FHWA/IN/JTRP-2008/22

Final Report

REAL-TIME SOCIO-ECONOMIC DATA FOR TRAVEL DEMAND MODELING AND PROJECT EVALUATION

Kevin M. Ford
Jon D. Fricker

July 2009

Final Report

FHWA/IN/JTRP-2008/22

REAL-TIME SOCIO-ECONOMIC DATA FOR TRAVEL DEMAND MODELING AND PROJECT EVALUATION

by

Kevin M. Ford
Graduate Research Assistant

Jon D. Fricker, P.E.
Professor

Joint Transportation Research Program
Project No. C-36-54KKK
File No. 3-3-62
SPR-3213

Conducted in cooperation with the
Indiana Department of Transportation and the
Federal Highway Administration
U.S. Deprtment of Transportation

The contents of this report reflect the views of the author, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Indiana Department of Transportation at the time of publication. This report does not constitute a standard, specification, or regulation.

Purdue University
West Lafayette, IN 47907
July 2009

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. FHWA/IN/JTRP-2008/22	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Real-Time Socio-Economic Data for Travel Demand Modeling and Project Evaluation		**5. Report Date** July 2009 **6. Performing Organization Code**
7. Author(s) Kevin M. Ford and Jon D. Fricker		**8. Performing Organization Report No.** FHWA/IN/JTRP-2008/22
9. Performing Organization Name and Address Joint Transportation Research Program 1284 Civil Engineering Building Purdue University West Lafayette, IN 47907-1284		**10. Work Unit No.** **11. Contract or Grant No.** SPR-3213
12. Sponsoring Agency Name and Address Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		**13. Type of Report and Period Covered** Final Report **14. Sponsoring Agency Code**

15. Supplementary Notes

Prepared in cooperation with the Indiana Department of Transportation and Federal Highway Administration.

16. Abstract

Small and medium-sized communities often lack data that are complete, current, representative of the community, and appropriate for the travel demand models/software to be applied. To aid such communities, this research was conducted to meet the following objectives:

(1) Identify and facilitate the communication of data that is updated more frequently than the Decennial Census for Indiana Travel Demand Models;
(2) Allow the use of cross-classification techniques within trip generation for planning organizations that lack the time and resources to conduct travel demand/household surveys;
(3) Evaluate the sensitivity of travel demand model outputs to varied parameters and inputs so as to help focus data collection and guide parameter selection; and
(4) Apply risk management strategies to travel demand modeling so as to help planners program the most resource-efficient projects.

17. Key Words Trip Generation, Input Updating, Synthetic cross-classification, Sensitivity Analysis, Risk Analysis	**18. Distribution Statement** No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	135	

Form DOT F 1700.7 (8-69)

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES x

ABSTRACT xi

CHAPTER 1. REAL-TIME SOCIO-ECONOMIC DATA FOR TRAVEL DEMAND MODELING AND PROJECT EVALUATION 1

1.1. Background and Objectives 1
1.2. Socio-Economic Variables 2
1.3. Data Source Identification 4
1.4. Data Source Recommendations 14
1.5. Communication Plan 15
1.6. Non-continuous versus Continuous Input Estimation 16
1.7. Cost-free versus Cost-associated Input Estimation 18
1.8. Conclusion 19

CHAPTER 2. RESOURCE-EFFICIENT APPROACH TO TRIP GENERATION BY CROSS-CLASSIFICATION 20

2.1. Background and Objectives 20
2.2. Literature Review 22
2.3. Creation of Synthetic Membership Matrices 23
2.4. Trip Rate Matrix Options 26
2.5. CAMPO Case Study 30
2.6. Verification of Results 35
2.7. Conclusion 37

CHAPTER 3. USING SENSITIVITY ANALYSIS TO GUIDE DATA COLLECTION IN SMALL- AND MEDIUM-SIZED COMMUNITIES 38

3.1. Background and Objectives 38
3.2. Trip Generation 41
3.3. Trip Distribution 50
3.4. Trip Assignment 53
3.5. Data Discrepancies 55
3.6. Conclusion 56

Page

CHAPTER 4. APPLYING RISK MANAGEMENT TO THE TRAVEL DEMAND MODEL 58

4.1. Background and Objectives 58
4.2. Introduction to Risk Management Practices 59
4.3. Risk Identification 66
4.4. Risk Assessment 66
4.5. Risk Management 78
4.6. Risk Monitoring 81
4.7. Conclusion 82

CHAPTER 5. SUMMARY AND FUTURE RESEARCH 83

5.1. Overall Conclusions 83
5.2. Future Research 84

LIST OF REFERENCES 84

APPENDICES

Appendix A: Travel Demand Model Parameter Databases 96
Appendix B: Input Data Availability 111
Appendix C: Recommended CAMPO Travel Demand Model Parameters 117
Appendix D: Current CAMPO Input Data 118
Appendix E: Risk Assessment Results - Top 10 Most Congested Links and VHT Values by Model run with Random Normal Inputs 123
Appendix F: Change in Inputs 132

LIST OF TABLES

Table Page

1.1 Typical Regression Variables used in Trip Generation 3

1.2 Currently Planned ACS Data Release Schedule (Chiu, 2008)..... 5

1.3 Projected versus Updated Aggregate CAMPO Travel Demand Model Inputs 17

1.4 Free versus Pay-site Aggregate CAMPO Travel Demand Model Inputs 18

2.1 Collected Trip Rate Means and Standard Deviations by Distribution 28

2.2 K-S Log-normality and Normality Test ρ-values..... 29

2.3 Trip Rate Ranges by Confidence Level for 1-Person Households..... 30

2.4 CAMPO ‘Seed Table’..... 31

2.5 CAMPO 'Marginals'..... 31

2.6 Initial CAMPO ‘Membership’ Probability Matrix..... 32

2.7 Final CAMPO ‘Membership’ Probability Matrix 33

2.8 Synthetic ‘Membership Matrix’ for CAMPO CT 105, BG 5 34

2.9 CAMPO CT 105 Trip Rates from the NHTS Transferability Program 34

2.10 Trip Productions Table with NHTS rates..... 35

2.11 Trip Productions Table with Average rates 35

2.12 Trip Production Estimates with Different ‘Seed Tables’ 36

2.13 NCHRP 365 versus NHTS Estimated Daily Person Trip Rates 37

3.1 Recommended Models and Evaluation Criteria for Each Planning Step within a Sensitivity Analysis 40

3.2 Cross-classification -- Range of Trips Predicted with Varied Parameters 42

Table Page

3.3 Cross-classification – Household ‘Membership’ Means and Standard Deviations 44

3.4 Cross-classification -- Range of Trips Predicted with Varied Inputs..... 45

3.5 Regression -- Range of Trips Produced by Purpose 46

3.6 Regression -- Range of Trips Attracted by Purpose 47

3.7 Regression -- Variable Input Means and Standard Deviations 48

3.8 Regression -- Range of Balanced Trips Predicted with Varied Inputs..... 48

3.9 Comparing Cross-Classification and Regression Techniques for Trip Generation..... 49

3.10 Average Trip Lengths in Minutes by Impedance Function 51

3.11 Total Peak-Hour VMT by Link Performance Function..... 53

4.1 Links with the top 10 v/c ratios using the Current Data..... 71

4.2 Risk Assessment Data for Links Appearing in the Top 10 Links as sorted by the v/c ratio for 10 Model Runs..... 75

4.3 CAMPO Likelihood, Consequence, and Expected Risk Values by Link 77

4.4 Recommended versus Likely CAMPO Strategies..... 81

LIST OF FIGURES

Figure Page

2.1 2001 NHTS Classification of Eligible Census Tracts (Hu et al., 2007) 22

2.2 Standard 'Membership' and Trip Rate Matrix Format 23

2.3: Kolmogorov-Smirnov Test Plot for 1-Person, 1-Vehicle Households 27

4.1 Risk Management Continuous Framework (Verzuh, 2005) 60

4.2 Defining Risk Value with Utility Curves (Sinha and Labi, 2007) 62

4.3 Risk Matrix Format (Alexander and Marwill, 2006) 63

4.4 Use of a Risk Tolerance Line within a Risk Matrix (OGC, 2008) 63

4.5 Assigning Risk Management Strategies to the Risk Matrix (Alexander and Marwill, 2006) 65

4.6 Risk Matrix Location for Sample Calculation 71

4.7 Locations for the Top 10 Links sorted by v/c ratio using Current Data 72

4.8 Location of all Links appearing in the Top 10 through 10 Model runs 74

4.9 CAMPO Risk Matrix 78

4.10 CAMPO Risk Management Recommendations 79

ABSTRACT

Small and medium-sized communities often lack data that are complete, current, representative of the community, and appropriate for the travel demand models/software to be applied. To aid such communities, this research was conducted to meet the following objectives:

(1) Identify and facilitate the communication of data that is updated more frequently than the Decennial Census for Indiana Travel Demand Models;

(2) Allow the use of cross-classification techniques within trip generation for planning organizations that lack the time and resources to conduct travel demand/household surveys;

(3) Evaluate the sensitivity of travel demand model outputs to varied parameters and inputs so as to help focus data collection and guide parameter selection; and

(4) Apply risk management strategies to travel demand modeling so as to help planners program the most resource-efficient projects.

CHAPTER 1. REAL-TIME SOCIO-ECONOMIC DATA FOR TRAVEL DEMAND MODELING AND PROJECT EVALUATION

1.1. Background and Objectives

A forecast is only as reliable as the data on which it is based. To make accurate forecasts, planners require data that are representative of current conditions. Yet, small and medium-sized planning organizations often lack the resources needed to continuously update socio-economic and demographic input data for their travel demand models. Instead, such organizations rely on the Decennial Census, leaving the actual data values in between the censuses uncertain. To estimate current conditions, planners use growth rate and trend analysis, as opposed to increased data collection. Metropolitan Planning Organizations (MPOs) in non-attainment and maintenance areas are an exception, pursuant to federal law, such communities must update their inputs and long-range transportation plans every three to five years (23 CFR450C, Sec.450.322). The Indiana Statewide Travel Demand Model (ISTDM) relies on MPO inputs. Without accurate inputs, the Indiana Department of Transportation (INDOT) risks programming construction projects that may not be the most critical, leading to possible inefficient utilization of state resources.

Currently, MPOs are asked to regularly inform INDOT of significant changes to their socio-economic and demographic data. However, this channel of communication has proven unreliable. For instance, roadway improvement plans were developed in the vicinity of Marion, Indiana in Grant County; due to major manufacturing plant closures, however, a sudden drop in travel demand was experienced, eliminating the need for the project. INDOT was unaware of

the plant closures for a period of time and proceeded to implement the project, resulting in the inefficient expenditure of resources.

Manufacturing job losses have hit the Midwest particularly hard since the 2001 recession, leading to unprecedented home foreclosure rates. Indiana has had the second highest foreclosure rate in the nation since 2003 (Tatom, 2007). The sudden loss of employment, income, and home ownership could not have reasonably been predicted using growth rate and trend analysis from the 2000 Census. These forecasting techniques become less and less reliable as the analysis year is further removed from the 'base year', or the year at which data was collected. By the time the 2010 census data is statistically verified and released, the 'base year' data, used in many MPO models, may be twelve to thirteen years old.

To reduce the uncertainty associated with out-dated 'base year' data in the travel demand model, continuous 'real-time' data collection and communication are needed. To address this issue, this chapter is organized as follows: (1) discussion of common socio-economic inputs used in the trip generation step of the travel demand model; (2) identification of data sources that are updated more frequently than the Decennial Census is updated; (3) data source recommendations, (4) a description of a communication plan for disseminating updated socio-economic inputs; (5) an evaluation of applying non-continuous versus continuous input estimation; and (6) a comparison of the estimated inputs using cost-free versus cost associated data sources. The Columbus Area MPO (CAMPO) region will be used as a case study.

1.2. Socio-Economic Variables

Numerous socio-economic variables are used across the country for regression equations at the trip generation step of the travel demand model. After a search of nationally-available publications and with private consultant travel demand models made available for this study, a database of regression equations (Appendix 1) used in trip generation was created (Grecco et al., 1976;

Caliper Corporation, 2005; ITE, 1994; Martin and McGuckin, 1998; ODOT, 2004; ODOT, 2008; Kanafani, 1983; Proussaloglou, 2004; B.L.A., Inc., 2004; and B.L.A., Inc., 1996). Four general categories of input variables were identified, using home-based work (HBW), home-based other (HBO), non-home-based (NHB), and home-based shopping (HB-Shop) equations (Table 1.1).

Table 1.1: Typical Regression Variables used in Trip Generation

Household Counts	*Employment Counts*
Total and Group Quarters Population	Total, Basic, Service, Retail and Non-retail, Industrial, Construction, Manufacturing, Wholesale, Public Administration, Educational, Health, and Social Services, and Arts, Entertainment, Recreation, Accommodation, and Food Services Employment
Average Number of Children Under Age 5	Average Number of Employees per Household
Number of Total, 1, 2, 3, 4, and 5+ Person Households	Total Labor Force
Average Household Size	Average Number of Labor Force per Household
Number of Vehicles Available	
Number of 0, 1, 2, and 3+ Vehicle Households	
Number of 0, 1, 2, and 3+ Worker Households	

Educational Enrollment	*Floor Areas*
Total, Primary, and Secondary School Enrollement	Retail, Manufacturing, Trade, Services, and Public and Quasi-public Floor Area
College Enrollment	

The most commonly used variables are *population*, *number of households*, *vehicle ownership*, *total employment*, *retail* and *non-retail employment*, and *school enrollment*. Various data source options are available to update these inputs at a frequency that exceeds that of the Decennial Census.

1.3. Data Source Identification

Several data sources exist, each with different data types, geographic levels, updating schedules, and user-access costs. The sources presented in this section are described according to these characteristics, organized by data that can be accessed without cost (Census and non-Census sources), with cost (consultant companies), as well as potential data from Indiana state departments. Sources with data available at or below the county level, updated more often than the Decennial Census, are discussed briefly.

Census Sources

American Community Survey (ACS) and the ACS Public Use Microdata Sample (ACS-PUMS) (http://www.census.gov/acs/www/)

These programs were created for the purpose of filling in the gaps between the decennial censuses. The survey is updated annually, needing one to two years for statistical verification, and provides the same demographic and socio-economic data as the Decennial Census; in fact, the ACS is slated to replace the long form of the census beginning 2010. The ACS is distributed by mail to 2.5% of the nation's population each year. Those receiving the ACS are required by law to complete the questionnaire. While the ACS places weights on survey responses to come up with imputed estimates, the original survey responses at both the individual and household levels can be found in the PUMS files.

The availability of ACS data depends on the population of the geographic level being sought. Due to the lack of five-year estimates, data are available only for school districts or counties with populations greater than 20,000. For all other areas, the smallest geographic level available is the Public Use Microdata Area (PUMA), a contiguous area with a population greater than 100,000. With five-year estimates becoming available in 2010, according to the release schedule in Table 1.2, the smallest level of geography is expected to eventually match that of the Decennial Census.

Table 1.2: Currently Planned ACS Data Release Schedule (Chiu, 2008)

	Release Year (Year Data was collected)							
Population	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*
> 65 k	(2004)	(2005)	(2006)	(2007)	(2008)	(2009)	(2010)	(2011)
> 20 k				(2007)	(2008)	(2009)	(2010)	(2011)
All						(2009)	(2010)	(2011)

Planners must choose between timeliness versus stability when selecting one-, three-, or five-year estimates. The one-year estimate is less stable (more influenced by short-term trends) but more timely; the five-year estimate is more stable (trends are masked by taking an average across a larger sample collected over a longer time-frame) but less timely (Blodgett and Burson, 2004).

Economic Census (http://www.census.gov/econ/census02/)

This data, released every five years, provides employment, establishment, and sales data at the place level (areas with a population greater than 2,500).

Population Estimates Program (http://www.census.gov/popest/estimates.php)

This annual program has been designated the official estimate for population and housing units at the county level. All Census data sources use these values as control totals.

CenStats Databases (http://censtats.census.gov/index.shtml)

These databases provide general demographics, building permits, international trade, and race-based statistics. Of particular interest for trip generation are the following three databases.

- USA Counties

  This program reports data from the Decennial Census, Economic Census, and Population Estimates program for every county in the United States. The program also allows for the comparison of each county to national averages.

- County Business Patterns

  This program provides the number of establishments, employees, and first quarter and annual payrolls as categorized by the North American Industry Classification System (NAICS) at the zip-code and county levels. All data is updated annually with an approximate one to two year verification lag. Comparisons to surrounding zip-codes and counties are also available.

- Building Permits

  This program provides construction cost and number of building permits issued by household size at the place or county level. This data is updated monthly with a typical one month lag.

Longitudinal Employer-Household Dynamics Program (LEHD) and Local Employment Dynamics (LED) Data Tools (http://lehd.did.census.gov/led/)

The LEHD program was created for the purpose of combining federal and state employment data with the Decennial Census. The LED program was then created to further combine local data with the LEHD. The LED is a voluntary program with participants such as Indiana's Department of Workforce Development. Several data tools are available at the LEHD/LED website including the following.

- Community Economic Development (CED) HotReport

  This report compiles data at the county level from several sources, including the Census programs (Decennial Census, ACS, Population Estimates Program, and USA Counties) and non-Census programs (Home Mortgage Disclosure Act (HMDA), Common Core of Data (CCD), Quarterly Census of Employment and Wages (QCEW), and County Business Patterns (CBP)).

- Quarterly Workforce Indicators (QWI)

  As the name implies, this data source provides quarterly employment data broken down at the metro level by sex, age, and NAICS

employment type. Other useful data from this site include net job flow, job creation, new hires, separations, turnover, and average monthly and new hire earnings. The main disadvantage to QWI is that the data does not include self-employed individuals.

- OnTheMap

  This source basically is a visual representation of QWI data. Users can interact with a map of the United States to show layers of employment data with points and thermal imaging scaled to show the densest locations for jobs and residencies.

- Industry Focus

  Similar to the previous two LEHD sources, this source focuses on metro-level quarterly employment totals and percent changes over time. The source is also unique in that it reports the top local NAICS sub-sector industries based on QWI data variables.

Non-census Sources

National Household Travel Survey (NHTS) (http://nhts.ornl.gov/)

The NHTS is a survey funded by the Federal Highway Administration (FHWA) to collect demographic, socio-economic, and traveling behavior data every 5 to 7 years. Estimates for the number of households and vehicles are available at the TAZ and census tract levels. The estimates are based on travel diary responses from groupings of 'similar' census tracts depending on their land use and average household income (Hu et al., 2007). For specific local data, the NHTS allows communities to participate in their 'add-on' program, as the Indianapolis MPO has purchased for the 2008 NHTS. This 'add-on' costs roughly $175 per household.

Quarterly Census of Employment and Wages (QCEW) (http://www.bls.gov/cew/data.htm)

The QCEW is the basis for the LEHD program and replacement for the former ES-202. This source developed by the Bureau of Labor Statistics provides quarterly and sometimes monthly NAICS employment data, depending on the business sector. As with the QWI, self-employed individuals are not included.

Foreclosures.com (http://www.foreclosures.com/)

This site lists foreclosures and pre-foreclosures (the grace period between initial delinquency and foreclosure) at the zip code level, providing the street the residence is on, why it was foreclosed, number of bedrooms and baths, and home price. This database is updated daily and could be useful for updating housing totals. Similar websites exist including RealityTrac.com which the Indiana Department of Workforce Development uses.

Pay-Sites and Consultants

DemographicsNow (http://www.demographicsnow.com/)

This company offers to perform trend analysis, make five-year forecasts, and develop executive reports with 'census variable' data, traffic counts, and marketing data for different business types at the block group level, for a price. Different packages are available. As of November 6, 2008, the basic package costs $159/month with a minimum four month subscription and includes census variables and reports. The pro-plus monthly package costs $329/month with a minimum four month subscription and provides traffic counts in addition to the basic package. Annual packages are likewise available ranging from $1,295 to $3,695. The state demographer's office located at Indiana University maintains a subscription. Smaller MPOs may have more difficulty in affording such a package.

SimplyMap (http://www.simplymap.com/index.php#)

SimplyMap is a pay-site, providing data similar to the LEHD – OnTheMap tool. From this program, census data variables are reported via a map interface (with thermal imaging) of the United States at the block group level. Data can be downloaded in tabular form for the current year, as well. SimplyMap data partners include Easy Analytic Software, Inc. (EASI), Applied Geographic Solutions, Inc. (AGS), and Mediamark Research, Inc. (MRI). These consultants specialize in collecting demographic and marketing data. The cost of a subscription depends on the number of concurrent users and the package selected (e.g., the standard package, which combines data from all partners, with one user costs approximately $8,000 as of April 2009; individual packages, from selected partners, with one user costs anywhere between $1,000 and $2,000). The cost of adding a concurrent user varies from approximately $500 to $1,000.

InfoUSA (http://list.infousa.com/)

InfoUSA is primarily a marketing service, specializing in the building of business and consumer lists. For approximately $30 to 40/month, users can identify new businesses and homeowners within a specified radius (in miles) of the study area. With a subscription, users can also obtain email lists of participating employers.

MelissaData (http://www.melissadata.com/)

This site performs similar services to InfoUSA. As a marketing website, data concerning new businesses and homeowners are available. The price varies depending on the number of records requested. For instance, as of October 1, 2008, it would cost $27.82 to obtain the addresses of all purchasers (192 in total) of homes in Tippecanoe County over the last 30 days.

