the original model has been developed and validated, and the issues appear when the model is being used for analysis. It is often infeasible to go back and revise and revalidate the original model. While the model validation process should include sensitivity testing that anticipates the probable uses of the model, it is difficult to foresee everything, and modelers often make changes to a model when updating it that are unrelated to the new data on which the model update is based.

An example of such a change in assumptions is the change in free flow speeds for freeways in the BMC model (from Model A to Model B). The speed changes have a significant effect on both modeled volumes and modeled speeds for freeways, but Model B produces more accurate estimates due to the change.

It is, of course, a positive action to "fix" issues when updating a model. The main thing to be aware of is that all of the effects of the "fix" should be examined during the validation of the updated model. For example, a change in free flow speeds may improve the accuracy of modeled base year volumes, but the effects on modeled speeds must also be examined.

4.1.5 Model Calibration Changes

While it is a common (and sometimes necessary) practice to change, or "calibrate" model parameters during the validation process, it can be tempting to make changes mainly to achieve a better match between base year model results and observed travel behavior. Perhaps the best approach to model calibration is to consider changes from estimated parameters only as needed to make the model a better forecasting tool. While achieving a better match to observed data sometimes can help in this objective, matching observations should not be the sole or main purpose of calibration changes.

It has been known that calibration changes can have the effect of "masking" errors or other deficiencies in models. A better match between model results and observed data can be easily achieved through targeted adjustments in model parameters, but if a model is to produce better forecasts, it is better for such adjustments to be made in response to a specific issue or shortcoming of the estimated model parameters.

For example, say that before validation and calibration, a model is producing VMT that is lower than observed. There could be several reasons for this, including the following:

- Socioeconomic data inputs, such as employment, may be underestimated.
- Trip generation rates may be too low for some market segments.
- Average trip lengths may be too short (which could be a result of various factors, such as incorrect trip distribution model parameters, inaccurate origin-destination patterns, or network speeds or capacities that are too low).
- Non-auto mode shares may be too high.
- External travel may be underestimated.
- Truck trips or trip lengths may be too low.
- Some travel markets may be underestimated or not even considered (for example, travel by non-residents of and visitors to the model region).

Or, a combination of such factors may be in play.

In the early days of travel modeling, it was common practice to address such an issue by adjusting the components of the model that were deemed "weakest," by virtue of data insufficiency. For example, since travel surveys in those days often tended to underestimate travel for non-work or non-home based purposes, trip rates for these trip purposes were often simply factored up to provide a better match to observed travel volumes. This type of change would help achieve a better match between modeled base year results and observed data but could actually make the model's predictive ability worse if the underestimation was due to some other model issue.

Model validation practice has improved over the years, with an emphasis on validating every component of a model, not only the overall results, and a focus on sensitivity testing and temporal validation. And it should be noted that there is no indication that calibration changes to either the BMC or OKI model were improper. But it is important to remember that the effects of calibration changes in forecasting are unknown, and it is difficult to estimate these effects.

4.1.6 Greater Accuracy in Earlier Model Components?

As noted in Section 3.3, the BMC model results seem to hint that model components near the beginning of the model stream (e.g., trip generation) may be more stable in forecasting. Overall trip rates appeared to decline in the BMC model region from 2000 to 2008, and both model versions showed such a decline in this period based on the backcast and base year runs of Model B and the base year and forecast year runs of Model A. The results of downstream model components are less conclusive about travel trends.

There is not enough information from this study to definitively conclude that greater accuracy in earlier components is characteristic of travel demand models in general; however, it does make intuitive sense that this might be the case. It is known that errors can propagate from upstream to downstream model components, and earlier components have fewer upstream components from which errors can propagate. It is also known that sometimes errors in different model components can offset one another, and experience with base year model validation has not shown that later model components match observed data less accurately than earlier components. (This may in part reflect the fact that there are more available independent data sources for downstream components such as trip assignment (e.g., traffic counts) than for earlier components, where sometimes it is necessary in validation to rely on information from the model estimation data set (i.e., household surveys).

Although this is not an original conclusion of this study, but the study does reinforce the need to recognize error propagation from upstream to downstream model components. It also serves as a reminder for developers of more complex models, such as activity based models, to pay close attention to the issue of error propagation since the available independent validation data for the multitude of components in these models can be much thinner than for simpler conventional models.

4.1.7 Use of Fixed Factors

Fixed factors derived from survey data or other sources are often used as parameters in travel demand models in cases where it is difficult to develop models that are sensitive to variables that could influence types of travel behavior. For example, it is still common practice in conventional four-step models to use fixed factors to convert daily trips to trips by time period.

The implication of using fixed factors is that the behavior being modeled is unaffected by the assumptions of the modeled scenario, including that the behavior does not change over time. If the modeled percentage of travel occurring in the a.m. peak period is determined by fixed factors, then time of day choice is unaffected by factors that might change from one scenario to another. So in such a case, modeled future congestion would not affect peak spreading, which is contrary to observations on congested highways.