Regional Economic Models, Inc. (REMI) (http://www.remi.com/)

The software developed by REMI is currently used by INDOT to model the short-term and long-term effects that local economic and transportation impacts have on the regional economy. Previously, INDOT used IMPLAN for this purpose. REMI combines 'input-output', 'general equilibrium', 'econometric', and 'economic geography' modeling approaches. Data for the model primarily come from the Bureau of Economic Analysis, Bureau of Labor Statistics, QCEW, and various Census programs with a two-year data lag. The model is highly complex, allowing for the interactions between industries, and then compounding these changes annually, to result in a final forecast. To run this rather complex model, technical expertise is essential. If technical staff is unavailable, the use of the software for local applications can be outsourced to the REMI Corporation. The software is also expensive, with a base purchase of roughly $53,000 and another $10,000 annually to update the database. Such costs make use of this model by small and medium-sized MPOs impractical.

Indiana State Departments (http://www.in.gov/index.htm)

Data is collected by several state departments that may have cross-over interest for other departments. The sharing of such information has proven difficult, primarily due to concerns raised by the Disclosure Review Board. Due to privacy agreements with their survey participants, certain departments are reluctant to loan collected data, particularly in disaggregated form. For this reason, the Census only releases economic information at the block group level as opposed to the block level. Departments violating privacy agreements are subject to litigation. The Indiana Department of Workforce Development (DWD), however, has been able to gain access to the following data:

- Indiana Family and Social Services Administration (FSSA)

  This administration supplies monthly welfare data at the metropolitan level, including counts on Temporary Assistance for Needy Families (TANF) and the number of people collecting food stamps.

- Indiana Department of Education (IDOE)

  The IDOE supplies data concerning the number of high school dropouts and graduates, high school graduates seeking college degrees, statistics from the free and reduced lunch program, and annual school enrollment data at the metro level with a one-year lag time.
- Indiana Secretary of State (SOS)

  The SOS publicly provides the number of monthly business starts and dissolutions at the state level but may privately possess data at a smaller geographic level.
- Indiana Professional Licensing Agency (PLA)

  The number of health professionals is available at the metro level each year through STATS Indiana. The agency itself possesses further data for other licensing professions.
- Indiana State Department of Health (ISDH)

  Population data in terms of the number of annual deaths, births, and marriages at the county level with a two-year lag are available through the ISDH.
- Indiana Department of Revenue (IDOR)

  The Indiana DOR can provide residence and place of work information at the county level. Sales tax information and the number of tax returns by zip-code and income level are also available.
- Indiana Department of Local Government Financing (IDLGF)

  Currently available through the DLGF are local government budget and property tax rate data at the county level. The department is also working on a parcel database that would collect “real-time” small business data. Once the parcel database is complete, INDOT could use the DLGF as a source to update employment data.
- Indiana Bureau of Motor Vehicles (BMV)

  Annual total vehicle, passenger car, and truck registrations are available at the metropolitan level with a two- to three-year lag. This data

can be used to update the *number of vehicles* variable. If permission from the BMV could be obtained, further information regarding license plate and home address data could be sought for origin and destination studies. Some modifications to the data are necessary to ensure that the vehicles are actually available at the home address. For instance, with college students, vehicles are often used far from the registration address for part of the year.

- Indiana Department of Workforce Development (IDWD)

  The Indiana DWD has been the leading proponent of organizing a single Indiana database. For this cause, the DWD has combined their data into two websites that cross-reference one another.

  STATS Indiana (http://www.stats.indiana.edu/)

  This source is the website on which the DWD would like to build on to. Data is available from federal, state, local, and private agencies. Besides the previously mentioned state departments, data are available from the Census, Bureau of Economic Analysis (BEA), Bureau of Labor Statistics (BLS), National Center for Education Statistics (NCES), National Science Foundation (NSF), National Center for Health Statistics (NCHS), U. S. Patent and Trademark Office (USPTO), U.S. Courts, Council for Community and Economic Research (ACCRA), The Conference Board, MoneyTree.com, RealityTrac.com, and Economy.com are available. With this vast compilation of data sources, most employment, population, vehicle, and household data are available.

  Hoosier Data (http://www.hoosierdata.in.gov/)

  This website compiles quarterly employment data from the QCEW, as well as providing annual commuting patterns, and basic demographics. Furthermore, the DWD is working to create a database that will combine

their public Indiana Economic Digest and their private collection of Warn Notices.

Indiana Economic Digest (http://www.indianaeconomicdigest.com/)

This publicly available source summarizes newspaper reports throughout the state. These summaries include articles concerning business openings and closings and can be used to update employment values. Newspaper articles concerning Census/demographic shifts are also identified through the Digest. State officials are recommended to pay particular attention to shifting employment and demographics within locations where projects are scheduled for construction. To ease the burden on local Indiana planners, the DWD has offered to process and summarize key articles, pertaining to such changes, on a weekly or monthly basis.

Warn Notices

These notices require companies with 50 or more employees to announce major layoffs and closings with a minimum of 60 days notice. With this information, employment data could be updated, and situations such as the Grant County case, discussed in the 'Background and Objectives' section, could be avoided. With such employment information, the Indiana DWD has began work on a workforce integration data project, resulting in the creation of the Indiana Workforce Intelligent System (IWIS).

Indiana Workforce Intelligent System (IWIS)

This project, yet to be completed, will provide similar data as the OnTheMap program. To avoid disclosure issues, synthetic modeling will be used to estimate place of residency and job locations. This data may be of interest for origin-destination studies.

Information on data availability, for the regression variables in Table 1.1, using selected data sources are presented in Appendix 2. With a multiplicity of data options, a planner must make his/her own judgments as to their reliability. A discussion of what sources to use is herein presented.

1.4. Data Source Recommendations

When selecting data sources, a planner must consider some tradeoffs. For instance, a planner may prefer the most recent data, but this data may not be statistically verified. Likewise, a planner may prefer local data, but the collection may have an insufficient sample size.

In these cases, the author would recommend using current data that are compared to annual control values, and to use local data that is merged with national databases through Bayesian processes (Dey and Fricker, 1994; and Mahmassani and Sinha, 1981).

The most current local data sometimes come at a cost. Therefore, planners must consider their budget. With this in mind, three recommendations for local data collection and forecasting follow, depending on the budget of the planning organization.

Large Budget

With a large budget, it is recommended that the planning organization use DemographicsNow to obtain input data. This is preferred because it provides the same range of data as the census at the block group level and can be updated monthly. This data could then be run through REMI software to make forecasts.

Medium Budget

For a medium budget, it is again recommended to use DemographicsNow. To make forecasts, planners can continue to use growth rate and trend analysis.

Small Budget

Smaller planning organizations are recommended to rely on a mixture of STATS Indiana, Hoosier Data, Population Estimates, and the ACS to obtain input data. STATS Indiana and Hoosier Data are preferred, due to the advantage of being able to use local data. However, some of the STATS Indiana data has longer lag times than other data sources. Therefore, the ACS should be used to fill in any outdated data, and particularly should be used in 2010 when block group level data is expected to become available. To update the input data further, alternative sources are recommended. For instance, to update housing totals, it is recommended that planners use a combination of the CenStats – Building Permits program and Foreclosures.com or RealtyTrac.com. Planners can add the building permit results and subtract the foreclosure results to the ACS or STATS Indiana totals in order to obtain a current or 'real-time' total. For employment updates, planners can monitor local newspapers to add and subtract business openings and closures or use the Indiana Economic Digest. For business closings in particular, planners can request Warn Notice information from the Indiana Department of Workforce Development (DWD). With base-year data and updates, growth rate and trend analysis are recommended to make forecasts.

To ensure that the most recent local data is shared with INDOT officials, a communication plan is recommended.

1.5. Communication Plan

To avoid the programming of less critical projects, INDOT requires the most recent inputs for the ISTDM. The current strategy of relying on MPOs to

communicate updated inputs has proven unreliable. While this communication should not be abandoned, an alternative option is available to state officials.

INDOT can apply the same updating strategies as detailed in the previous section. One approach would be for state officials to either subscribe or obtain a license from the DWD for DemographicsNOW, so as to have monthly input values for the state. A cost-free approach would be to monitor the Indiana Economic Digest for business opening and closings, obtain Warn Notices from the DWD, and update housing totals with the CenStats – Buillding Permits program and Foreclosures.com or RealtyTrac.com.

Further improvements to the ISTDM can be obtained with improved interdepartmental communication. While the efforts of the DWD in developing a cross-section of data for STATS Indiana are commendable, more data sharing is recommended. With improved communication, barriers such as privacy laws could become less of a hindrance. For instance, through negotiation, parcel data obtained by the IDOR and vehicle registration addresses from the Indiana BMV may be available to INDOT to update state inputs.

The importance of updating inputs follows, using the Columbus Area Metropolitan Planning Organization (CAMPO) as a case study. The CAMPO study area consists of Bartholomew County, Johnson County census tract 6112, block group 3, and Shelby County census tract 7109, block group 4.

1.6. Non-continuous versus Continuous Input Estimation

CAMPO is one of several Indiana MPOs that rely on the Decennial Census with growth rate and trend projection techniques to make forecasts. To demonstrate the potential error in this approach, the strategy discussed in the 'Data Source Recommendations' section for small-budget organizations will be compared to 2007 CAMPO estimates. Rough estimates, as of 2007, have been provided for this research by CAMPO director Kent Anderson.

To begin this analysis, the distribution of inputs between the traffic analysis zones (TAZs) will be assumed to be the same as the CAMPO data

provided for this research, while the totals that are distributed to the zones will vary.

CAMPO population and household data for the county is available from the Population Estimates program as of July 1, 2007. The total employment for the county, as of September, 2008, was reported in the Indiana Economic Digest (Ladwig, 2008). Retail employment can be obtained for the first quarter of 2008 from Hoosier Data. The number of vehicles for Bartholomew County is available from the 2007 ACS. Input data for the additional block groups that make up CAMPO can be taken from the 2000 Census.

To update these values to the present (December 2008), the Indiana Economic Digest, Building Permits program, and Foreclosures.com can be used. From the Digest, knowledge of furniture store and inn closings (Showalter, 2008; and Blair, September 2008), and engineering firm and hospital openings (The Republic, 2008; and Blair, October 2008) can be used to update employment values. From the Building Permits program, it can be found that 160 homes have been built in Bartholomew County over the past year, while from Foreclosures.com it can be found that 92 homes were foreclosed over the past year.

Table 1.3.compares the results of this 'updated' approach against the 'projected' values used by CAMPO based on the Decennial Census.

Table 1.3: Projected versus Updated Aggregate CAMPO Travel Demand Model Inputs

	Projected	*Updated*
Population	77,986	77,215
Total Employment	70,215	39,265
Retail Employment	10,754	4,641
Non-Retail Employment	59,461	34,624
Number of Vehicles	58,792	57,971
Number of Households	30,380	32,235

Comparing these two sets of inputs, it is apparent that employment estimates differ greatly. This difference may have been caused by the unexpected slowing of Indiana's economy, user error when inputting the data, but also due to an inaccurate number of jobs being reported to CAMPO. The number of employees reported may correspond to the location where the paychecks are originated, instead of the location where the employees actually work. Self-employed individuals must be considered as well. Care must be taken in order to have the most current and representative estimates. To obtain the estimates, planners can rely on cost-free (no access fee) or cost-associated (with access fee) techniques.

1.7. Cost-free versus Cost-associated Input Estimation

Considering the relatively small budget of CAMPO, efficient resource allocation is essential. For this analysis, the 'updated' or cost-free data (Table 1.3) are compared against the DemographicsNow pay-site, a cost-associated database. From Table 1.4, slight differences are observed for the current year (December 2008). The DemographicsNow pay-site estimates a 2% higher population, 6% higher employment, 8% higher number of vehicles, and 12% lower number of households. Without the 'business add-on', retail employment was unavailable to be accessed through the Purdue Libraries System version of DemographicsNow.

Table 1.4: Free versus Pay-site Aggregate CAMPO Travel Demand Model Inputs

	Free	*Pay-Site*
Population	77,215	78,884
Total Employment	39,265	41,456
Retail Employment	4,641	--
Non-Retail Employment	34,624	--
Number of Vehicles	57,971	62,748
Number of Households	32,235	28,522

Planners should use their judgment when deciding if pay-site data are more representative and worth the investment. While pay-sites require an access cost, planners will save time during data collection.

1.8. Conclusion

Having the most current demographic and socio-economic data is critical to accurate travel predictions. Without accurate predictions, less critical construction projects may be selected. As demonstrated in the CAMPO case study, it can be seen that forecasting inputs from the Decennial Census is not always a reliable approach. Instead, it is recommended to continuously update input values. To update values, planners can invest in data from pay-sites, such as DemographicsNow. To ensure the updated inputs are represented in the ISTDM, closer coordination with state MPOs and state departments is recommended. A centralized state data source such as STATS Indiana can be used to pool Indiana data resources into a single location. With ‘real-time’ data available, more accurate travel demand forecasts can be made, leading to the efficient allocation of state resources.

The next chapter outlines an approach to trip generation by cross-classification that can use continuous input updates with a mathematical model to obviate the need for intensive household surveys.

CHAPTER 2. RESOURCE-EFFICIENT APPROACH TO TRIP GENERATION BY CROSS-CLASSIFICATION

2.1. Background and Objectives

The accurate forecasting of trips is essential for the safety, reliability, and efficiency of the U.S. transportation system (Hu et al., 2007). Policy makers rely on these estimates to strategically plan and finance capital investments at the local, state, and federal levels. Planning agencies in small- and medium-sized communities have a limited budget for producing such estimates. Such budgetary constraints often lead to limited data collection and force planners to rely on borrowed models. To mitigate these consequences, this research describes a resource-efficient approach that can be applied during the trip generation step of the traditional travel demand model.

For this first step, there are two primary approaches: Zone-Based and Household-Based trip generation. The former is an aggregate methodology that applies linear regression, and the latter is a disaggregate methodology that uses cross-classification tables. A disaggregate method is generally favored, particularly for trip production estimates, but not all planning organizations have sufficient resources available to conduct the desired travel demand surveys. Instead, such planners rely on regression techniques. Therefore, the goal of this research is to create an approach that mitigates the need for expensive data collection, allowing small and medium-sized planning organizations to use the cross-classification approach.

The main advantage of using cross-classification is the clarity provided by basing trips on an actual trip-making unit (the household), as opposed to aggregated geographic units (traffic analysis zones). This disaggregation is particularly important when predicting trip productions within central business

districts (CBDs). When regression is applied to dense areas, an inordinate number of trips are predicted, leading to a breakdown at the zonal level (Dobson and McGarvey, 1977). Because regression equations are typically calibrated for an entire study area, average values are used. Individual zones will vary greatly around this average, leading to the overestimation of trips in dense zones and the underestimation of trips in others. When aggregated, however, studies confirm that regression results are comparable to cross-classification (Dobson and McGarvey, 1977). This is documented in Chapter 3. Studies recommend the use of cross-classification for the aforementioned reasons, as well as for the ease of calibration (Dobson and McGarvey, 1977). Other advantages of cross-classification include its ability to account for non-linearities, the elimination of regression-related assumptions, independence of the zonal system, and the simpler interpretation of variable strength (Abraham, undated; Dixon, 2008; and Fricker, 2007).

The main disadvantages of cross-classification are the need for extensive data collection, lack of trip rates for cells without observations, assumption of variable independence regardless of their true relationship, lack of extrapolation, difficulty in adding stratifying variables, and the difficulty in choosing household groups (Abraham, undated; Dixon, 2008; and Fricker, 2007).

The method provided in this study seeks to obviate the need for intensive data collection and to identify potential trip rates to use. This can be accomplished by creating synthetic ‘membership matrices’ along with the borrowing of nationally available trip rate matrices. The term ‘membership matrices’ is used throughout this study to refer to the table of households as stratified by two classifying variables (*household size* and *vehicle ownership*).

The creation of a synthetic cross-classification approach has been attempted before, as is detailed in the ‘Literature Review’ section. This chapter describes how to create a synthetic ‘membership matrix’, discuss what trip rates can be borrowed, and present a case study showing how to apply the

recommended approach. Finally, quantitative and qualitative checks to verify the accuracy of this method are offered.

2.2. Literature Review

The idea of creating synthetic matrices has been applied by the National Household Travel Survey (NHTS, formerly known as the NPTS). The NHTS Transferability Program applies cluster and regression techniques to the national survey to create both 'membership' and trip rate matrices (Hu et al., 2007). This is accomplished by organizing survey data into clusters based on the land use and income level of each census tract. Each cluster then has a unique regression equation that uses local census tract data to predict 'membership' and trip rate matrices. For instance, the clusters from the 2001 survey can be seen in Figure 2.1.

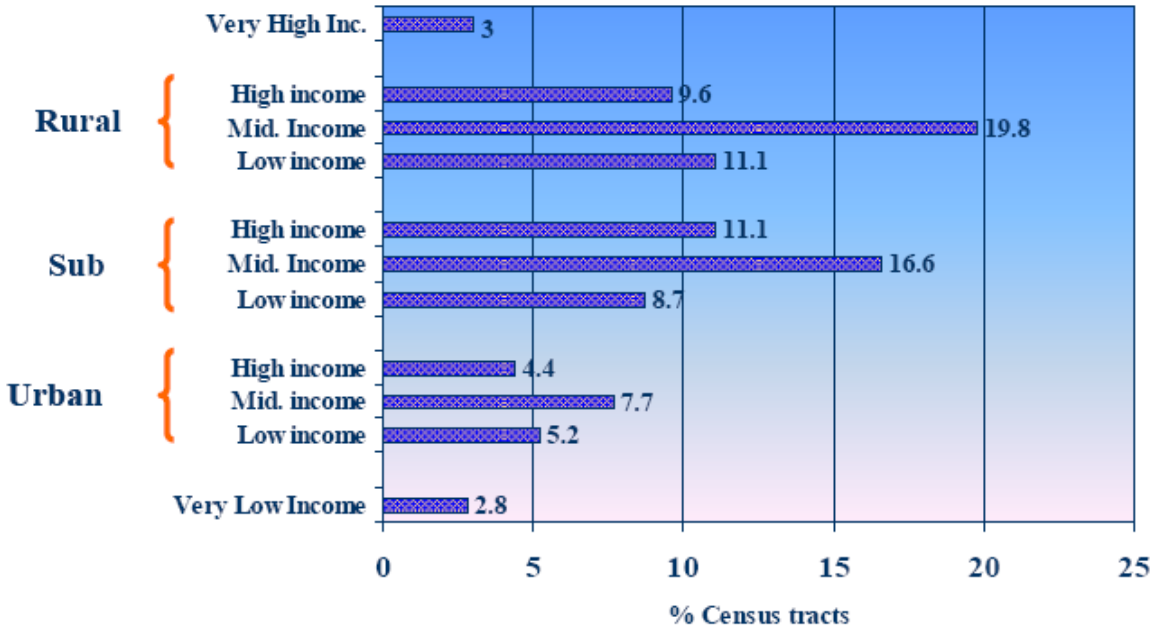

Figure 2.1: 2001 NHTS Classification of Eligible Census Tracts (Hu et al., 2007)

While the goal of this chapter is the same as the NHTS Transferability program, a different approach is recommended for creating household 'membership matrices'. This is due to the concern about the error created by

grouping tens of thousands of U.S census tracts into only ten clusters based on only two variables. Local data is preferred over this reliance on such national clusters. The 'membership matrices' created by the NHTS Transferability program, are also only singly-constrained to Decennial Census data. While the total number of *households by household size* matches the Decennial Census totals, the total number of *households by vehicle ownership* does not. The NHTS cross-classification method relies on regression, as well, which has its own disadvantages. Furthermore, these 'membership matrices' can only be updated as often as the national household survey is performed, typically anywhere from five to seven years.

This study demonstrates how to produce similar synthetic household 'membership matrices', but by using local data that can be updated, a doubly-constrained mathematical model, and at a smaller geographic level.

2.3. Creation of Synthetic Membership Matrices

To create a synthetic 'membership matrix', a mathematical method is used to merge the data sources for the 'seed table' and 'marginals', as defined in Figure 2.2.

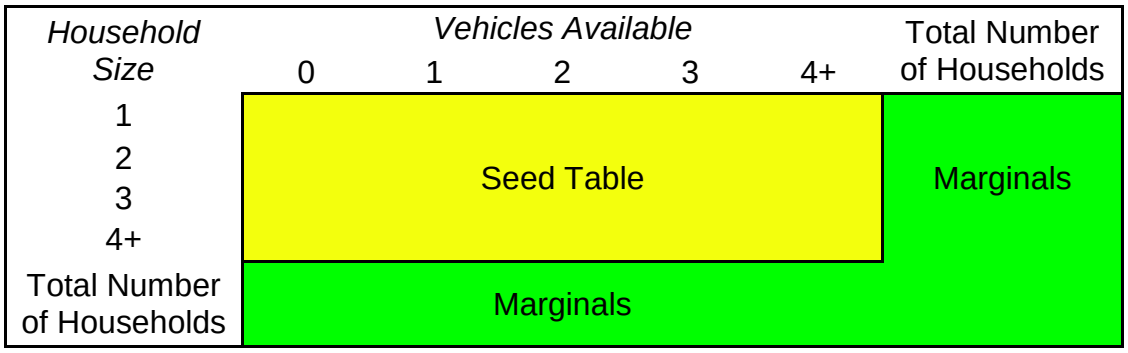

Figure 2.2: Standard 'Membership' and Trip Rate Matrix Format

The 'seed table' represents a starting point for how a planning organization believes the households in their study area are distributed. This 'seed table' can

be based on any preliminary travel surveys completed or by borrowing nationally available data. The 'marginals' represent the geographic area's totals for each row and column. With a 'seed table' and the 'marginals', a mathematical method can then be applied to modify the 'seed table' so as to produce row and column sums that match the 'marginals'.

The 'marginals' are best obtained via the U.S. Census tool called American FactFinder (factfinder.census.gov/). Using data sets H16 and H44 from Summary File 3 of the Decennial Census, aggregate totals are available for the number of households broken down by household size and by vehicles available, respectively. These totals could also be updated over time as discussed in Chapter 1.

Another potential 'marginal' variable is *household income*, although the *vehicles available* variable is preferred. This is because the two variables are highly correlated. Households with more income tend to own more vehicles. Survey respondents are also more likely to share vehicle ownership information on travel surveys than personal income information. Furthermore, *vehicle ownership* is a consistent variable between different geographic levels. Due to the volatility of income across geographic levels, the ACCRA Cost of Living Index would have to be applied to any borrowed data (Hu et al., 2007).

With the 'marginals' readily available, the only other data needed is that for the 'seed table'. The Decennial Census does not give details about how the 'marginals' are distributed, therefore another source is needed. Several options are available, including using locally collected survey responses, borrowing a 'membership matrix' from another planning organization, through the ACS, or by using the aforementioned NHTS Transferability Program. While a local survey is generally preferred, budget constraints do not always allow this choice. Instead, this study focuses on using ACS data, which are cost-free. While this data has a relatively low sampling rate, it is considered a starting point and could even be combined with local data using Bayesian methods (Dey and Fricker,1994; Mahmassani and Sinha, 1981).

Through ACS dataset B08201, data on *household size* by *vehicles available* at the county, Public Use Microdata Area (PUMA), or school district geographic level are able to be accessed. While these geographic levels are larger than the desired block group level, it makes sense to rely on local 'membership' information as opposed to clusters of national data. With the county being a smaller geographic level than the PUMA for counties with less than 100,000 people, ACS, as opposed to PUMS data, is recommended for the 'seed table'.

The data for the 'marginals' and 'seed table' need to be prepared for merging. Because the data come from different geographic levels, the 'seed data' will not add up to the 'marginal' values. Therefore, the two each need to be normalized by converting their data to probabilities.

A mathematical model is needed to merge the 'marginals' and 'seed table' together. To have both the row and column sums equal their corresponding 'marginal' values, a doubly-constrained model is needed. Two such models available are the Average Growth Factor Method and Fratar's Method (de Dios Ortúzar and Willumsen, 2001 and Fratar, 1954). These two methods have been applied during trip distribution, but they can also be used to create the synthetic matrices. Both methods produce the same results; however Fratar's Method does so in fewer trials. The merged result will be the final probability matrix, where the 'seed' probabilities sum to the known 'marginal' probabilities within a tolerance level determined by the planner. To obtain the final synthetic 'membership matrix', all that is needed is to multiply the probabilities by the number of households within the desired geographic area.

To obtain the final trip production estimates, a trip rate matrix is needed. With the 'marginals' of a trip rate matrix being unknown without travel demand surveys, a different approach is needed. A discussion of borrowing trip rates, as well as using a distribution for each cell of a trip rate matrix follows.

2.4. Trip Rate Matrix Options

Each of the collected trip rate sources in Appendix 1 has its advantages and disadvantages to borrowing (Purvis, 1994; Marshment, 1995; Martin and McGuckin, 1998; and Hu et al., 2007). With the majority of the rates coming from larger study areas (e.g., Seattle and San Francisco), transferability may be a concern, but the rates are based on larger sample sizes. The NHTS Transferability Program clusters trip rates for 'similar' census tracts and relies partially on local data, but uncertainty lies in the clustering process (Hu et al., 2007). Regardless, the NHTS rates are recommended. However, if a planner is more interested in a combination of these sources, then a distribution can be developed.

Instead of adopting a single borrowed set of trip rates, the multiple sources collected in Appendix 1 can be used to develop distributions for each cell of the trip rate matrix. With a distribution, planners could then identify the volatility of each cell and develop a range of acceptable trip rates based on different confidence levels.

To determine which distribution best fits the collected trip rates in Appendix 1, the Kolmogorov-Smirnov (K-S) Test is recommended (Kirkman, 1996). Using this 'goodness of fit' test, a specified distribution's cumulative frequency distribution can be tested against the collected data. The first step in performing the K-S test is to fit a cumulative uniform distribution to the collected data. This is done by sorting the data in ascending order and then giving each data point an equal probability (one divided by the sample size). Next, the standard deviation and mean of the data can be applied to a known distribution, such as the normal cumulative distribution. For instance, with regard to the collection of trip rates for households with 1-person and 1-vehicle, the cumulative distributions in Figure 2.3 show that a normal distribution is fairly reasonable for the data.

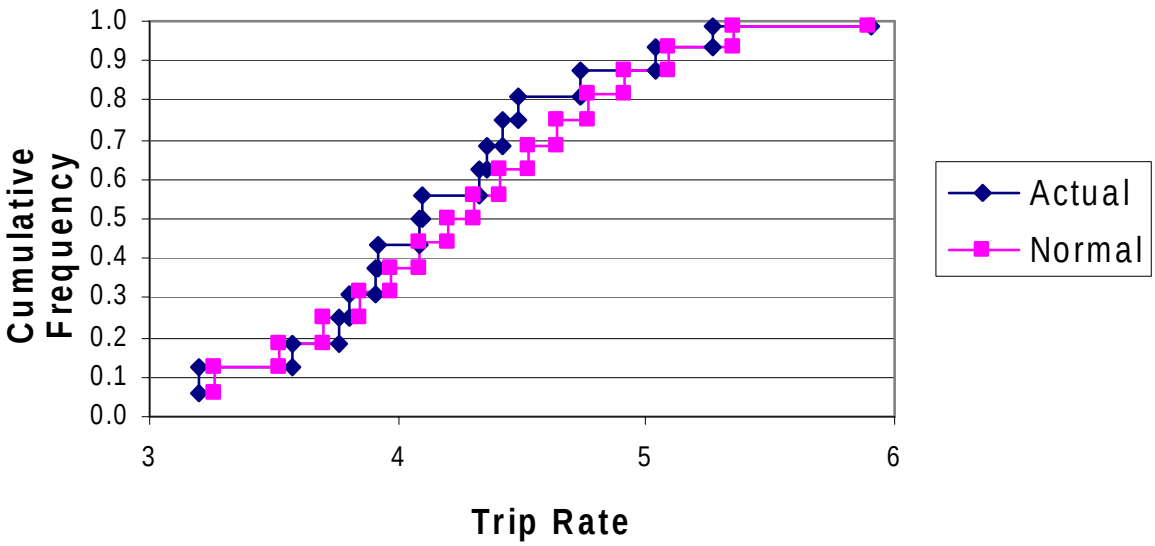

Figure 2.3: Kolmogorov-Smirnov Test Plot for 1-Person, 1-Vehicle Households

With these plots, planners can visually or statistically compare the distribution fit by calculating the corresponding ρ-value (the probability of the null hypothesis being true). In the case of Figure 2.3, the null hypothesis being tested is that the collected data is normally distributed. The larger the ρ-value, then the more confident we are in the distribution being a proper fit. A simple way to develop the plot and calculate the ρ-value is to use the on-line calculator developed by Dr. Kirkman at Saint Benedict and Saint John's University (Kirkman, 1996). The calculator requires two inputs: (1) the actual trip rates from our collection in ascending order and (2) the corresponding trip rate from a statistical distribution.

For this study, log-normal and normal distributions were tested for each cell. The log-normal distribution has been fit to average trip rates with success in previous research (Kockelman and Zhao, 2001). With normal distributions assumed by the Census Bureau for ACS 'membership' data, the normal distribution was also tested (U.S. Census Bureau, 2006). The sample means and standard deviations for the log-normal distribution and normal distributions in Table 2.1 are needed to calculate the corresponding trip rates.

Table 2.1: Collected Trip Rate Means and Standard Deviations by Distribution

Input	*Log-normal* Mean	*Log-normal* Standard Deviation	*Normal* Mean	*Normal* Standard Deviation
1-Person Household:				
No Vehicle Available	0.56	0.67	2.18	1.13
1 Vehicle Available	1.45	0.16	4.31	0.68
2 Vehicles Available	1.48	0.15	4.52	0.71
3+ Vehicles Available	1.50	0.35	4.84	1.59
2-Person Household:				
No Vehicle Available	1.12	0.56	3.73	1.86
1 Vehicle Available	1.81	0.27	6.42	1.70
2 Vehicles Available	1.94	0.21	7.13	1.41
3+ Vehicles Available	2.00	0.24	7.66	1.78
3-Person Household:				
No Vehicle Available	1.70	0.34	6.18	2.48
1 Vehicle Available	2.04	0.29	8.23	2.51
2 Vehicles Available	2.19	0.28	9.44	2.47
3+ Vehicles Available	2.32	0.26	10.58	2.38
4+-Person Household:				
No Vehicle Available	2.05	0.44	9.10	4.06
1 Vehicle Available	2.31	0.37	11.08	4.01
2 Vehicles Available	2.52	0.37	13.40	4.38
3+ Vehicles Available	2.63	0.35	14.68	4.23

Using the cumulative uniform distribution probabilities for the collected data, the corresponding ‘log-normal’ or ‘normal’ trip rate can be calculated. This is done using spreadsheet software. For instance in Excel, the LOGNORMDIST() or NORMINV() command can be used. This procedure will produce what the trip rate would be at that probability if drawn from the log-normal or normal distribution. Note, however, that because the last collected data point will have a cumulative probability of 1.00, the corresponding log-normal or normal value will be skewed. Instead a probability of 0.99 is recommended to be applied to the last data point (or highest trip rate for that cell).

The actual and normal trip rates for each cell can then be run through Dr. Kirkman’s software (Kirkman, 1996). Using this program, the plots for each cell

can be created and the ρ-value reported. The ρ-values for each cell in Table 2.2 indicate that a log-normal distribution provides the better fit.

Table 2.2: K-S Log-normality and Normality Test ρ-values

Input	*Log-normal*	*Normal*
1-Person Household:		
No Vehicle Available	0.998	0.890
1 Vehicle Available	0.999	0.912
2 Vehicles Available	0.862	0.862
3+ Vehicles Available	0.786	0.786
2-Person Household:		
No Vehicle Available	0.997	0.997
1 Vehicle Available	0.890	0.890
2 Vehicles Available	0.589	0.308
3+ Vehicles Available	0.998	0.890
3-Person Household:		
No Vehicle Available	0.828	0.828
1 Vehicle Available	0.998	0.890
2 Vehicles Available	0.890	0.890
3+ Vehicles Available	0.589	0.589
4+-Person Household:		
No Vehicle Available	0.985	0.985
1 Vehicle Available	0.890	0.890
2 Vehicles Available	0.890	0.890
3+ Vehicles Available	0.890	0.890

From this analysis, it can be seen that for nearly every cell (except for the 2-person, 2-vehicle and 3-person, 3+-vehicle cells), the cumulative distributions match with a confidence level of at least 78.6%. Therefore, based on the K-S Test, it is reasonable to assume a log-normal distribution for the majority of the trip rate cells.

With a distribution of trip rates, the planner may ask: “How sure are we that the trip rate for this cell is between x and y?” The answers to such a question can be obtained by using the mean and standard deviation for each cell of the trip rate table. For example, Table 2.3 shows the range of trip rates that a

planner can expect for 1-person households, depending on the desired confidence level.

Table 2.3: Trip Rate Ranges by Confidence Level for 1-Person Households

Confidence Level	*1-Person, 0- Vehicles*	*1-Person, 1- Vehicle*	*1-Person, 2- Vehicles*	*1-Person, 3+- Vehicles*
95%	0.47 ≤ x ≤ 6.54	3.12 ≤ x ≤ 5.81	3.26 ≤ x ≤ 5.97	2.27 ≤ x ≤ 8.80
90%	0.58 ≤ x ≤ 5.29	3.28 ≤ x ≤ 5.53	3.42 ≤ x ≤ 5.69	2.53 ≤ x ≤ 7.90
75%	0.80 ≤ x ≤ 3.79	3.55 ≤ x ≤ 5.11	3.69 ≤ x ≤ 5.27	3.01 ≤ x ≤ 6.66

*Where x represents the selected trip rate

This table shows, for instance, that there is a 95% chance that the trip rate for 1-Person, 1-Vehicle households is between 3.12 and 5.81, based on a log-normal distribution for the collected rates.

Trip rates from these distributions or a single source can be applied to a synthetic 'membership matrix' to develop trip production predictions. To demonstrate how this works, the CAMPO region is studied.

2.5. CAMPO Case Study

CAMPO, similar to other small to medium-sized planning organizations, does not have the resources required to conduct a significant number of travel demand surveys. In order to use trip generation by cross-classification, the CAMPO would need synthetic 'membership matrices'. To create such matrices, the method described in Section 2.3 is applied to Bartholomew County, IN, census tract 105, block group 5.

The first step in the process is to determine the 'marginals' and the 'seed table'. The 'seed table' from the 2007 ACS at the county level (Table 2.4) and the 'marginals' can be pulled from the 2000 Census at the block group level (Table 2.5).

Table 2.4: CAMPO ‘Seed Table’

Household Size	*Vehicles Available*					
	0	1	2	3	4+	Row Sum
1	533	3,802	876	69	60	5,340
2	436	2,002	6,084	2,476	287	11,285
3	47	1,321	1,532	1,305	559	4,764
4+	0	746	3,134	1,292	1,289	6,461
Column Sum	1,016	7,871	11,626	5,142	2,195	27,850

Table 2.5: CAMPO 'Marginals'

Household Size		*Vehicles Available*	
1	72	0	9
2	240	1	103
3	111	2	338
4	113	3	110
5	54	4	30
6	0	5+	8
7+	8		
Σ =	598	Σ =	598

Due to the fact that the ‘seed table’ and ‘marginals’ are for different geographic levels, there is a need to normalize them first. This is done by calculating the probability of each cell (dividing by the total households for the respective geographic level). For instance, there is a 72 / 598 = 0.120 chance of having a one person household at the block group level, and a probability of 533 / 27,850 = 0.019 of having a 1-person, 0-vehicle household at the county level. The initial probabilities after such normalization, in Table 2.6, show that the row and column sums for the ‘seed table’ do not add up to the ‘marginals’.

Table 2.6: Initial CAMPO 'Membership' Probability Matrix

Household Size	*Vehicles Available* 0	1	2	3	4+	Row Sum	Marginals
1	0.019	0.137	0.031	0.002	0.002	0.192	0.120
2	0.016	0.072	0.218	0.089	0.010	0.405	0.401
3	0.002	0.047	0.055	0.047	0.020	0.171	0.186
4+	0.000	0.027	0.113	0.046	0.046	0.232	0.293
Column Sum	0.036	0.283	0.417	0.185	0.079	1.000	1.000
Marginals	0.015	0.172	0.565	0.184	0.064	1.000	

Therefore, to generate the final probabilities, Fratar's Method is applied to modify the initial 'seed table' probabilities so that the row and column sums match the known 'marginals'.

The steps are as follows (Fratar, 1954; Hutchinson, 1954; and Fricker, 2007):

1.) Determine Convergence Criterion, ε and δ

For this study,

Let ε = 0.001, where:

ε ≡ Difference between Probability 'seed table' row and column sums and probability 'marginals'

Let δ = 1, where:

δ ≡ Difference between the final household totals from the merged matrix and the household totals from the 'marginals'

2.) Calculate column-specific growth factors, $F_j^k = \frac{A_j^h}{\sum_i T_{ij}^k}$

where:

A_j^h ≡ Probability 'marginals' for households with specified number of Vehicles Available

$\sum_i T_{ij}^k$ ≡ Probability 'seed table' column sum of households with specified number of Vehicles Available

3.) Calculate row-specific growth factors, $F_i^k = \dfrac{P_i^h}{\sum_j (T_{ij}^k F_j^k)}$

where:

$P_i^h \equiv$ Probability 'marginals' for households with specified Household Size

4.) Create new trip table, $T_{ij}^{k+1} = T_{ij}^k F_j^k F_i^k$

5.) Check convergence, $|\sum_i T_{ij}^k - A_j^h| \le \varepsilon \;\; and \;\; |\sum_j T_{ij}^k - P_i^h| \le \varepsilon$?

where:

$\sum_j T_{ij}^k \equiv$ Probability 'seed table' row sum of households with a specified Household Size

If satisfied, final membership probability matrix is assumed to be achieved. If not satisfied, repeat steps 2 through 4

6.) Verify that row and column sums match the 'marginals' within convergence, $n*(\sum_j T_{ij}^{k+1} - A_i^h) \le \delta$? $\quad n*(\sum_i T_{ij}^{k+1} - A_j^h) \le \delta$?

where:

$n \equiv$ Total number of households in study area

If satisfied, final 'membership matrix' is achieved. If not satisfied, repeat steps 2 through 5.

With use of spreadsheets, the final probability membership matrix is obtained in four trials, k=4 (Table 2.7)

Table 2.7: Final CAMPO 'Membership' Probability Matrix

Household Size	*Vehicles Available* 0	1	2	3	4+	Row Sum	Marginals
1	0.007	0.074	0.035	0.002	0.001	0.120	0.120
2	0.007	0.043	0.267	0.078	0.007	0.401	0.401
3	0.001	0.035	0.083	0.051	0.016	0.186	0.186
4+	0.000	0.021	0.179	0.053	0.039	0.293	0.293
Column Sum	0.015	0.172	0.565	0.184	0.064	1.000	1.000
Marginals	0.015	0.172	0.565	0.184	0.064	1.000	

The final step is to convert back from probabilities to actual households using the total household value obtained from the Decennial Census for the block group. By multiplying the final probabilities by the 598 households in Bartholomew County CT 105, BG 5, the final membership matrix is calculated as seen in Table 2.8.

Table 2.8: Synthetic 'Membership Matrix' for CAMPO CT 105, BG 5

Household Size	*Vehicles Available*						
	0	1	2	3	4+	Row Sum	Marginals
1	4	44	21	1	1	72	72
2	4	25	160	47	4	240	240
3	1	21	50	30	10	111	111
4+	0	12	107	32	24	175	175
Column Sum	9	103	338	110	38	598	598
Marginals	9	103	338	110	38	598	

With a membership matrix, trip rates can be applied.

Both the NHTS trip rates and the average trip rates from the collected database, in Appendix 1, are applied. With the trip rates for CAMPO CT 105 using the NHTS Transferability program in Table 2.9, the trip productions in Table 2.10 are predicted. For instance, with 44 one-person, one-vehicle households and a trip rate of 4.51 trips per household, 200 trips are predicted for this cell.

Table 2.9: CAMPO CT 105 Trip Rates from the NHTS Transferability Program

Household Size	*Vehicles Available*				
	0	1	2	3	4+
1	4.35	4.51	5.61	6.71	7.81
2	8.09	8.26	8.42	8.59	8.75
3	11.84	12.00	12.17	12.33	12.50
4	15.58	15.75	15.91	16.08	16.25
5+	19.33	19.49	19.66	19.83	19.99

Table 2.10: Trip Productions Table with NHTS rates

Household Size	*Vehicles Available*					Row Sum
	0	1	2	3	4+	
1	19	200	119	8	6	352
2	32	210	1,346	401	35	2,025
3	6	249	605	375	121	1,355
4+	0	218	1,909	570	427	3,124
Column Sum	58	878	3,979	1,354	589	6,856

If the average trip rates from the collected database, in Table 2.1, were applied instead, then the trip productions in are predicted as follows (Table 2.11).

Table 2.11: Trip Productions Table with Average rates

Household Size	*Vehicles Available*					Row Sum
	0	1	2	3	4+	
1	10	191	95	6	4	306
2	15	164	1,139	357	31	1,706
3	3	171	469	321	102	1,066
4+	0	137	1,438	466	345	2,387
Column Sum	28	663	3,142	1,150	482	5,465

The NHTS rates lead to more trips than the average rates, with an average daily person trip rate per household of 11.5 versus 9.1. A discussion of the reasonableness of these trip productions and of the synthetic 'membership matrix' follows.

2.6. Verification of Results

To verify the accuracy of the synthetic 'membership matrix' approach, travel demand surveys would be required. Because small and medium-sized communities cannot afford such surveys, an alternative verification approach is needed. This can be done by assessing the volatility of trip production estimates due to applying different 'seed tables'. For instance, by creating synthetic 'membership matrices' from the 'seed tables' in Appendix A of the NCHRP 365

and multiplying by the NHTS rates, the results in Table 2.12 are found to vary by only 0.1 daily person trips per household (Hu et al., 2007 and Martin and McGuckin, 1998).

Table 2.12: Trip Production Estimates with Different 'Seed Tables'

Seed Table	*Total Trips Generated*	*Daily Person Trips per Household*
Pittsburgh	299,552	10.4
Nashua	299,638	10.4
Charlotte	300,081	10.4
St. Louis	300,340	10.4
New Jersey	301,783	10.5
ACS 2007	300,668	10.5
NPTS	300,707	10.5
San Diego	300,785	10.5
Reno	301,088	10.5
Phoenix	301,184	10.5
Atlanta	301,783	10.5
Vancouver	301,905	10.5
Austin	303,108	10.5
Seattle	303,344	10.6
Range	3,792	0.1

This shows that the cross-classification results are not sensitive to the 'seed table'. Therefore, errors in the 'seed table' may be accepted without consequence in the final 'membership matrix' if the 'marginals' are known.