It is obvious that the use of fixed factors can have significant effects on model results. Since it is relatively straightforward to develop these factors, it is typical to develop new factors each time a model is updated. The changes from one model version to the next can be notable. The BMC model validation data indicate that the percentage of travel occurring in peak periods increased from 2000 to 2008 by a modest amount (on the order of three to five percent). However, since both BMC model versions used fixed factors to convert daily trips to peak period trips, the comparison of results for both scenarios from either Model A or Model B shows almost no change in peak travel percentages. In a region where the peak travel changed more substantially, this could pose a problem for planners using peak period model forecasts.

It should be noted that the term "fixed factors" includes a variety of model parameters, and there are some of these factors in every travel model. These fixed factors also include constant terms in discrete choice models and K-factors in trip distribution models. So, <u>when using fixed factors, this is another motivator for checking the sensitivity of model results to the factors that will vary between planning scenarios and to determine how the constants in models contribute to the lack of proper sensitivity</u>.

4.2 Recommendations for Modelers

Based on the lessons documented in Section 4.1, the following recommendations for analysts involved in developing and validating travel demand models are offered:

1. **Model validation should include temporal validation and sensitivity testing**. This recommendation has been made elsewhere, but this study has shown that insights can be provided well beyond what can be learned only from comparisons of base year model results to observed data
2. **If possible, temporal model validation should include a backcast and/or a forecast year application**. A backcast can provide a valuable "second data point" for comparing model results to observed data and either a backcast or a forecast can help identify the effects of changes in model assumptions and procedures from the previous model version.
3. **Recognize that in a model update, changes in model procedures, assumptions, and input data can result in changes in model results** that can go well beyond changes in travel behavior over time. The effects of such changes on model results and forecasts can be examined during the model update process. These differences should be documented, and one should be prepared to deal with these differences when presenting model results.
4. **Model inputs, including networks and socioeconomic data, need to be thoroughly checked during model development and validation**. This recommendation is also not new, but this study reinforces it and has helped in understanding the possible effects of such errors.
5. **Whenever model parameters are changed or recalibrated during validation, the effects of these changes should be estimated if possible, and should be recognized** in any case. Sensitivity tests can be structured to examine such effects.
6. **Recognize the effects of error propagation from model components to subsequent components, and test for error propagation** whenever possible.

Understand that downstream components may have more error associated with them than upstream components.

7. **If it is necessary to use fixed factors or constants in models, recognize and test for the effects of model sensitivity of such factors**. When using results of models that use fixed factors, recognize the limitations associated with insensitivity to factors that are not included in the models.

5.0 Opportunities for Further Research

As noted in Chapter 4, the lessons that came out of this study were somewhat different than those that were anticipated before the study began. This was due in part to the models chosen for the study. The selection of BMC and OKI was based on a number of factors, including the availability of two model versions that the study team could install and run, the availability of model input data for the backcast year, and the ability of the agency staff to assist the study team by providing the model files and data and reviewing model results and comparisons.

Additional value could be generated by extending the study to more regions and models. Some additional considerations that could be used to choose models for further study might include:

- Regions with faster growth rates, where the differences in the amount of travel between the base years of the two model versions would be greater, and we would expect additional insights into the differences between base year and backcast/ forecast scenario results.
- Regions where there have been substantial changes in the transportation system between the base years of the previous and updated models. This would provide an opportunity to examine the sensitivity to a greater range of transportation level of service changes between the two analysis years.
- Regions where the model update included more substantial changes in model structure. The most obvious case would be where the current model is activity based and the previous model was trip based. We could compare the differences in forecasting ability between the two model types, which has been an area of great interest among modeling analysts since activity based modeling has become more the standard in large U.S. metropolitan areas.
- Regions where the level of highway congestion is greater, so that the sensitivity to differences in travel times could be better examined.
- Regions with higher levels of non-auto travel, so that the effects of forecasting transit and non-motorized travel could be better examined.
- Regions with a variety of managed lanes and toll roads, where the sensitivity to road pricing could be considered.

There are also opportunities to “dive deeper” into model results than we were able to do within the constraints of the current study. The resource constraints of the project relative to the amount of time it took to work with the agencies get the models operational for the project and to analyze the results limited the testing that could be done. Specifically, we would like to have been able to look into model results for market segments below the overall regional level. Such segments could include geographic subregions, land use area types, demographic segments such as income levels.

A “deeper dive” might also allow more detailed analysis of the sensitivity of model results to specific input variables. For example, if transit fares changed significantly between the two analysis years (as was the case with the two BMC analysis years), the mode choice model results could be examined in more detail to estimate the effects of the fare change on transit demand, perhaps by examining segments of the model where there were few other changes in transit service.

6.0 References

Allen, W.G. (2006). *Travel Forecasting Model Calibration Report*. Prepared for Maryland Transportation Administration.

Baltimore Metropolitan Council (2007). Baltimore Region Travel Demand Model Version 3.3 2000 Validation. Task Report 07-8.

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

NOTICE

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

November 2016

FHWA-HEP-17-048

U.S. Department of Transportation
Federal Highway Administration