To decide upon trip rates selected values can be compared to locally collected data or to publications such as the *Model Validation and Reasonableness Checking Manual* or *NCHRP 365* (Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc., 1997; and Martin and McGuckin, 1998). For instance, from Table 9 of *NCHRP 365*, the selected NHTS Transferability program trip rates can be compared to national averages as classified by urban population and household size (Martin and McGuckin, 1998). The *NCHRP 365* and NHTS average trip rates in Table 2.13 differ by 2.3 daily person trips per household.

Table 2.13: NCHRP 365 versus NHTS Estimated Daily Person Trip Rates

Household Size	*NCHRP 365*	*NHTS*
1	3.7	4.9
2	7.6	8.4
3	10.6	12.2
4 (4+)	13.6	(17.9)
5+	16.6	--
Weighted Average	9.2	11.5

The reasonableness of the NHTS rates is up to the judgment of the planner. If the planner believes that the selected block group in this case is likely to have more trips than the national average, then the rates may be deemed reasonable.

2.7. Conclusion

Compared to regression, the cross-classification approach can be performed at a more disaggregate level, making the latter the preferred method. Due to insufficient resources, however, this method is not always an option for planners. Therefore, the development of synthetic 'membership matrices' (Figure 2.2) is deemed useful and likely cost-effective. Using the CAMPO case study, this chapter showed that the 'membership matrix' is not sensitive to the 'seed table' (how households are distributed by two stratifying variables). This means that local household surveys may not be necessary to be collected. The NHTS Transferability Program may be used as an alternative to collecting local trip rate information. While locally collected surveys are always preferred, the use of synthetic 'membership' and trip rate matrices is a viable alternative to using regression for small and medium-sized planning organizations.

The next chapter, will discuss the uncertainty involved in such household surveys, as well as various other inputs and parameters in the travel demand model.

CHAPTER 3. USING SENSITIVITY ANALYSIS TO GUIDE DATA COLLECTION IN SMALL- AND MEDIUM-SIZED COMMUNITIES

ASIDE: Slight modifications to this chapter have been made since being presented at the 11th National Tools of the Trade Conference organized by TRB Committee ADA 30.

3.1. Background and Objectives

Efficient data collection is critical for small- and medium-sized planning communities that lack significant resources. To expend these resources wisely, it would be helpful to know which input data (e.g., regression variables) or model parameters (e.g., coefficients in regression equations) have the greatest impact on the model output. With this knowledge, planners can better focus their efforts to refine data or parameter values. For instance, should a planning agency spend money to obtain trip length data for trip distribution or instead invest resources in obtaining volume-delay parameters to replace standard values? To help answer such questions, sensitivity analysis is used in this research.

In the literature, sensitivity analyses have often focused on how the output changes due to variations of input data within each step of the planning process. For instance, in a study by *Barton-Aschman, Inc. and Cambridge Systematics, Inc., 1997*, the effects of varying utility function inputs on mode choice outputs were observed. In the same study, the sensitivity of trip generation outputs caused by the variation of socio-economic inputs was also examined. Similarly, *Fehr and Peers, 2005* detailed the effect of varying the capacity input within the link performance function on the trip assignment output.

For the sensitivity analysis in this study, a two-tiered approach is recommended. First, the sensitivity of the output to borrowing various model parameters is discussed. By knowing which parameters are the most critical to the output, a decision can be made to decide which inputs to study further. Secondly, the output sensitivity due to varied input values is analyzed. The effects of these varied inputs can be tracked throughout the entire model and not just within the corresponding planning step. As a result of this two-tiered approach, the sensitivity of the output to model selection and input variation can be quantified.

To accomplish this, it is recommended to collect as many different parameters as possible within each of the four standard planning steps and then to apply local data. For instance, a database of regression equations or link performance functions can be built. When such equations and functions are applied to local input data, various outcomes are produced. In so doing, a range of results are created, from which a planner can make a more informed model selection, as well as learn what the consequences are from borrowing various models. The volatility of this range of outputs can help guide data collection. For instance, if the results do not vary significantly when applying various parameters, then data collection should be focused elsewhere. To further guide data collection, we can also assess the volatility of the selected model(s) by observing the change in output caused by multiple inputs.

To help demonstrate this approach, a sample sensitivity analysis will be presented using data from the Columbus Area Metropolitan Planning Organization (CAMPO), a small Indiana MPO. This study area, with a population of 73,900, as of the 2000 Census, is located approximately 46 miles south of Indianapolis.

To assess sensitivity for the CAMPO study area or any other planning jurisdiction, the following procedure is recommended at each of the standard four planning steps:

1. Build a database of equations/models with varying parameters.

2. Apply local data to develop a range of outputs.
3. Identify the sets of parameters yielding the minimum, average, and maximum outputs based on a predetermined evaluation criterion.
4. Select a model that is deemed most appropriate for the study area. For this research, two selections will be made:
    a. The 'original selection' represents the models chosen independently of this sensitivity analysis, based on the characteristics of the borrowed source;
    b. The 'revised selection' represents the models chosen based on the results of this sensitivity analysis using local data.
5. Where appropriate, apply a distribution to the input data and evaluate how the output varies based on different input confidence levels.
6. Carry on the selected, minimum, average, and maximum cases, as well as any input variations, to the next planning step.

With this procedure, it is possible to assess the impact of applying each set of parameters and input data values on the output for five different cases.

Before applying the six steps, a planner first needs to define what equations/models to assess and which evaluation criteria should be used to determine the minimum, average, and maximum cases. Such decisions left to the planners, but recommendations are given in Table 3.1.

Table 3.1: Recommended Models and Evaluation Criteria for Each Planning Step within a Sensitivity Analysis

Planning Step	*Equation/Model*	*Evaluation Criteria*
Trip Generation		
Cross-Classification	Trip Rate Matrices	Total Trips,
Regression	P and A Equations	by purpose
Trip Distribution	Friction Factor Functions, Vehicle Occupancy, and Time of Day Factors	Average Trip Length
Mode Choice	Utility Functions	Percent Auto Trips
Trip Assignment	Link Performance Functions	VMT, PRMSE

All of these steps, excluding that of mode choice, are evaluated in this case study.

3.2. Trip Generation

Planners primarily rely on two basic approaches when determining the number of trips generated for a community: (1) Cross-Classification and (2) Regression. While regression is more commonly used, both methods will be analyzed for sensitivity in this study.

Trip Generation by Cross-Classification

Assuming that a planning organization knows how its households are distributed by two stratifying variables, a trip rate matrix is then needed to predict trips. To assess the sensitivity involved in borrowing trip rates, a database of trip matrices, in Appendix 1, was created, with the vast majority coming from Appendix 2 of NCHRP 365. Next, each set of trip rates was applied to the CAMPO region household ‘membership matrix’ as stratified by household size and vehicle ownership. The resulting range of predicted trips is shown in Table 3.2.

Table 3.2: Cross-classification -- Range of Trips Predicted with Varied Parameters

Trip Rate Source	*Total Trips Predicted*	*Average Daily Person Trips per Household*	
Seattle	141,971	4.7	Minimum
Vancouver	167,824	5.5	
Houston	199,604	6.6	
New Jersey	224,525	7.4	
San Francisco	236,471	7.8	
Austin	240,091	7.9	
Nashua	241,362	7.9	
Phoenix	254,621	8.4	Average
Reno	257,395	8.5	
Albany	257,272	8.5	
Atlanta	268,325	8.8	
St. Louis	270,319	8.9	
Charlotte	282,924	9.3	
Pittsburgh	283,682	9.3	
NHTS Transferability Program – CAMPO	300,954	9.9	Original and Revised Selections
San Diego	412,983	13.6	Maximum

Depending on the rates that were “borrowed”, the number of trips predicted for the CAMPO region varies from 141,971 with Seattle trip rates to 412,983 with San Diego trip rates. The average daily person-trips per household vary between 4.7 and 13.6. Such large variations indicate that caution must be applied when deciding which of the published trip rates are most appropriate to be transferred to a study area. This may in part be due to the size of the geographic area from which the rates are transferred, but also due to sampling errors within the surveys collected in each study area. For instance, the Seattle rates from Appendix A of the NCHRP 365 appear relatively low, ranging from 3.86 to 5.60, despite the variations in household size and vehicles available (Martin and McGuckin, 1998). If local travel survey data on average trip rate is not available, the volatility in the outputs from borrowed models justifies an effort to collect local data.

Having assessed the sensitivity involved in choosing a model, the sensitivity of our selected model to variations in input data will be focused on. In this case, our ‘original’ and ‘revised’ selections happen to be the same. This is due to the NHTS Transferability program’s use of local data combined with census tracts of similar land use types and income levels (Hu et al., 2007). By quantifying how sensitive the model output is to varied input data, a planner can better guide potential data collection. To do this, it would be helpful to have statistical distributions for each cell in the household ‘membership matrix’. These distributions could then quantify the sampling uncertainty or volatility of each cell.

To establish a distribution for household membership, the American Community Survey (ACS) is recommended. The ACS provides ‘margin of error’ data which represents the sampling error at the 90% confidence level based on the assumption that the data are normally distributed (U.S. Census Bureau, 2006). With this information the standard deviation of each cell can be calculated (‘margin of error’ / 1.645) and other confidence levels explored (U.S. Census Bureau 2006). As applied to CAMPO, the mean (based on local travel demand surveys) and standard deviation for each household ‘membership’ cell can be seen in Table 3.3.

Table 3.3: Cross-classification – Household 'Membership' Means and Standard Deviations

Input	*Mean*	*Standard Deviation*
1-Person Household:		
No Vehicle Available	1,281	672
1 Vehicle Available	4,326	1,048
2 Vehicles Available	1,464	705
3 Vehicles Available	47	97
4+ Vehicles Available	52	58
2-Person Household:		
No Vehicle Available	295	453
1 Vehicle Available	2,058	1,077
2 Vehicles Available	6,871	1,220
3 Vehicles Available	1,274	552
4+ Vehicles Available	386	264
3-Person Household:		
No Vehicle Available	55	229
1 Vehicle Available	1,193	773
2 Vehicles Available	1,761	604
3 Vehicles Available	1,574	731
4+ Vehicles Available	418	371
4+-Person Household:		
No Vehicle Available	6	180
1 Vehicle Available	935	744
2 Vehicles Available	3,044	1,088
3 Vehicles Available	2,093	887
4+ Vehicles Available	1,248	519

When the standard deviation is larger than the mean, a lower bound of zero should be used to avoid applying negative inputs. With a distribution for the household 'membership matrix', the planner can start exploring how total trip predictions vary with different inputs For instance, if the lower and upper bounds of the 90% confidence interval are applied to the household 'membership matrix', how would the output change? With CAMPO data, the difference between the two bounds was found to range from 199,415 total trips to 410,482 total trips, as compared to the 300,954 total trip value found using our current data or mean (Table 3.4,).

Table 3.4: Cross-classification -- Range of Trips Predicted with Varied Inputs

Input Variable	*90% Lower Bound*	*Mean*	*90% Upper Bound*	*Range*
Total:	199,415	300,964	410,482	211,067
No Vehicle Available	2,804	7,867	20,674	17,870
1 Vehicle Available	35,810	63,767	91,724	55,914
2 Vehicles Available	102,789	134,723	166,657	63,868
3 Vehicles Available	41,265	64,723	88,404	47,140
4+ Vehicles Available	16,747	29,885	43,023	26,276
1-Person Household:	22,571	32,174	42,001	19,430
No Vehicle Available	2,804	4,932	7,060	4,257
1 Vehicle Available	15,000	18,734	22,467	7,467
2 Vehicles Available	4,739	7,848	10,957	6,218
3 Vehicles Available	0	295	813	813
4+ Vehicles Available	28	365	703	675
2-Person Household:	64,819	87,641	111,045	46,226
No Vehicle Available	0	2,211	5,003	5,003
1 Vehicle Available	9,124	16,023	22,922	13,798
2 Vehicles Available	47,503	55,628	63,754	16,252
3 Vehicles Available	6,760	10,503	14,246	7,485
4+ Vehicles Available	1,431	3,275	5,119	3,688
3-Person Household:	34,054	58,628	85,825	51,771
No Vehicle Available	0	626	3,876	3,876
1 Vehicle Available	6,402	13,700	20,999	14,597
2 Vehicles Available	14,744	20,542	26,340	11,596
3 Vehicles Available	11,537	18,671	25,805	14,268
4+ Vehicles Available	1,371	5,088	8,805	7,434
4-Person Household:	77,971	122,521	171,611	93,640
No Vehicle Available	0	97	4,735	4,735
1 Vehicle Available	5,285	15,310	25,336	20,051
2 Vehicles Available	35,803	50,704	65,605	29,802
3 Vehicles Available	22,967	35,253	47,540	24,573
4+ Vehicles Available	13,916	21,156	28,396	14,479

This wide range of outputs indicates the caution that must be taken when deciding on inputs.

To guide the data collection for these inputs, it is recommended to identify those cells that exhibit the most volatility in terms of trip productions. For CAMPO data, the most volatile cell was determined to be the 4+-person, 2-vehicle cell, which varied by nearly 30,000 trips, when applying the lower and

upper bounds of the 90% confidence interval. Therefore, data collection should be focused on households with these characteristics.

Trip Generation by Regression

For those communities without sufficient data for cross-classification, regression equations are commonly used. These equations are typically separated by trip purpose and by productions and attractions. For this research, the home-based work (HBW), home-based other (HBO), and non-home-based (NHB) trip productions and attractions will be evaluated for sensitivity.

To test sensitivity, the database of regression equations in Appendix 1 was first created. Of these regression models, all those whose input variables matched the socio-economic data available at the traffic analysis zone level in the CAMPO study area were included in the sensitivity analysis. After each regression equation was run through a software package, the productions and attractions were balanced. The results with the minimum, average, and maximum total trips, along with the 'original' and 'revised selections' are displayed in Table 3.5 for productions and Table 3.6 for attractions.

Table 3.5: Regression -- Range of Trips Produced by Purpose

Model	*HBW*	*HBO*	*NHB*
Minimum	31,170	73,825	44,736
Average	71,059	113,883	57,066
Original Selection	32,719	79,515	136,723
Revised Selection	44,088	159,914	57,066
Maximum	119,334	218,493	136,723
% Difference (Range / Maximum)	74%	66%	67%

Table 3.6: Regression -- Range of Trips Attracted by Purpose

Model	*HBW*	*HBO*	*NHB*
Minimum	31,194	41,481	44,412
Average	93,355	84,759	50,748
Original Selection	88,997	79,515	136,723
Revised Selection	47,267	167,634	136,723
Maximum	128,871	167,634	185,333
% Difference (Range / Maximum)	76%	75%	76%

The 'original selection' models are those from Madisonville, Kentucky as reported in NCHRP 167 (Grecco et al., 1976). The characteristics of this city are believed to best match those of the CAMPO region. The 'revised selection' models were selected based on the trip rates calculated for each purpose from the NHTS Transferability program (Hu, et al., 2007).

Once each case was balanced, the range of total trips predicted for CAMPO was found to vary from 133,409 to 478,194; and the range of average daily person-trips per household was found to vary from 4.4 to 15.7. As with cross-classification, this range is indicative of the care that must be taken when borrowing trip generation models. It should also be noted that the 'original selection' predicted approximately 30,000 fewer trips than the 'revised selection'. Therefore, borrowing a model based solely on the source may not lead to the most representative results.

While the output is deemed sensitive to the regression parameters, we will now assess the sensitivity of the output to the input data. For this analysis, the 'revised selection' model will be studied. As with the cross-classification method, we can create a normal distribution for each input variable with the 'margin of error' values provided by the ACS. The mean (based on local survey data) and standard deviation, calculated from the ACS, can be seen in Table 3.7 for each input variable.

Table 3.7: Regression -- Variable Input Means and Standard Deviations

Input	*Mean*	*Standard Deviation*
Total Employment	70,215	3,288
Retail Employment	10,754	4,492
# of Vehicles	58,792	3,134
# of 0 Vehicle HHs	1,736	836
# of 1 Vehicle HHs	8,575	1,610
# of 2 Vehicle HHs	12,989	1,447
# of 3+ Vehicle HHs	7,081	1,771
# of HHs	30,380	1,546
Median HH Income	$47,371	$6,042

As before, it is possible analyze how the output changes by applying data at the lower and upper bounds of their respective 90% confidence intervals. In doing so, the total trips, in

Table 3.8, were found to vary between 270,940 and 341,753 total trips as compared to the 306,346 value obtained using the current data.

Table 3.8: Regression -- Range of Balanced Trips Predicted with Varied Inputs

Model	*HBW*	*HBO*	*NHB*	*Total*	*Average Daily Person Trip Rates per Household*
90% Lower Bound	42,218	141,430	87,291	270,940	9.3
Mean	45,678	163,774	96,894	306,346	10.1
90% Upper Bound	49,130	186,123	106,501	341,753	10.8
Range	6,912	44,692	19,209	70,814	1.5
% Difference (Range/Lower Bound)	16%	32%	22%	26%	16%

This range is more acceptable than that found during cross-classification, yet a more narrow variation would be preferred.

With more focused data collection, a planner can be even more confident in the number of total trips generated. To guide the collection, it is recommended to focus on the most volatile model in terms of the range of trips predicted

between the two bounds. For CAMPO it can be determined that the HBO trip purpose is this most volatile model. This is due to the combination of HBO being the most common trip purpose and the use of retail employment as an input variable in the attraction model. Therefore, data collection should be focused on obtaining accurate retail employment values, particularly in those zones with the highest concentration of retail.

To compare the difference in outputs between cross-classification and regression, the total balanced trip results of the two approaches are shown in Table 3.9.

Table 3.9: Comparing Cross-Classification and Regression Techniques for Trip Generation

Model	*Cross-Classification*		*Regression*	
	Total Trips	*Average Daily Person Trips per Household*	*Total Trips*	*Average Daily Person Trips per Household*
Minimum	141,971	4.7	133,409	4.4
Average	254,621	8.4	223,485	7.4
Original Selection	300,957	9.9	277,096	9.1
Revised Selection	300,957	9.9	306,346	10.1
Maximum	412,983	13.6	478,194	15.7
Range	271,012	8.9	344,785	11.3

Both methods yield similar total trip results for the ‘revised selection’ case. The cross-classification method was found to be more sensitive to input variations, while the regression method was found to be more sensitive to parameter variations. With cross-classification, the sampling error involved with travel demand surveys was found to be more significant than the difference among the borrowed trip rates. To reduce the volatility due to sampling errors, the approach in Chapter 2 is recommended. With knowledge of the ‘marginals’ and applying Fratar’s Method, the ‘seed table’ or ‘membership’ cells will have less of an effect

on the output. For regression, the parameter strength was more significant than the inputs. To reduce the volatility of regression results, more data is needed to calibrate the models.

Because regression is considered to be the more common trip generation method used by small- and medium-sized planning organizations, the regression results are carried forward in the remaining planning steps for this study.

3.3. Trip Distribution

With the five cases from trip generation by regression, as well as the two input variations from the 'revised selection' case, trip distribution was then applied. To assess sensitivity, variations within the friction factor function were analyzed. In particular, the a, b, and c parameters in the Tanner form of this function, as seen in Equation 1, were varied. The impedance for this study is assumed to be the free-flow travel time between zones, t_{ij}.

$$Friction\ \ Factor = a * t_{ij}^{b} * e^{\left(c * t_{ij}\right)} \tag{1}$$

Note that the 'a' parameter has no effect when used in the gravity model and is only used to scale up the friction factors, so they are more manageable to work with. Also, the more recognizable exponential form of this function is obtained by setting 'c' to zero and 'b' to two.

After compiling a database of published Tanner functions (Appendix 1), each of the five cases was run for every collected Tanner function. The average trip length results for each equation, trip type, and case are shown in Table 3.10, with the highlighted values to be carried on to the trip assignment step.

Table 3.10: Average Trip Lengths in Minutes by Impedance Function

	Minimum			*Average*			*Original Selection*		
Equation	*HBW*	*HBO*	*NHB*	*HBW*	*HBO*	*NHB*	*HBW*	*HBO*	*NHB*
1	12.2	16.3	10.5	12.6	16.1	10.8	12.2	16.2	9.7
2	12.2	16.3	10.5	12.6	16.1	10.8	12.2	16.3	9.7
3	12.1	16.3	10.2	12.4	16.1	10.8	12.1	16.2	9.5
4	11.8	15.6	7.9	12.1	15.4	10.4	11.8	15.5	7.7
5	12.8	12.2	10.7	13.1	12	8.2	12.8	12.1	9.8
6	12.1	16.3		12.5	16.1	10.9	12.1	16.2	
7	11.9			12.2			11.9		
8	12			12.4			12		
9	11.8			12.2			11.8		
10	11.7			12			11.7		
Range	1.1	4.1	2.7	1.1	4.1	2.7	1.1	4.1	2.1

	Revised Selection			*Maximum*		
Equation	*HBW*	*HBO*	*NHB*	*HBW*	*HBO*	*NHB*
1	12.9	13.9	10.3	7.7	13.9	9.2
2	12.9	13.9	10.3	7.7	13.9	9.2
3	12.7	13.9	10	7.7	13.9	9
4	12.4	13.4	8.1	7.5	13.4	7.3
5	13.5	11.2	10.4	7.9	11.2	9.2
6	12.8	13.9		7.7	13.9	
7	12.5			7.6		
8	12.6			7.7		
9	12.4			7.5		
10	12.2			7.4		
Range	1.3	2.8	2.3	0.5	2.8	1.9

After applying the various Tanner functions to the CAMPO data, it can be seen that the HBW models vary the least, being within 1.3 minutes of each other for every case. The most volatile collection of Tanner functions is for the HBO trip purpose, which varies by up to 4.1 minutes for three of the cases. Therefore, if data is to be collected for this step, then trip length information for HBO should be obtained so as to calibrate the Tanner function for this purpose. By varying

the ‘a’, ‘b’, and ‘c’ parameters, planners may be able to better represent the locally collected travel times.

The ‘original selections’ for this step are the models taken from an FHWA study (Ismart, 1990), in part because it is more recent than some other model sources. The ‘revised selections’ were chosen based on the average HBW trip length reported for the area by the ACS, and the NCHRP 365 recommendation that the HBO and NHB trip lengths be approximately 75 and 85% of the HBW trip length (Martin and McGuckin, 1998). The total average trip lengths between the selections differ by only 1.2 minutes, however, the NHB average trip lengths vary by nearly 3 minutes.

Because free flow travel time is considered to be a reliable input, no distributions will be applied in this step. Instead, we will track the effect on the average trip length caused by the two input variations examined during trip generation by regression. By applying the results from the 90% confidence bounds to the ‘revised’ Tanner function selection, it was found that the average trip length for CAMPO varies by approximately 2 seconds for HBW; 4 seconds for HBO; 18 seconds for NHB; and 9 seconds overall. Therefore, it can be determined that the variation of trip generation inputs has a negligible effect on average trip length.

The decision to spend resources at all on calibrating the Tanner function should be based on the planner’s personal preference. If a planner is willing to accept an error of x minutes in average trip length and the variation for each trip purpose is below this value, then no resources need be expended. Average trip length is a value that is not readily available to many small MPOs. If the trip assignment step results indicate a systematic overloading or under-loading of links, and trip generation is not the reason, then the value of average trip length must be further researched.

3.4. Trip Assignment

Transit ridership is very low in CAMPO, so the mode choice step can be skipped. The user equilibrium method will be applied for trip assignment. To assess sensitivity, the α and β parameters within the link performance function, shown in the standard form in Equation 2, are to be varied.

$$t = t_0\left(1+\alpha\left(\frac{v}{c}\right)^{\beta}\right) \tag{2}$$

where: t ≡ Congested travel time,

t_0 ≡ Free-flow travel time, and

v/c ≡ volume to capacity ratio

A database of link performance functions (LPFs) applying to all link types was first accumulated (Appendix 1).

With a collection of equations, each model can be applied to the trip distribution results for the five cases and two carry-over input variations from trip generation by regression. The results of this application in terms of total VMT during the peak hour can be seen in Table 3.11.

Table 3.11: Total Peak-Hour VMT by Link Performance Function

Equation	*Minimum*	*Average*	*Original Selection*	*Revised Selection*	*Maximum*
1	56,700	128,633	109,058	132,734	201,413
2	56,832	131,128	110,479	135,033	212,673
3	56,827	132,731	111,025	137,345	232,995
4	56,775	129,571	109,586	133,416	207,642
5	56,741	129,392	109,568	133,275	207,509
6	56,805	131,362	110,377	135,295	216,534
% Difference (Range / Minimum Value)	0.2%	3.2%	1.8%	3.5%	15.7%

The results for total VMT vary from 56,700 to 232,995 between the cases. These values differ by approximately a factor of four, showing the consequences from choosing different models.

As for the selected cases, the ‘original selection’ uses the standard FHWA parameter values of α=0.15 and β=4.0, and the ‘revised selection’ was made based on VMT data obtained from the NHTS transferability program (Hu et al., 2007). The results of these selections differ by 28,287 total peak-hour VMT, which is a percent difference of 26%. This finding further highlights the consequences of making model selections based solely on the parameter source.

With regards to the variation of the LPF within each case, there appears to be only a slight effect on the total VMT. For every case except the maximum, the LPFs result in outputs that differ by less than 3.5%. This suggests that collecting VMT data for the purpose of refining the LPF parameters is not critical in most cases for CAMPO. To confirm this supposition, there is need to examine the loadings at the link level. Before the standard calculation of deviation in terms of the percent root mean squared error (PRMSE) between modeled and observed link volumes is done, a check for unusually large link flow rates ought to be made. The PRMSE for each case is not calculated here due to the lack of counts on a sufficient number of links in the CAMPO area. The results of such an analysis could even be used instead of total VMT as a basis for the ‘revised selection’.

As with trip distribution, the inputs for trip assignment, such as capacity values, are deemed reliable and therefore it would not be beneficial to apply statistical distributions. Instead, the sensitivity of the output due to applying the two 90% input bounds from trip generation was examined. In terms of total VMT, the two variations resulted in values of 121,190 and 154,669 as compared to the 137,345 value obtained using the current data. Therefore, the trip generation inputs are significant when it comes to total VMT resulting in a percent difference around 25%. This percent difference is roughly the same value, 26%, found for the percent difference of total trips generated for the two bounds. Congestion effects on route choice are likely the reason the two values differ slightly. The

difference, due to inputs, provides further proof that data collection is indeed critical at the trip generation step.

3.5. Data Discrepancies

As discussed in Chapter 1, Section 6, a large discrepancy was discovered between the provided CAMPO travel demand model employment inputs and the predicted employment inputs using the 'small budget' approach outlined in Chapter 1, Section 4 (Table 1.3). It was mentioned that this difference may have been due to unexpected slowing of Indiana's economy, user error when inputting the data, and/or an inaccurate number of jobs reported to CAMPO.

After the completion of the sensitivity analysis, outlined in the previous sections and the risk analysis in Chapter 4, it was discovered that the discrepancy occurred due to user error when inputting the data. The employment value inputted for CAMPO zone #24, in particular, was incorrect; the total employment value for zone #24 was more likely closer to 6,615 instead of 36,615. The most recent total employment value, updated after the completion of this study, is estimated to be 5,948 for zone #24. With the change in inputs, the CAMPO travel demand model aggregated totals are much closer to the predicted inputs presented in Table 1.3.

To quantify the effects of this error, the sensitivity analysis outlined in this report was conducted with updated employment values. The changes in input values are outlined in Appendix F. Using the revised inputs, the results of the sensitivity analysis changed (Table 3.12).

Table 3.12: Change in Sensitivity Analysis Results due to Revised Inputs

	Total Trips Generated		*Daily Person Trips per Household*		*Avg. Trip Length (Min)*	*Total Peak Hour VMT*
	X-Class	*Regression*	*X-Class*	*Regression*		
Minimum	0	-4,855	0.0	-0.2	1.68	8,004
Original Selection	0	-2,378	0.0	-0.1	1.52	-8,805
Average	0	-70,626	0.0	-2.3	0.01	-26,423
Revised Selection	0	-12,613	0.0	-0.4	-0.25	-25,249
Maximum	0	-99,856	0.0	-3.3	2.83	-16,399

In general, the reduction in employment resulted in a decreased number of regression trips, increased average trip length, and decreased peak hour VMT. The increased trip length is considered a result of the spreading out of trips that were originally concentrated in zone #24. The same general conclusions using the original inputs were found to still be applicable. Still, such findings demonstrate the need for the careful verification of data inputs for "reasonableness".

3.6. Conclusion

From this study, three basic lessons are learned: (1) sensitivity analyses can guide data collection, (2) it is possible to quantify output volatility by fitting distributions to input data, and (3) there can be significant consequences from selecting different models. For the CAMPO region in particular, it was learned that:

- Borrowing parameter data from communities with similar socio-economic and geographic characteristics does not guarantee that the data are the most appropriate for the area being studied.
- Cross-classification outputs are more sensitive to varied input data, while regression outputs are more sensitive to varied parameters;
- To increase the confidence in trip generation outputs by regression, careful attention to the accuracy of retail employment data should be

given, particularly in the central business district (CBD) or other employment centers.

- The variation of trip generation inputs have a negligible effect on average trip length. Because of congestion effects on route choice, the change in total VMT is slightly greater than the change in total trips generated.
- To increase confidence in trip distribution outputs, emphasis should be placed on the accuracy of HBO trip lengths, so as to calibrate the Tanner Function.
- The link performance function (LPF) parameters are not that critical to the total peak hour VMT outputs and therefore LPF calibration data may not be the most critical to obtain.

Such determinations allow for the more efficient use of resources. For this reason, among others, sensitivity analyses are strongly recommended (TRB, 2007).

With the input distributions and lessons learned in this chapter, the next chapter will focus on the how to manage the risk of programming less critical capacity-building projects due to uncertain travel demand model inputs and parameters.

CHAPTER 4. APPLYING RISK MANAGEMENT TO THE TRAVEL DEMAND MODEL

4.1. Background and Objectives

The amount of uncertainty in the travel demand model (TDM) may lead to inaccurate results. Planning organizations risk the inefficient use of resources when programming projects based on an uncertain TDM solution. To deal with this risk, management strategies can be applied.

While commonly applied in the private sector, risk management is an area of increasing research in the transportation planning field. Recent studies have predominantly focused on improving risk management techniques for 'risk due to disasters' and 'risk due to uncertainties in [project] estimation' (Sinha and Labi , 2007). Several studies have focused upon disaster evacuation (Moulton, 2007; Renne, 2006; Kalaftas, 2005; and Wolshon, 2002). Other studies have focused upon uncertainties with project construction costs, schedule, and performance (Sanderson, 2008; Damnjanovic and Zhang, 2008; and Flyvbjerg, 2006). These latter studies focus on 'project risk', which is distinctively different from 'business risk'. As defined in *Verzuh, 2005*, "Selecting the right project is business risk. Managing uncertainty to meet the stakeholder's objective is project risk."

For this research, the 'business risk' of choosing the right projects will be analyzed. By managing this initial risk, resources can be focused on projects that are considered more critical. 'Project risk' management strategies can then be applied.

To guide the selection of projects based on the TDM results, risk management strategies used in the private sector can be applied. A discussion of risk, typical management frameworks, and general strategies follows.

4.2. Introduction to Risk Management Practices

Risk can be defined as (1) “the possibility of suffering damage or loss in the face of uncertainty about the outcome of actions, future events, or circumstances” or (2) “a condition in which there exists a quantifiable dispersion in the possible outcomes from any activity” (CBS Interactive Inc., 2008). More generally, risk can be considered as “the chance that harm will actually occur” (Cefic, 2003). This concept is distinctively different from the terms ‘hazard’ and ‘exposure’. While often mistaken for risk, ‘hazard’ is “the way in which an object or situation may cause harm” and ‘exposure’ is “the extent to which the likely recipient of harm is exposed to – or influenced by – the hazard” (Cefic, 2003). Risk is the probability of harm, hazard is the instrument of harm, and exposure is the time/spatial interval during which harm may occur. For example, there is a risk of falling down a flight of stairs, but the hazard is the stairs, and the exposure is the time it takes to walk down the stairs. Without both a hazard and an exposure, there is no risk (Cefic, 2003). The identification of hazards and exposures are critical before proceeding with a risk management framework.

Risk management frameworks can be either continuous or non-continuous: a continuous approach is preferable, where risks are assessed at each phase; a non-continuous approach considers risks only once in the planning process (Carnegie Mellon University – SEI, 2008). Several variations of the continuous framework exhibited in Figure 4.1 are applied in industry.

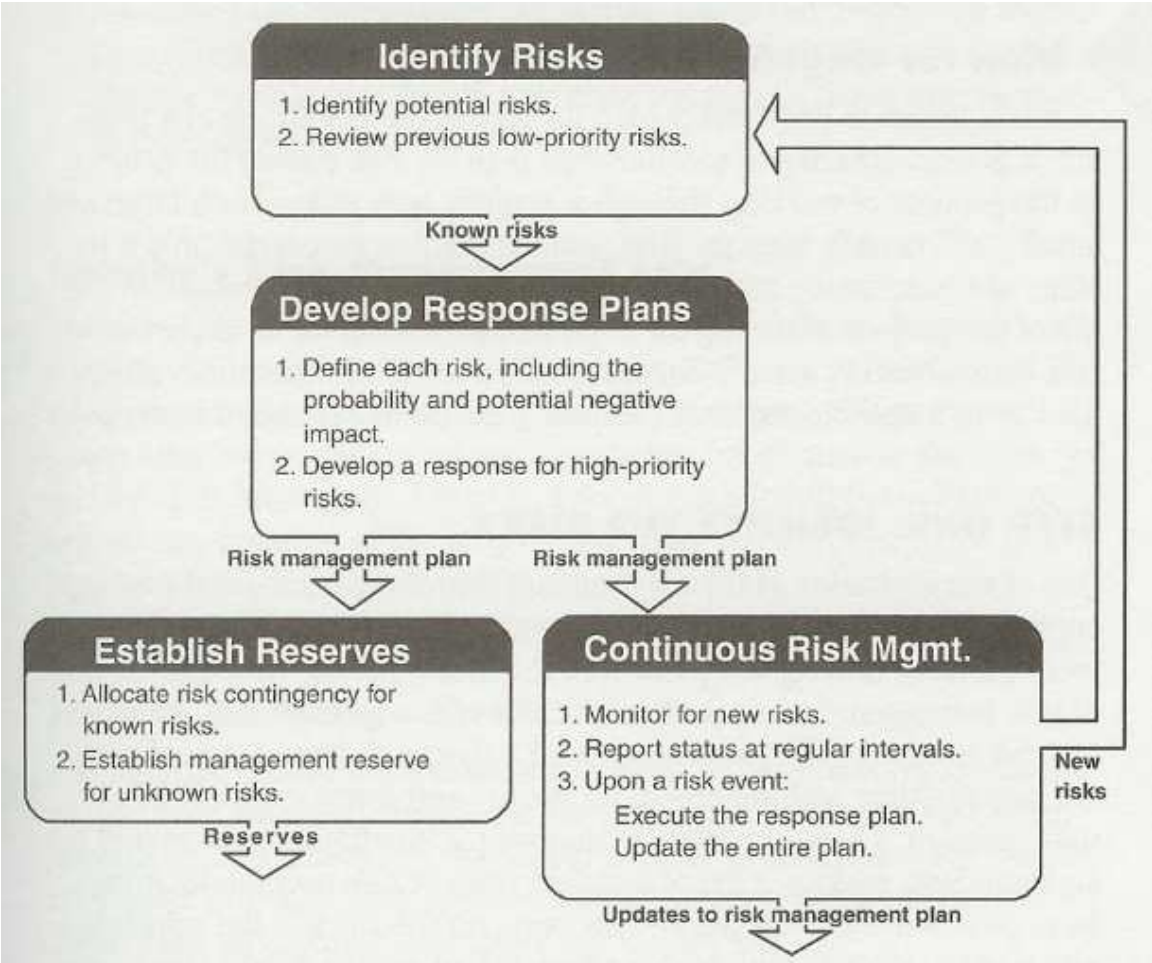

Figure 4.1: Risk Management Continuous Framework (Verzuh, 2005)

The principal steps of any risk analysis are to (1) identify, (2) assess, (3) manage, and (4) review/monitor (MID-BEDS, 2006).

Risk identification should include the description of conditions that may lead to a loss and a rough description of that loss (Carnegie Mellon University – SEI, 2008). To determine the loss more precisely, risk assessment is needed.

As described by the EPA, risk assessment is the step used to gather 'information', while risk management is the 'action' applied to reduce the occurrence or exposure to a risk (EPA, 2008). During the risk assessment step, quantitative and qualitative data are collected to identify the likelihood and consequences of 'risk occurrence' (the realization of the risk). Also during this step, the value stakeholders place on avoiding 'risk occurrence' is decided.

Quantitative methods to assess risk include: sensitivity analyses, fitting statistical distributions, forecasting, simulation, mathematical programming, and econometrics; qualitative methods include: obtaining expert opinion, determining risk value, and risk-cost-benefit trade-offs (Goodman and Hastak, 2006).

Qualitative analyses are typically subjective and depend on personal values. 'Experts' and 'the public' often perceive risks differently. For instance, *Goodman and Hastak, 2007* contend that 'the public' may be' risk neutral' to automobile accidents but 'risk averse' to earthquakes. This value scale indicates that 'the public' is more willing to accept a more frequent but expected event such as an automobile accident than a random catastrophic event such as an earthquake. 'Experts', however, may decide to expend more resources to limit automobile accidents than to protect structures against earthquakes.

The size of a decision-making body may also influence the value placed on a risk. A large organization may consider the failure of a single project as 'risk neutral', whereas a smaller organization may consider the same failure as 'risk averse', due to having fewer resources to expend for other projects. Therefore, before moving on to risk management, all stakeholders should be consulted on the value placed on the defined risk. This decision can be considered as a matter of utility. Figure 4.2 presents different risk value scales based on the utility or satisfaction obtained versus the performance level of a specified criterion.

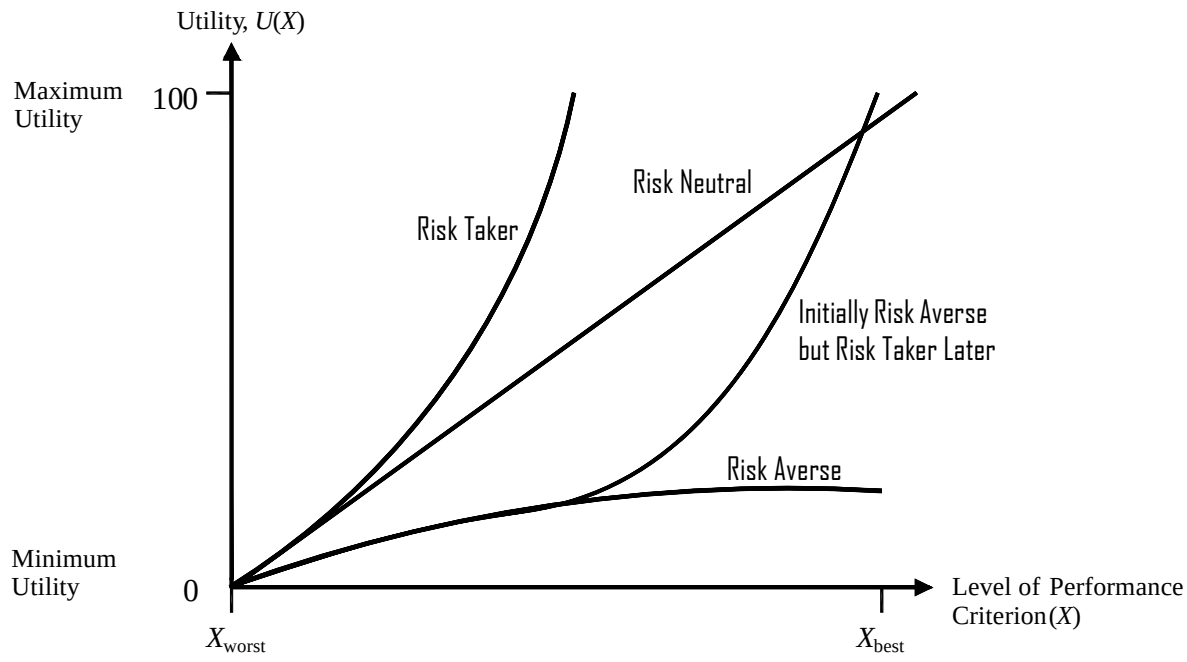

Figure 4.2: Defining Risk Value with Utility Curves (Sinha and Labi, 2007)

To prioritize which risks should be managed first, quantitative and qualitative data can be combined using the equation: Probability (or Likelihood) * Consequence = Expected Risk Value (Verzuh, 2005). Figure 4.3 is a graphical representation of a 'risk matrix' that shows the relationship between probability and consequence, and how management strategies and unknown events can change the risk matrix. For instance, the arrow in the figure represents a management strategy that reduced the consequence of the risk, and the points represent potential projects.

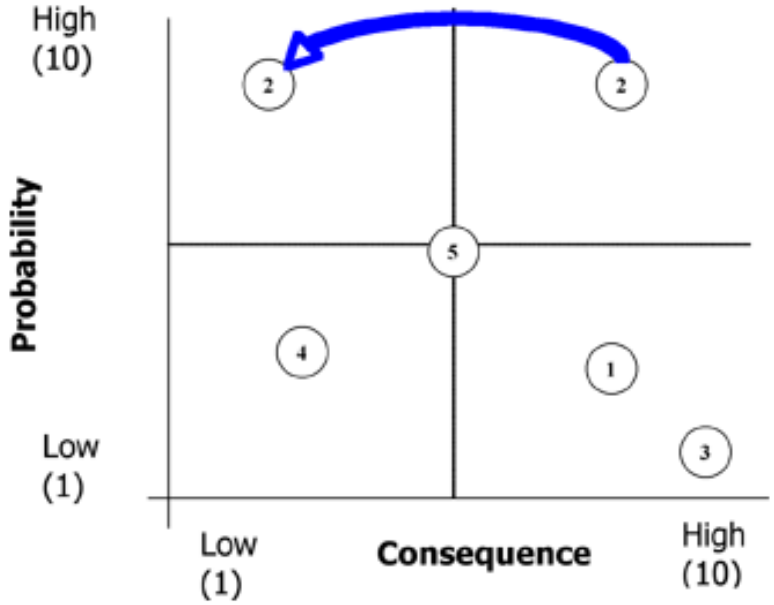

Figure 4.3: Risk Matrix Format (Alexander and Marwill, 2006)

To decide if the location in the risk matrix calls for particular management strategies, a risk tolerance line, like that in Figure 4.4, can be imposed on the matrix.

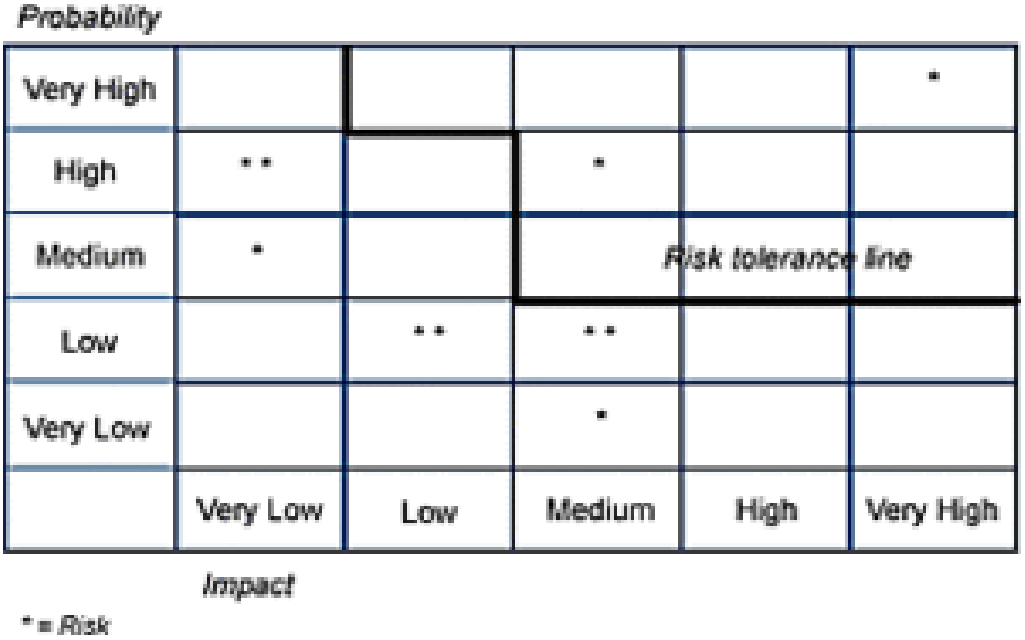

Figure 4.4: Use of a Risk Tolerance Line within a Risk Matrix (OGC, 2008)

Risk management is a term applied in a diverse range of disciplines for a diverse range of risks. For environmentalists, it may mean managing for the safe transport of oil tankers. For financial institutions, it may mean creating a contingency fund in case the rate of return of a project is lower than expected. For insurance agencies, it may mean varying coverage costs to help pay for potential damage caused by natural disasters. For factory safety officials, it may mean modifying plant processes to reduce accidents. More generally, risk management can be defined as the "means by which uncertainty is systematically managed to increase the likelihood of meeting project objectives" (Verzuh, 2005).

The most common management strategies (or action plans) typically applied are: (1) avoid, (2) reduce, (3) retain, and (4) transfer (Alexander and Marwill, 2006). These strategies are (Alexander and Marwill, 2006):

(1) Avoid ≡ Business chooses to not undertake risky activity

(2) Reduce ≡ Business takes action to reduce probability and/or consequence

(3) Retain ≡ Business accepts risk due to low consequence

(4) Transfer ≡ Business purchases insurance policy in case risk occurs

When to apply these strategies depends upon the location of the risk in the 'risk matrix' (Figure 4.5).

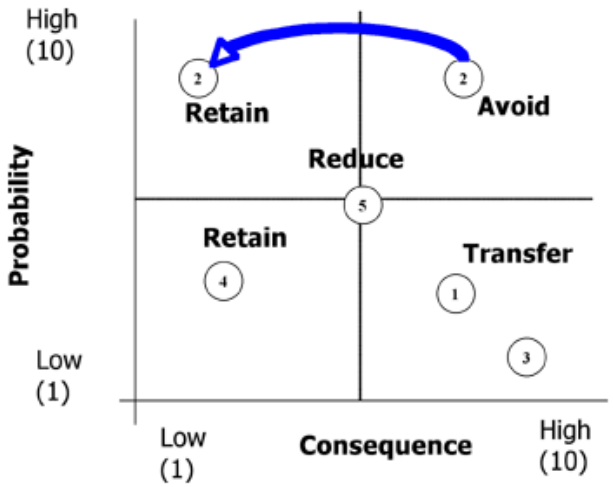

Figure 4.5: Assigning Risk Management Strategies to the Risk Matrix (Alexander and Marwill, 2006)

More specific management techniques include the use of decision rules and trees, heuristics, incremental strategy, strategic choice approach, multi-objective, multi-attribute theory and goal programming, expected utility theory, surveys, and the formulation of clearer goals, aims, objectives, and policy guidelines (Goodman and Hastak, 2006).

To pursue a risk strategy, a process must be established that allocates responsibility, facilitates compliance, and raises awareness (Durham University, 2008). Personnel should be assigned risks to assess and action plans to monitor, as well as be provided with a clear communication plan that informs others of his/her findings.

Risk logs and more regular business meetings are common ways to monitor the effectiveness of action plans for continuous risk management (Verzuh,2005). With new information, planners can efficiently plan resources, enhance focus on internal operations, establish responsibility/accountability, identify new opportunities, and modify action plans (Durham University, 2008).

These concepts are applied in this chapter with respect to project programming as guided by the results of the travel demand model.

4.3. Risk Identification

The overall 'business risk' in terms of transportation planning can be described as follows:

Risk: Resources may be applied to less critical projects based on the results of the travel demand model (TDM).

Loss: The inefficient use of limited financial and human resources on projects that do not optimize the reduction of congestion or other goals.

Conditions: TDM inputs and modeling parameters are uncertain. The input uncertainty occurs due to insufficient sample sizes and procedural bias as constrained by data collection resources. The parameter uncertainty occurs due to the lack of calibration data. Small and medium-sized planning organizations often transfer parameters that may not best fit the community.

The 'hazard' is the process of programming projects based on the results of a single TDM output. Oftentimes, this hazard is enhanced due to the outsourcing of the development of the TDM. Metropolitan Planning Organizations (MPO) risk using the model as a 'black box', being unaware of the uncertainty involved in the model development and unreasonably placing faith in the outcome. The 'exposure' is the period of time from model development to project programming. During model development, planners are exposed to the risk of using inaccurate inputs and parameters. During project programming, planners are exposed to the risk of programming inefficient projects based on inaccurate outputs.

4.4. Risk Assessment

To assess the defined risk, the following quantitative and qualitative techniques are recommended:

Quantitative

Inputs: Statistical analysis

Parameters: Sensitivity analysis

Qualitative

Inputs and Parameters: Expert Opinion for reasonableness/validation checks

Outputs: Trade-off analysis

The objective of these techniques will be to define the potential 'loss' in more detail by focusing on the peak-hour volume-to-capacity (v/c) ratios and possible reductions in vehicle hours traveled (VHT). Projects are typically scheduled for links with the most congestion, as indicated by a high v/c ratio, and by those with the greatest benefits, such as the largest reduction in VHT. This research will sort projects by v/c ratio to determine the 'likelihood' and assess by expected VHT savings to determine the 'consequence'.

To assess risk, the following procedure is recommended:

1. Unless sufficient local data has been collected, make model parameter selections by applying borrowed parameters (sensitivity analysis) and choosing those that best match local estimates (expert opinion) as demonstrated in Chapter 3.
2. Establish a 'base case' by running modeling software with current input estimates to determine the top *x* links with the highest v/c ratios. In this example, set *x* = 10, assuming that a planning organization has sufficient resources to complete 10 projects in one programming cycle.
3. Develop normal distributions (statistical analysis) for each of the trip generation data inputs at the zonal level, as demonstrated in Chapter 3, where:

   $\mu \equiv$ the mean or best estimate (expert opinion) achieved through local data collection or through data sources and updating techniques such as those discussed in Chapter 1.

   $\sigma \equiv$ the zonal standard deviation, taken as:

   $$\sigma = (\text{ACS } \sigma / \text{ACS } \mu) * \mu, \text{ with}$$

   $$\text{ACS } \sigma = (\text{ACS 'Margin of Error'} / 1.645).$$

The *ACS μ* and *ACS 'Margin of Error'* variables represent the American Community Survey (ACS) estimate and margin of error values for the smallest geographic level available in the study area.

4. Establish multiple 'variable cases' by running modeling software *n* times with randomly selected values pulled from the normal distributions established in the previous step. Record the top *x* links with the highest v/c ratios (the most congested links) and the corresponding link and network peak-hour vehicle hours traveled (VHT) values. In this example, set *n* = 10. This step is similar to a previous approach, where randomly selected log-normal values were used to construct histograms of potential peak-hour flows for links (Kockelman and Zhao, 2001).
5. Calculate the 'likelihood' of the links appearing in the top *x* being overlooked during the programming step. As the 'likelihood' increases, the risk increases. To calculate the 'likelihood' of each link, the following equations are recommended:

   Likelihood = (1 – Probability of Appearance in *n* model runs) * (Highest Link Score calculated – Link Score)

   Normalized Likelihood = 100 * Likelihood / highest of all Likelihoods calculated

   The score is taken as the sum of points awarded at the end of *n* model runs, where after each model run, the link with the highest v/c ratio is given 10 points; the second most congested link is given 9 points, and so on. Links with a lesser score and less appearances in the top *x* are more likely to be overlooked.

   For example, the likelihood of a link appearing in the top 10 (as sorted by v/c ratio), through 10 model runs, with the following breakdown can be calculated.

   1st most congested: appeared once
   3rd most congested: appeared twice
   5th most congested: appeared thrice

The Probability of Appearance = (1 + 2 + 3) / 10 = 0.60

The Score = (10 * 1) + (8 * 2) + (6 * 3) = 44

Assume another link scored a value of 90 (the highest score of all links that appeared in the top 10 during the 10 model runs)

Likelihood = (1 – 0.60) * (90 – 44) = 18.4

Assume another link had a likelihood value of 50 (the highest likelihood of all links that appeared in the top 10 during the 10 model runs).

Normalized Likelihood = 100 * 18.4 / 50 = 37

6. Calculate the 'consequence' of choosing one link over another (trade-off analysis) for capacity-building projects in terms of expected VHT savings on the link, as well as on the network. In this analysis, the capacity-building project, for simplicity, will be assumed to be the addition of one lane per direction. Using modeling software, the new link and network VHT values can be compared to the original link and network VHT values. The difference between the new and old values represents the VHT savings expected for each appearance in the top *x*. The consequence then represents the 'loss' of potential VHT savings on the link and network, if a less critical capacity-building project was constructed. To calculate the 'consequence' for each link, the following equations are recommended.

Link Consequence = average Link VHT savings – the smallest of all average Link VHT savings calculated

Network Consequence = average Network VHT savings – the smallest of all average Network VHT savings calculated

Normalized Consequence = (50 * Link Consequence / the largest of all Link Consequences calculated) + (50 * Network Consequence / the

largest of all Network Consequences calculated)

For example, assume a link with the following characteristics.

Lanes per Direction = 2

Original Peak-hour Capacity C = 338 vph

New Peak-hour Capacity C =338 * (2+1)/2 = 507 vph

Link VHT average savings = 500 VHT

Network VHT average savings = 4,000 VHT

Also assume, a link has a calculated average Link VHT savings of 50 (the smallest of all average Link VHT savings calculated) and a calculated average Network VHT savings of 100 (the smallest of all average Link VHT savings calculated).

Link Consequence = 500 – 50 = 450 VHT

Network Consequence = 4,000 – 100 = 3,900 VHT

Assume the largest Link Consequence calculated was found to be 700, and the largest Network Consequence of 7,000 was calculated.

Normalized Consequence = (50 * 450 / 700) + (50 * 3,900 / 7,000) = 60

7. Calculate the ‘expected risk value’ with

   Expected Risk Value = Normalized Likelihood * Normalized Consequence.

   For example, Expected Risk Value = 37 * 60 = 2,220.

A plot or ‘risk matrix’ of ‘likelihood’ on the ordinate against ‘consequence’ on the abscissa can graphically represent the risk value as the rectangular area bounded by the axes and a specific link point. For example, this project would be located on a risk matrix shown in Figure 4.6.

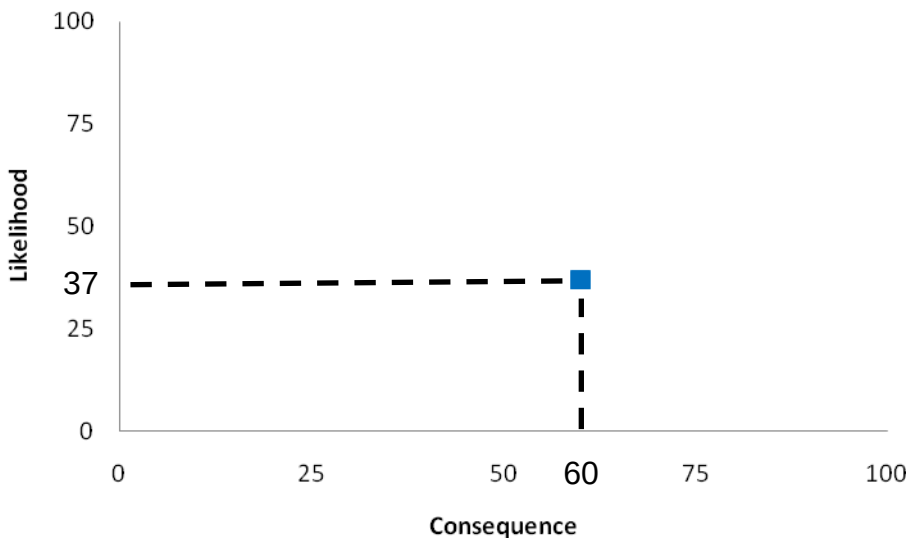

Figure 4.6: Risk Matrix Location for Sample Calculation

To demonstrate further how this procedure works, data from the Columbus Area Metropolitan Planning Organization (CAMPO) study region will be used.

The 'base case' can be established by first applying the model parameter selections in Appendix 3 (discussed in Chapter 3) to the CAMPO input data in Appendix 4. Using travel demand modeling software, the top ten v/c links in Table 4.1 would likely be recommended for project programming.

Table 4.1: Links with the top 10 v/c ratios using the Current Data

Link ID	*Street Name*	*v/c*	*Priority*
10871959	Lincoln Park Dr.	1.32	1
665041	Franklin St.	1.27	2
10873324	Lowell Rd.	1.25	3
10873327	Lowell Rd.	1.25	3
672163	Washington St.	1.23	5
664902	S.R. 46	1.21	6
10871342	S.R. 46	1.20	7
705467	Wildwood Pl.	1.20	8
10873321	Indianapolis Rd.	1.18	9
10873328	Indianapolis Rd.	1.18	9
10873325	Indianapolis Rd.	1.18	9
665831	Indianapolis Rd.	1.18	9

The locations of these links, primarily in the central business district (CBD), are shown in Figure 4.7.

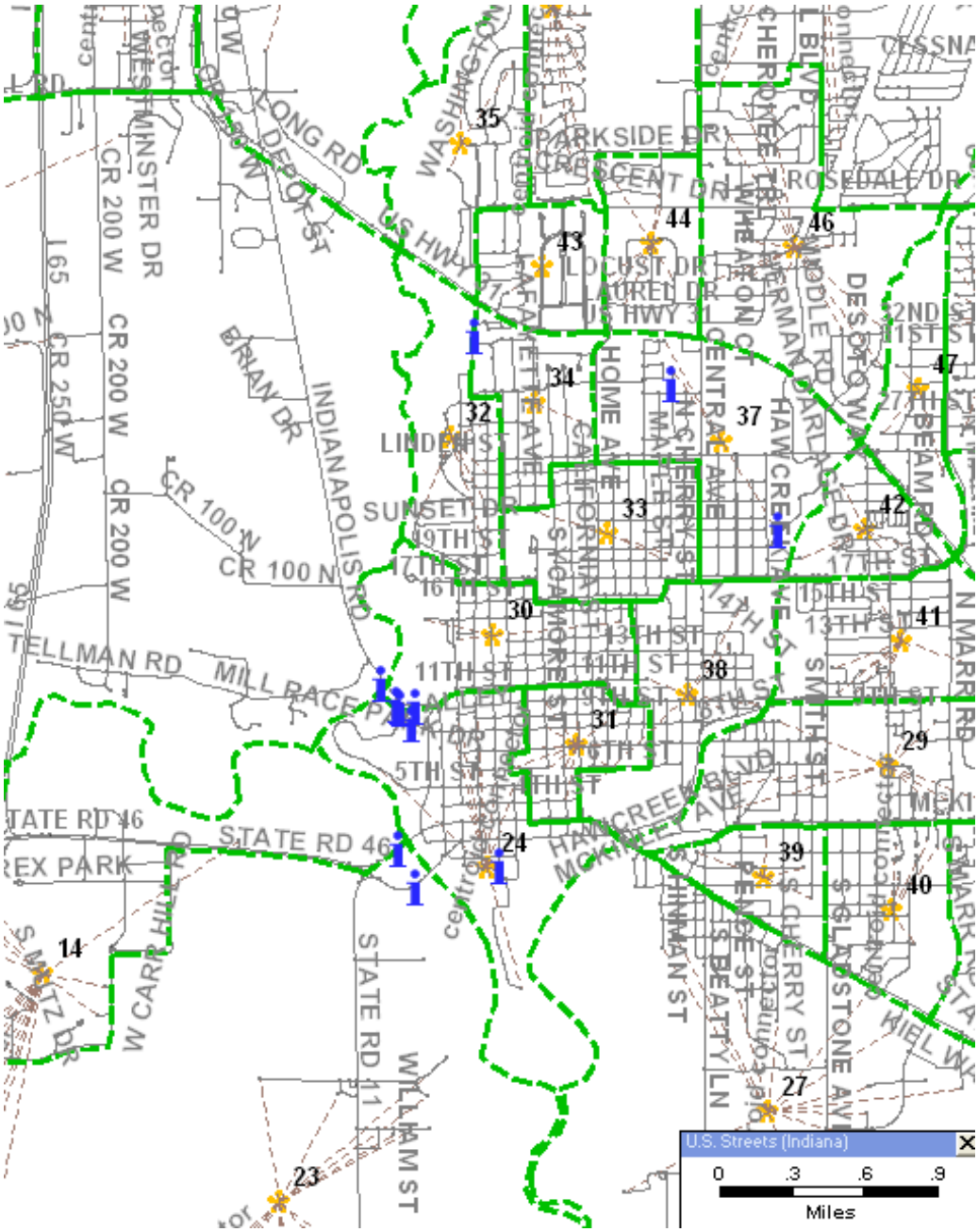

Figure 4.7: Locations for the Top 10 Links sorted by v/c ratio using Current Data

This solution, however, is just one of several possible solutions. Considering the sampling error for the inputs, multiple outputs can occur. To demonstrate the volatility of such results, statistical analysis can be used. By using input values within the sampling error, each 'variable case' represents a feasible solution. With the normal distributions developed for each input variable in Table 3.7, the 'variable case' analysis begins by randomly generating normal inputs. This can be done with spreadsheet software such as Excel or the @Risk add-on to Excel (Palisade, 1998).

In Excel, the following function is recommended to randomly generate a value from the standard normal distribution:

x = SQRT(-2*LN(RAND()))*SIN(2*PI()*RAND())

This equation, derived from the Box-Muller transformation, is preferred over traditional Excel functions because of inaccuracies in the NORMINV() and NORMDIST() functions (Kyd, 2008 and Weisstein, undated).

To scale the random value up from the standard normal distribution the *z*-value is needed. This is calculated knowing the mean and standard deviation for a standard normal distribution are 0 and 1, respectively.

$$z = (x - \mu) / \sigma$$

This equation can be rearranged, with the calculated *z*-value, zonal mean and standard deviation, to calculate random normal *x* values for each input. For those zonal inputs with $\sigma > \mu$, the lower limit that can be randomly selected should be set to zero, so as to avoid negative inputs.

The 'variable cases' are developed by applying a set of random normal values to the selected model parameters for each of the ten model runs. Having done so, the top 10 v/c links, along with corresponding VHT values, for each of the model runs can be found in Appendix 5. Twenty-one different links, as located in Figure 4.8, appeared in the top 10 for the 10 model runs. This suggests that a broader range of projects should be considered when programming.

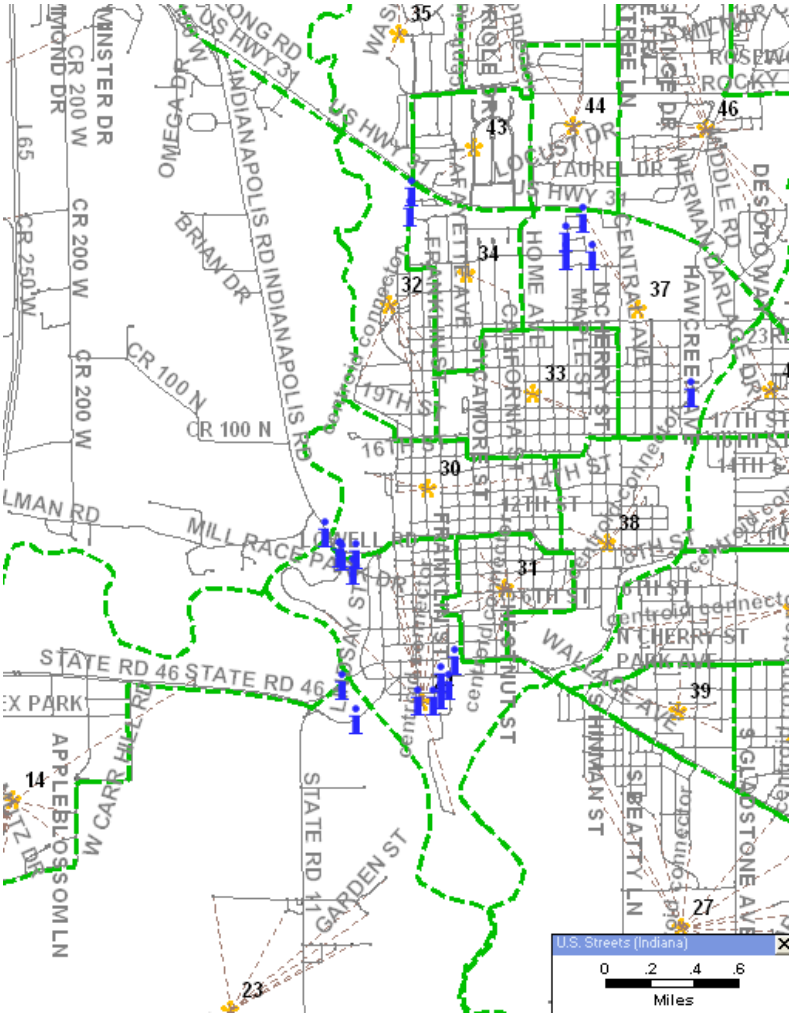

Figure 4.8: Location of all Links appearing in the Top 10 through 10 Model runs

Table 4.2 summarizes the probability of appearance, score, and average link and network VHT Savings for each of the 21 links. For example, link 10871342 (S.R. 46) appeared in the top ten in 9 out of 10 model runs. This is a probability of appearance of 90%. With one appearance as the 4th, 5th, and 9th most congested links (7, 6, and 2 points respectively) and two appearances in the 3rd, 7th, and 10th positions (16, 8, and 2 points respectively), the final score for link 10871342 is 41. The link VHT savings, over the 10 appearances, were found to be 122 (Model Run 1), 111 (Model Run 2), 348 (Model Run 4), 143 (Model Run 5), 366 (Model Run 6), 131 (Model Run 7), 258 (Model Run 8), 237 (Model Run 9), and 197 (Model Run 10). The average link VHT savings is then 213. The

network VHT savings, over the 10 appearances, were found to be 1,362 (Model Run 1), 663 (Model Run 2), 2,543 (Model Run 4), 1,785 (Model Run 5), 1,344 (Model Run 6), 1,997 (Model Run 7), 1,565 (Model Run 8), 2,112 (Model Run 9), and 1,565 (Model Run 10). The average network VHT savings is then 1,659.

Table 4.2: Risk Assessment Data for Links Appearing in the Top 10 Links as sorted by the v/c ratio for 10 Model Runs

Link ID	*Street Name*	*Probability of Appearance*	*Score*	*Average Link VHT Savings*	*Average Network VHT Savings*
664902	S.R. 46	100%	56	1,456	3,867
672163	Washington St.	60%	34	524	4,188
10871342	S.R. 46	90%	41	213	1,659
10871959	Lincoln Park Dr.	80%	70	2,789	8,131
10873324	Lowell Rd.	100%	67	881	482
10873327	Lowell Rd.	100%	66	290	149
665547	S.R. 46	20%	10	226	-371
665041	Franklin St.	50%	24	289	-193
10873321	Indianapolis Rd.	50%	13	366	210
10873328	Indianapolis Rd.	50%	15	59	-263
10873325	Indianapolis Rd.	40%	11	48	-81
665831	Indianapolis Rd.	40%	13	481	-90
705467	Wildwood Pl.	60%	38	3,715	12,005
672206	Washington St.	10%	3	448	-491
705386	Maple St.	40%	34	696	693
705492	Maple St.	40%	34	3,272	4,626
705552	Tipton Ln.	40%	34	1,529	1,619
665015	Washington St.	20%	14	346	721
665537	S.R. 46	30%	10	211	156
665434	1st St.	10%	2	358	-493
665528	Lafayette Ave.	10%	2	369	-447

The ‘likelihood’ of each link being overlooked increases with a lower probability of appearance and a lower score. Planning organizations are less

likely to consider programming projects with a high ‘likelihood’ value. The ‘consequence’ of not programming a link increases with higher potential VHT savings. The largest ‘expected risk value’, shown in Table 4.3, increases when a project likely to be overlooked is considered to have the highest savings. Resources may be inefficiently used if projects with a high ‘consequence’ are overlooked when programming. With link 10871342 (S.R. 46) the ‘likelihood’ is found by multiplying the probability of not appearing (100 – 90 = 10%) by the trade-off of a higher score being possible (70 (the highest score found) – 41 = 29). This ‘likelihood’ comes out to be 2.9, as compared to the highest calculated ‘likelihood’ of 61.2 for another link. However, when normalized on a scale of zero to 100, the likelihood becomes 5 (2.9 / 61.2 * 100).

The ‘consequence’ for the link is found by the trade-off of a higher average link and network VHT savings being possible. With the smallest average link VHT savings, calculated to be 48, and an average link VHT savings for link 10871342 (S.R. 46) of 213, the ‘link consequence’ is 213 – 48 = 165. Once normalized on a scale of 50, with the highest ‘link consequence’ of 3,668 being calculated, the ‘link consequence’ is found to be 2 as well (165 / 3,668 * 50). The smallest average network VHT savings, calculated to be -493, and an average network VHT savings for link 10871342 (S.R. 46) of 1,659, the ‘link consequence’ is 1,659 – -493 = 2,153. Once normalized on a scale of 50, with the highest ‘network consequence’ of 12,498 being calculated, the ‘network consequence’ is found to be 9 (2,153 / 12,498 * 50). The overall ‘consequence’, taken as the sum of the average link and network VHT savings, was found to be 11 (2 + 9). Multiplying the ‘likelihood’ and ‘consequence’ together yields a risk value of 51 (5 * 11, not 55 due to rounding).

Table 4.3: CAMPO Likelihood, Consequence, and Expected Risk Values by Link

Link ID	*Street Name*	*Likelihood*	*Consequence*	*Expected Risk Value*
705492	Maple St.	35	64	2,274
705467	Wildwood Pl.	21	100	2,092
705552	Tipton Ln.	35	29	1,011
665015	Washington St.	73	9	653
672163	Washington St.	24	25	593
672206	Washington St.	99	5	539
705386	Maple St.	35	14	479
665528	Lafayette Ave.	100	5	457
665434	1st St.	100	4	424
665831	Indianapolis Rd.	56	8	420
10873321	Indianapolis Rd.	47	7	333
665537	S.R. 46	69	5	332
665547	S.R. 46	78	3	230
665041	Franklin St.	38	4	169
10873325	Indianapolis Rd.	58	2	95
10871342	S.R. 46	5	11	51
10873328	Indianapolis Rd.	45	1	48
664902	S.R. 46	0	37	0
10871959	Lincoln Park Dr.	0	72	0
10873324	Lowell Rd.	0	15	0
10873327	Lowell Rd.	0	6	0

These results can also be expressed graphically with the ‘risk matrix’ in Figure 4.9, where the points represent the individual links or potential projects. The darker points represent the top 10 most congested links using the current data.

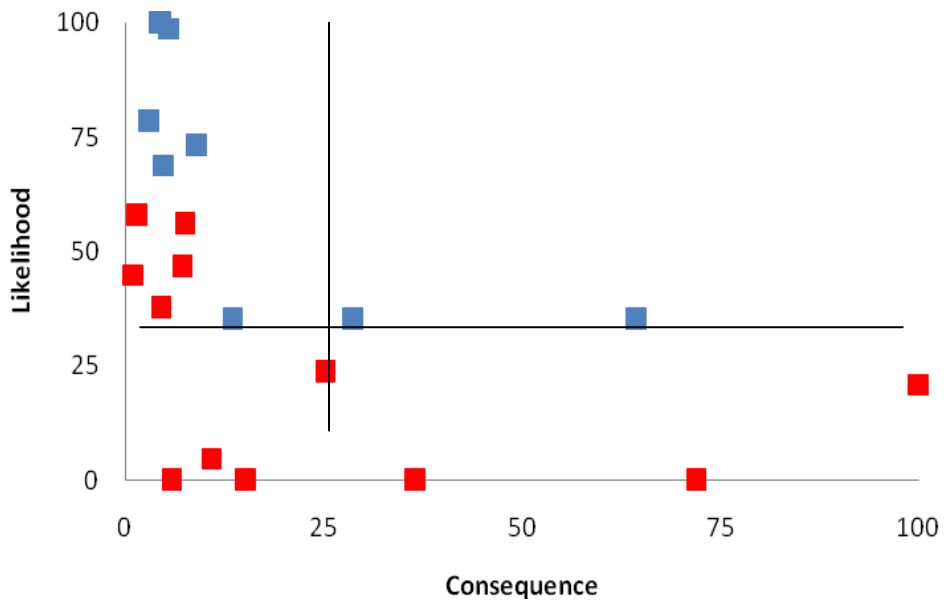

Figure 4.9: CAMPO Risk Matrix

As demonstrated in *Alexander and Marwill, 2006*, the risk matrix can be used to determine management strategies once values are established.

This research, primarily aimed at small and medium-sized communities, assumes a 'risk averse' attitude as to pursuing less efficient projects, whereas larger communities are probably 'risk neutral' due to having more resources to expend.

4.5. Risk Management

The four common risk management strategies, applied to the travel demand model (TDM) results, can be interpreted as follows:

(1) Avoid

Do not pursue the project.

(2) Reduce

Consider pursuing the project. Collect more input data using the methods detailed in Chapter 3, and perform traffic counts on the 27

links to verify the v/c ratios. Expert opinion, field observations, and improved public involvement could also be used to validate the consideration of each project for programming. The link capacity, travel times, and travel speeds entered into the TDM should also be validated.

(3) Retain

Pursue the project. This strategy represents the common practice of accepting model outputs and using expert opinion for reasonableness/validation checks.

(4) Transfer

Consider pursuing the project with contingency funds. Planning organizations can seek more state and federal funding such as for Regional Surface Transportation Program (RSTP) and/or Congestion Mitigation and Air Quality (CMAQ) Improvement Program projects (Lysy, 2005).

The location of each link in a 'risk matrix' (Figure 4.10) can be used to decide which strategy to apply (Alexander and Marwill, 2006).

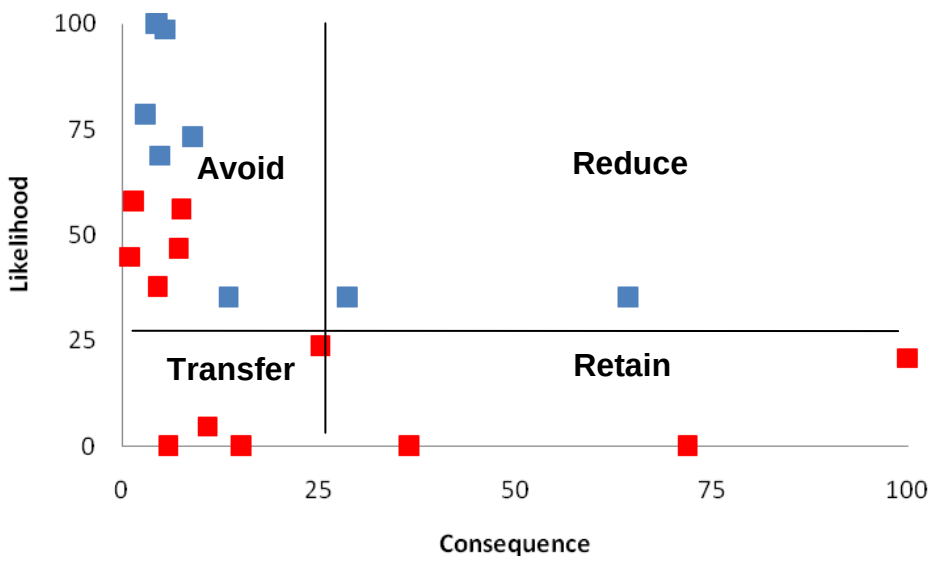

Figure 4.10: CAMPO Risk Management Recommendations

For CAMPO, it is recommended to *avoid* projects that are not likely to appear among the most congested links in the TDM and also have low expected VHT savings. Those links with a high probability of appearing among the most congested links, but with low expected VHT savings, are recommended to be *transferred* or only pursued with excess funding, because they may not be as critical. Projects with potentially high VHT savings that are also more likely to appear are recommended to be *retained*. Projects not likely to appear in the TDM results, but likely to have high VHT savings, are recommended to have their risks *reduced*. Once reduced, through data collection or expert opinion, the points can be redistributed to the other quadrants in the 'risk matrix'.

Ideally, running the TDM using the current data would result in all critical projects falling in the *retain* quadrant of a risk matrix, due to a high 'consequence' and low 'likelihood. This means that the projects with the highest return (or VHT savings) are the most likely to appear in the output of the TDM. Basically, the calculation of 'likelihood' is used as a screening process, while the calculation of 'consequence' is used for the decision-making process.

In tabular form, Table 4.4 further shows the recommended versus the likely CAMPO strategies, where the 'likely strategy' is to accept the single solution of the TDM, with current data, and retain the projects on the most congested links.

Table 4.4: Recommended versus Likely CAMPO Strategies

Link ID	Street Name	Recommended Strategy	Likely CAMPO Strategy
705492	Maple St.	Reduce	--
705467	Wildwood Pl.	Retain	Retain
705552	Tipton Ln.	Reduce	--
665015	Washington St.	Avoid	--
672163	Washington St.	Retain	Retain
672206	Washington St.	Avoid	--
705386	Maple St.	Avoid	--
665528	Lafayette Ave.	Avoid	--
665434	1st St.	Avoid	--
665831	Indianapolis Rd.	Avoid	Retain
10873321	Indianapolis Rd.	Avoid	Retain
665537	S.R. 46	Avoid	--
665547	S.R. 46	Avoid	--
665041	Franklin St.	Avoid	Retain
10873325	Indianapolis Rd.	Avoid	Retain
10871342	S.R. 46	Transfer	Retain
10873328	Indianapolis Rd.	Avoid	Retain
664902	S.R. 46	Retain	Retain
10871959	Lincoln Park Dr.	Retain	Retain
10873324	Lowell Rd.	Transfer	Retain
10873327	Lowell Rd.	Transfer	Retain

The recommended and 'likely' strategies differ significantly. Several projects retained under the 'likely strategy' are recommended to be *avoided* or *transferred*. By avoiding risk, planners can reserve more resources for projects that are considered more likely to have higher VHT savings. By transferring risk, planners can risk another organization's money or use non-essential funds to pursue a project that may not yield the highest VHT savings.

4.6. Risk Monitoring

A continuous risk management framework involving with risk monitoring is recommended for planning organizations. Planning organizations can assign responsibility to staff to follow up on the effectiveness of the recommended risk

management strategies. With more information obtained for links under the *reduce* strategy, the 'likelihood' or 'consequence' will likely change. For instance, with more retail employment data, the sampling error can be reduced, leading to more consistent model outputs, which affects the 'likelihood' of each link appearing in the top 10 v/c links. With a lower 'likelihood', the project may be recommended to be *retained*.

To keep track of the risks, a risk log can be used to track the 'likelihood', 'consequence', and risk management strategy pursued (Verzuh, 2005). Communication of the risk log is critical to make all stakeholders aware of the risk.

4.7. Conclusion

Considering the number of uncertainties associated with use of the traditional travel demand model, it seems unreasonable to make policy decisions based on a single model output. Instead, risk analysis can avoid the inefficient use of resources. By borrowing risk frameworks commonly used in private industry, the number of links considered for programming can be expanded to ensure that the most optimal projects are undertaken. Once the most efficient projects have been programmed, planners can then focus on 'project risk' management.

CHAPTER 5. SUMMARY AND FUTURE RESEARCH

5.1. Overall Conclusions

The travel demand model (TDM) is an imperfect tool used to predict transportation patterns and guide construction project programming decisions. To ensure that the most resource-efficient projects are planned, model inputs and parameters must be current and representative of the planning area.

Considering the dynamic nature of travel demand, inputs should be updated continuously to reflect the most current planning area characteristics (Chapter 1). This can be done through cost-free data sources, such as the American Community Survey (ACS), or through cost-associated data sources, such as the DemographicsNow pay-site.

One approach to obtaining accurate inputs and travel predictions is to apply synthetic cross-classification (Chapter 2). This approach to trip generation can obviate the need for household surveys, by applying a doubly-constrained mathematical model to ACS data and relatively well-known ‘marginals’ or household totals by *household size* and *vehicles available*.

TDM inputs and parameters are uncertain. By analyzing such inputs and parameters stochastically, planners can focus data collection and make better-informed model selections (Chapter 3).

To manage the uncertainty of TDM inputs and parameters, risk management frameworks from the private sector can be applied (Chapter 4). Such an approach can expand the number of capacity-building projects being considered when programming, as well as help decide which projects are the most likely to reduce congestion and yield the highest link and network vehicle-hours-traveled (VHT) savings.

5.2. Future Research

Future research will focus upon:

- How to verify employment values reported in dense urban areas.
- How the synthetic cross-classification approach compares to cross-classification tables developed using significant household travel surveys.
- How to develop distributions for travel demand model parameters.

The travel demand model yields uncertain results, but by quantifying the uncertainties in the model inputs and parameters, planners can help ensure that the most critical projects are being scheduled for construction.

LIST OF REFERENCES

LIST OF REFERENCES

2002 Economic Census (August 28, 2008). U.S. Census Bureau. Retrieved September 24, 2008 from HTTP://WWW.CENSUS.GOV/ECON/CENSUS02.

Abraham, John K. *CE 7630 Course Lecture: Trip Generation*. Wayne State University. Retrieved September 24, 2008 from HTTP://WWW.ENG.WAYNE.EDU/LEGACY/FORMS/4/POWERPOINT%20_TRIP_GENERATION.PDF.

ACS: Public Use Microdata Sample Main Page (September 22, 2008). U.S. Census Bureau: American Community Survey Office. Retrieved September 24, 2008 from HTTP://WWW.CENSUS.GOV/ACS/WWW/PRODUCTS/PUMS/INDEX.HTM.

Alexander, Corinne and Maria I. Marwill (April 2006). *The Risk Matrix: Illustrating the Importance of Risk Management Strategies*. Purdue University - Department of Agricultural Economics. Journal of Extension, Volume 44 Number 2, Article Number 2TOT1. Retrieved October 29, 2008 from HTTP://WWW.JOE.ORG/JOE/2006APRIL /TT1.SHTML.

American Community Survey (September 15, 2008). U.S. Census Bureau: American Community Survey Office. Retrieved September 24, 2008 from HTTP://WWW. CENSUS.GOV/ACS/WWW/.

American FactFinder. U.S. Census Bureau. Retrieved November 10, 2008 from FACTFINDER.CENSUS.GOV/.

Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc. (February 1997). *Model Validation and Reasonableness Checking Manual*. Prepared for the Travel Model Improvement Program of the Federal Highway Administration. Retrieved May 2, 2008 from HTTP://TMIP.FHWA.DOT.GOV/ CLEARINGHOUSE/DOCS/MVRCM/ MVRCM.PDF.

Basic Information | Risk Assessment Portal | US EPA (Last updated March 27, 2008). U.S. Environmental Protection Agency. Retrieved October 29, 2008 from HTTP:// WWW.EPA.GOV/RISK/BASICINFORMATION.HTM #ARISK.

Bernardin, Lochmuller, and Associates, Inc and Transportation Support Group, Inc.. *Indiana Reference Modeling System*. Indiana Department of Transportation, Transportation Planning Division, December 1996.

Bernardin, Lochmuller, and Associates, Inc and Transportation Support Group, Inc.. *Lexington Area TransCAD Model Upgrade*. Lexington-Fayette County Government, Division of Planning, July 2004.

Blair, Brian (October 3, 2008). *Columbus Regional Hospital Reopening October 27, in Better Shape Than Ever* . The Republic. Obtained through the Indiana Economic Digest. Retrieved November 8, 2008 from HTTP://WWW.INDIANA ECONOMICDIGEST.COM/MAIN.ASP?SECTIONID=31&SUBSECTIONID=70 &ARTICLEID=43954.

Blair, Brian (September 3, 2008). *Historic Columbus Inn Closed*. The Republic. Obtained through the Indiana Economic Digest. Retrieved November 8, 2008 from HTTP://WWW.INDIANAECONOMICDIGEST.NET/MAIN.ASP?SECTION ID=31&SUBSECTIONID=62&ARTICLEID=42975.

Blodgett, John and Ryan Burson (July 2004). *Q & As About the ACS*. Adapted from a May 2004 U.S. Census Bureau presentation, developed by Cynthia Tauber and Dave Hubble, for a Missouri Census Data Center Meeting. Retrieved December 3, 2008 from MCDC2.MISSOURI.EDU/PRESENTATIONS/MCDCCONFERENCE_JULY04/ACS0704_PERTAEUBER.PPT.

Caliper Corporation. *Utown Training Model*. TransCAD 4.7 Software, Fall 2005.

Cefic – The European Chemical Industry Council (August 2003). *Risk and Hazard: How They Differ*. Retrieved October 29, 2008 from HTTP://WWW.DEHP-FACTS.COM/UPLOAD/DOCUMENTS/WEBPAGE/DOCUMENT52.PDF.

CenStats Databases (March 14, 2008). U.S. Census Bureau. Retrieved September 24, 2008 from HTTP://CENSTATS.CENSUS.GOV/INDEX.SHTML.

Chiu, Melissa. *FHWA Data Workshop*. As presented September 17, 2008 at the 11th National Tools of the Trade Conference organized by TRB Committee ADA 30.

Code of Federal Regulations (April 1, 2008). *Title23-Highways Chapter 1-Federal Highway Administration Department of Transportation Part 450-Planning Assistance and Standards*. Section 322, pgs. 114-117. Retrieved October 2, 2008 from HTTP://WWW.ACCESS.GPO.GOV/NARA/CFR/WAISIDX_08/23CFR450_08.HTML.

Damnjanovic, Ivan and Zhanmin Zhang. *Risk-based Model for Valuation of Performance-specified Pavement Maintenance Contracts*. Journal of Construction Engineering and Management v. 134 no. 7 (July 2008) p.492-500.

de Dios Ortúzar, Juan and Luis H. Willumsen. *Modelling Transport*. John Wiley and Sons, 2001, pgs. 140-143. Retrieved October 9, 2008 from HTTP://BOOKS. GOOGLE.COM/.

DemographicsNow (2008). SRC, LLC. Retrieved September 24, 2008 from HTTP:// WWW.DEMOGRAPHICSNOW.COM/.

Dey, Soumya S. and Jon D. Fricker. *Bayesian Updating of Trip Generation Data: Combining National Trip Generation Rates with Local Data*. Transportation, Volume 21, December 1994, p. 393-403.

Dixon, Michael P (January 12, 2008). *CE 573 Transportation Planning Course Lecture: Cross-Classification Based Trip Generation*. University of Idaho. Retrieved September 24, 2008 from HTTP://WWW.WEBS1.UIDAHO. EDU/CE573/.

Dobson, Ricardo and William E. McGarvey. *An Empirical Comparison of Disaggregate Category and Regression Trip Generation Analysis Techniques*. Charles River Associates Inc. and University of Southern California. Transportation, Volume 6, Number 3. Springer Netherlands, September 1977.

Fehr and Peers (2005). *Las Vegas Travel Demand Model Guidelines for Estimation, Calibration, and Validation*. Regional Transportation Commission of Southern Nevada.

Feng, Li-Yang and Lisa C. Gion (June 1995). *Traffic Assignment: BPR to HCM (Part I) - Denver's Application*. Denver Regional Council of Governments and the Federal Highway Administration. Presented at the 1995 TRB Planning Applications Conference, Seattle, Washington.

Flyvbjerg, Bent. *From Noble Prize to Project Management: Getting Risks* Right. Project Management Journal v. 37 no. 3 (August 2006) p. 5-15.

Fratar, Thomas J.. *Vehicular Trip Distribution by Successive Approximations*. Traffic Quarterly 8:1, January 1954, p. 53-65.

Fricker, Jon D. (September 19, 2007). *CE 566 Transportation Planning Course Lecture*. Purdue University.

Geographic Research, Inc. – SimplyMap (2007). Geographic Research, Inc. Retrieved December 22, 2008 from HTTP://WWW.GEOGRAPHICRESEARCH.COM/SIMPLYMAP.HTML.

Goodman, Alvin S. and Makarand Hastak. *Infrastructure Planning Handbook*. ASCE Press, McGraw-Hill, New York, 2006.

Grecco, W.L., Spencer, J., Wegmann, F.J., and Chatterjee, A.. *Transportation Planning for Small Urban Areas*. National Cooperative Highway Research Program Report 167, Transportation Research Board, 1976.

Hu, Patricia S., Timothy Reuscher, Richard L. Schmoyer, and Shih-Miao Chin (May 2, 2007). *Transferring 2001 National Household Travel Survey*. Prepared for the Federal Highway Administration. Retrieved May 15, 2008 from HTTP://NHTS-GIS.ORNL.GOV/TRANSFERABILITY/TRANSFERABILITYREPORT.PDF.

Hutchinson, B.G.. *Principles of Urban Transport Systems Planning*. McGraw-Hill, 1954, Section 4.1.

Ismart, Dane (December 1990). *Calibration and Adjustment of System Planning Models*. Federal Highway Administration. Retrieved May 21, 2008 from HTTP:// NTL.BTS.GOV/ DOCS/377CAS.HTML.

IN.Gov: Home (2008). Retrieved October 2, 2008 from HTTP://WWW.IN.GOV/.

Indiana Foreclosures – Foreclosures in Indiana (2008). ForeclosureFreeSearch, Inc.. Retrieved October 2, 2008 from HTTP://WWW.FORECLOSURE.COM/STATE/IN.HTML.

Indiana Workforce Development – Hoosiers by the Numbers Home. Indiana Department of Workforce Development. Retrieved September 24, 2008 from HTTP://HOOSIERDATA.IN. GOV/INDEX.ASP.

ITE Technical Council Committee 6Y-53 chaired by Mann, William W. *Travel Demand Forecasting Processes Used by Ten Large Metropolitan Planning Organizations*. An Informational Report of the Institute of Transportation Engineers, February 1994.

Kalafatas, Georgios. *Network Design for Evacuation Planning*. Thesis (MSCE) – Purdue University, CE, December 2005.

Kanafani, A.. *Transportation Demand Analysis*. McGraw-Hill Companies, Inc, 1983.

Kirkman, T.W. (1996). *Statistics to Us: Kolmogorov-Smirnov Test*. College of Saint Benedict and Saint John's University. Retrieved October 7, 2008 from HTTP:// WWW.PHYSICS.CSBSJU. EDU/STATS/KS-TEST.HTML.

Kyd, Charles W. (July 2006). *An Introduction to Excel's Normal Distribution Functions*. ExcelUser, Inc.. Retrieved October 3, 2008 from HTTP://WWW.EXCELUSER.COM/ EXPLORE/STATSNORMAL.HTM.

Ladwig, Boris (October 22, 2008). *Bartholomew County Adds, State Loses Jobs in September 2008*. The Republic. Obtained from the Indiana Economic Digest. Retrieved November 8, 2008 from HTTP://WWW.INDIANAECONOMICDIGEST.NET/MAIN.ASP?SECTIONID=31&SUBSECTIONID=156&ARTICLEID=43883.

Lima and Associates, Inc. (January 2006). *Lincoln MPO Travel Demand Model*. Prepared for the Lincoln Metropolitan Planning Organization. Retrieved September 22, 2008 from HTTP://WWW.LINCOLN.NE.GOV/CITY/PLAN/MPO/TRAVEL/.

Lysy, Daniel N.. *TAC Policy on use of RSTP and CMAQ Contingency Funds*. MPO Technical Advisory Committee (TAC) Resolution January 20, 2005 TAC Meeting. Retrieved November 5, 2008 from HTTP://WWW.RICHMONDREGIONAL.ORG/PUBLICATIONS/REPORTS%20AND%20DOCUMENTS/TAC_%20POLICY_ON_PROGRAMMING_CMAQ_RSTP_CONTINGENCY_FUNDS.PDF.

Mahmassani, Hani S. and Kumares C. Sinha. *Baysian Updating of Trip Generation Parameters*. Massachusetts Institute of Technology and Purdue University. Transportation Engineering Journal, Vol. 107, No. 5, September/October 1981, pp. 581-589.

Mailing List – Email, Business, Sales Leads, Consumer – InfoUSA (2008). infoUSA.com, Inc. Retrieved September 24, 2008 from HTTP://LIST.INFOUSA.COM/.

Marshment, Richard S. *Urban Travel Characteristics Database*. Prepared by the ITE Technical Council Committee 6Y-46, April 1995.

Martin, William A. and Nancy A. McGuckin. *National Cooperative Highway Research Program Report 365*. Prepared for the Transportation Research Board, National Research Council. National Academy Press: Washington, D.C. 1998.

Melissa Data | Data Quality Tools | Mailing Software | Lists } NCOA (2008). Melissa DATA Corp. Retrieved September 24, 2008 from HTTP://WWW.MELISSADATA. COM/.

MID-BEDS District Council (September 2006). *Risk Management Strategy (September 2006 – March 2009)*. Appendix A. Retrieved October 29, 2008 from HTTP://WWW.MIDBEDS.GOV.UK/IMAGES/RISK%20MANAGEMENT%20STRATEGY%20_APPENDIX%20A__TCM5-5832.PDF.

Moulton, Sam. *When Disaster Strikes*. Chicago (1975) v. 56 no. 3 (Martch 2007) p. 49.

NHTS Home. National Household Travel Survey and Center for Transportation Analysis, Oak Ridge Laboratory (ORNL) under funding from the Federal Highway administration. Retrieved September 24, 2008 from HTTP://NHTS. ORNL.GOV/.

OGC – Risk Management Strategy (Last Updated October 20, 2008). Office of Government Commerce, United Kingdom. Crown Copyright. Retrieved October 29, 2008 from HTTP://WWW.OGC.GOV.UK/DELIVERY_LIFECYCLE_RISK_MANAGEMENT_STRATEGY_.ASP.

Oregon Department of Transportation Planning Analysis Unit (August 4, 2004). *Grant Pass Model Development*. Provided for research purposes by Michael H. Gillett.

Oregon Department of Transportation Planning Analysis Unit and MW Consulting (May 2008). *Corvallis Area Metropolitan Planning Organization: 2000 Travel Demand Model Development Documentation*. Provided for research purposes by Michael H. Gillett.

Palisade (November 1998). *@Risk: Simulation Add-In For Microsoft Excel*. Windows Version Release 4.0. Newfield, NY: Palisade Corp.

Population Estimates – U.S. Census Bureau (August 21, 2008). U.S. Census Bureau. Retrieved September 24, 2008 from HTTP://WWW.CENSUS.GOV/ POPEST/ ESTIMATES.PHP.

Proussaloglou, K.E. Popuri, Y.D., Aunet, B.A., and Cipra, D.M.. *Enhancing State and MPO Transportation Planning Using National Household Travel Survey Add-On Data: The Wisconsin Experience*. Transportation Research Board, October 12, 2004.

Purvis, Charles L. (December 1994). *San Francisco Bay Area 1990 Regional Travel Characteristics*. Metropolitan Transportation Commission - Oakland, CA. Retrieved September 22, 2008 from HTTP://NTL.BTS.GOV/DOCS/SF. HTML.

Regional Transportation Commission of Southern Nevada (July 2006). *Regional Transportation Plan FY 2006-2030*. Retrieved September 22, 2008 from HTTP://WWW.RTCSOUTHERNNEVADA.COM/MPO/DOCUMENTS/ RTP0630.CFM.

REMI Home. Regional Economic Models, Inc.. Retrieved October 2, 2008 from *HTTP://WWW. REMI.COM/.*

Renne, John. *Evacuation and Equity*. Planning v. 72 no. 5 (May 2006) p. 44-46.

Risk: Definition and Additional Resources from BNET (2008). CBS Interactive Inc. Retrieved October 29, 2008 from HTTP://DICTIONARY.BNET.COM/ DEFINITION/ RISK.HTML.

Risk Management (2008). Carnegie Mellon University – Software Engineering Institute. Retrieved October 29, 2008 from HTTP://WWW.SEI.CMU.EDU/ RISK/.

Sanderson, Peter. *A Closer Look: Minnesota Department of Transportation*. PM Network v. 22 no. 5 (May 2008) p. 40-45.

Showalter, Brenda (August 28, 2008). *Flood, Economy Force Columbus, Edinburg Furniture Stores to Shut Doors*. The Republic. Obtained through the Indiana Economic Digest. Retrieved November 8, 2008 from HTTP://WWW.INDIANAECONOMICDIGEST.NET/MAIN.ASP?SECTIONID=31&SUBSECTIONID=62&ARTICLEID=42890.

Singh, Rupinder (June 1995). *Beyond the BPR Curve: Updating Speed-Flow and Speed-Capacity Relationships in Traffic Assignment*. Metropolitan Transportation Commission - Oakland, CA. Presented at the 1995 TRB Planning Applications Conference, Seattle, Washington.

Sinha, Kumares C. and Samuel Labi. *Transportation Decision Making: Principles of Project Evaluation and Programming*. John Wiley and Sons, Inc., Hoboken, New Jersey: 2007.

STATS Indiana. Indiana Department of Workforce Development, Lilly Endowment, Inc., and the Indiana Business Research Center at Indiana University's Kelley School of Business. Retrieved September 24, 2008 from HTTP://WWW.STATS.INDIANA. EDU/.

Tatom, John (August 28, 2007). *Why is the Foreclosure Rate so High in Indiana?*. Networks Financial Institute – Indiana State University. Retrieved November 5, 2008 from HTTP://MPRA.UB.UNI-MUENCHEN.DE/4674/1/ MPRA_PAPER_4674. PDF.

The Republic (September 6, 2008). *Columbus Engineering Company Expands to 2nd Local Location*. Obtained through the Indiana Economic Digest. Retrieved November 8, 2008 from HTTP://WWW.INDIANAECONOMIC DIGEST.NET/MAIN.ASP?SECTIONID=31&SUBSECTIONID=70&ARTICLEI D=43043.

The University: Risk Management Strategy – Durham University (Last Modified March 12, 2008). Durham University. Retrieved October 29, 2008 from HTTP://WWW.DUR.AC.UK/ABOUT/GOVERNANCE/OPERATIONAL/ RISK_MANAGEMENT/.

Transportation Research Board (2007). *Special Report 288: Metropolitan Travel Forecasting – Current Practice and Future Direction*. Committee for Determination of the State of the Practice in Metropolitan Area Travel Forecasting, Washington D.C.

United States Census Bureau: American Community Survey (2006). *Accuracy of the Data (2006)*. Retrieved August 1, 2008 from HTTP://WWW.CENSUS. GOV/ACS/WWW/DOWNLOADS/ACS/ ACCURACY2006.PDF.

U.S. Census Bureau – Longitudinal Employer-Household Dynamics (September 24, 2008). U.S. Census Bureau: Center for Economic Studies. Retrieved September 24, 2008 from HTTP:// LEHD.DID.CENSUS.GOV/LED/ INDEX.HTML.

Verzuh, Eric. *The Fast Forward MBA in Project Management, 2nd Edition.* John Wiley and Sons, Inc., Hoboken, New Jersey: 2005.

Weisstein, Eric W.. *Box-Muller Transformation*. From MathWorld – A Wolfram Web Resource. Retrieved November 2, 2008 from HTTP://MATHWORLD. WOLFRAM.COM/BOX-MULLERTRANSFORM ATION.HTML.

Wolshon, Brian. *Planning for the Evacuation of New Orleans*. ITE Journal v. 72 no. 2 (February 2002) p. 44-49.

Zhao, Yong and Kara M. Kockelman (August 2001). *The Propagation of Uncertainty Through Travel Demand Models: An Exploratory Analysis*. Prepared by The University of Texas at Austin for the Annals of Regional Science. Retrieved August 18, 2008 from HTTP://WWW. CE.UTEXAS.EDU/ PROF/KOCKELMAN/PUBLIC_ HTML/ ARS01ERRORPROPAGATION. PDF.

APPENDICES

Appendix A: Travel Demand Model Parameter Databases

Trip Generation

Trip Rates by Vehicles Available and Household Size

References: [1] Purvis, 1994, [2] Marshment, 1995, [3] Martin and McGuckin,1998, and [4] Hu et al., 2007

Regression Equations by HBW, HBO, NHB, and HB-Shop trip purposes and Total trips generated

References: [1] Grecco et al., 1976, [2] Caliper Corporation, 2005, [3] ITE, 1994, [4] Martin and McGuckin, 1998, [5] ODOT, 2004, [6] ODOT, 2008, [7] Kanafani, 1983, [8] Proussaloglou, 2004, [9] B.L.A., inc., 2004, and [10] B.L.A., Inc., 1996

Trip Distribution

Friction Factor (Tanner) Functions for HBW, HBO, NHB, and HB-Shop trip purposes

References: [1] Martin and McGuckin, 1998, [2] Lima and Associates, Inc., 2006, [3] Ismart, 1990, [4] RTC of Southern Nevada, 2006, and [5] Zhao et al., 2001

Synthetic Friction Factors

References: [1] Martin and McGuckin, 1998 and [2] B.L.A., Inc., 2004

Trip Assignment

Link Performance Functions for Single Purpose, Freeway, Arterial, Multilane, Signalized and Un-signalized Links

References: [1] Singh, 1995, [2] Martin and McGuckin, 1998, [3] Feng and Gion, 1995, and [4] B.L.A., Inc., 2004

Trip Rates

Reference	Household Size	1				2				3				4+			
	Vehicle Availability	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+
[1]	*San Francisco*	2.72	3.91	3.81	4.17	4.05	5.79	6.71	7.08	5.78	7.30	8.37	9.64	8.26	10.17	12.56	14.47
[2]	*Albany*	2.14	3.76	--	--	3.78	6.07	7.22	5.97	--	9.28	9.92	11.00	--	13.44	14.79	14.87
[2]	*Houston - Galveston*	0.30	3.80	4.36	5.83	0.67	4.49	6.29	6.95	--	5.05	8.18	10.47	--	5.34	9.14	11.18
[3]	*Nashua*	1.00	5.04	4.00	--	--	7.58	7.26	7.60	2.67	7.18	10.02	10.28	--	10.20	12.27	11.81
[3]	*Charlotte*	1.55	4.36	4.43	3.60	2.29	7.40	7.64	9.43	5.43	8.29	10.45	11.15	7.75	12.82	16.26	17.62
[3]	*Reno*	1.79	4.42	4.97	3.89	2.75	6.30	7.22	8.21	3.67	8.29	8.38	11.01	--	10.62	14.32	15.26
[3]	*Austin*	2.12	4.48	5.33	4.00	2.80	5.13	7.18	7.96	7.25	6.73	9.19	10.09	6.40	10.58	11.42	13.91
[3]	*Vancouver*	3.29	3.20	3.52	4.23	2.00	3.50	4.72	4.87	5.00	5.40	5.82	7.40	14.00	7.19	7.70	10.16
[3]	*Phoenix*	1.56	4.09	5.22	4.58	3.03	6.54	7.37	7.39	5.75	7.21	10.21	10.73	4.79	10.38	13.26	15.33
[3]	*Pittsburgh*	3.32	5.27	5.00	--	4.43	8.86	7.66	6.75	6.40	10.10	10.88	11.43	7.40	12.86	14.75	16.25
[3]	*Seattle*	4.08	4.73	4.34	4.88	4.37	4.50	4.49	4.77	5.60	5.27	4.76	4.81	3.86	4.94	4.75	4.93
[3]	*St. Louis*	1.70	4.10	4.16	6.20	5.08	6.48	7.60	8.02	6.09	9.46	9.52	12.05	12.86	10.40	14.98	16.63
[3]	*San Diego*	3.27	5.91	5.79	8.38	7.35	10.24	10.88	12.31	11.40	14.23	15.80	16.46	12.25	20.42	24.41	25.09
[3]	*Atlanta*	0.78	3.92	3.62	4.72	2.47	6.44	7.10	8.61	5.62	10.33	10.08	10.74	10.78	14.28	15.76	16.28
[3]	*New Jersey*	1.38	3.57	3.84	2.00	3.33	5.68	6.62	8.29	4.50	6.00	7.76	10.02	4.56	7.20	11.33	14.16
[3]	*NPTS*	1.47	3.16	3.51	3.82	3.29	4.63	5.53	6.00	4.82	6.66	7.53	8.20	6.85	9.10	11.07	11.37
[4]	*NHTS - CAMPO*	3.85	4.33	5.36	6.48	7.49	7.79	8.10	8.30	11.38	11.48	11.67	11.95	16.24	16.37	16.66	16.88

Home-based Work (HBW) Regression – Productions

Reference		Production Equations
[1]	1	Y = -2.90486 + 0.75377(*# of Vehicles*)
[1]	2	Y = 0.42(*Population*)
[1], [2]	3	Y = 8 + 1.2(*# of Vehicles*)
[1]	4	Y = 0.4(*Population*)
[3]	5	Y = 0.813(*Total Employment*)
[3]	6	Y = 0.831(*Total Employment*)
[3]	7	Y = 1.53(*Total Employment*)
[3]	8	Y = 1.574(*Total Employment*)
[3]	9	Y = 1.537(*Total Employment*)
[3]	10	Y = 1.7(*Total Employment*)
[3]	11	Y = 1.29(*Total Employment*)
[3]	12	Y = 17.04 + 1.14(*Labor Force*)
[1]	13	Y = 1.12(*Labor Force*)
[3]	14	Y = 0.6806(*# of Vehicles*) + 0.574(*# of Households*) + 0.18(*Group Quarters Population w/in 10 miles*) + 0.72(*Remaining Group Quarters Population*)
[4]	15	Y = 0.74(*# of 1-Person Households*) + 1.672(*# of 2-Person Households*) + 2.014(*# of 3-Person Households*) + 2.584(*# of 4-Person Households*) + 2.822(*# of 5+-Person Households*)
[5]	16	Y = 0.049(*0-Worker Households*) + 1.384(*1-Worker Households*) + 2.649(*2-Worker Households*) + 4.7(*3+-Worker Households*)
[6]	17	Y = 1.383(*1-Worker Households*) + 2.391(*2-Worker Households*) + 3.887(*3+-Worker Households*)

Home-based Work (HBW) Regression – Attractions

Reference		Attraction Equations
[1]	1	Y = 34.51357 + 1.47557(*Retail Employment*) + 0.4897(*Non-Retail Employment*)
[1]	2	Y = 1.268(*Total Employment*)
[1], [2]	3	Y = 54 + 0.9(*Total Employment*)
[4]	4	Y = 1.45(*Total Employment*)
[3]	5	Y = 1.6813(*Total Employment*)
[3]	6	Y = 1.7(*Total Employment*)
[3]	7	Y = 6.08 + 1.83(*Total Employment*)
[3]	8	Y = 1.48(*Total Employment*)
[3]	9	Y = 1.33(*Total Employment*)
[3]	10	Y = 1.22(*Total Employment*)
[7]	11	Y = -193 + 1.54(*Total Employment*)
[3]	12	Y = 1.478(*Total Employment*)
[3]	13	Y = 1.477(*Total Employment*)
[3]	14	Y = 1.447(*Total Employment*)
[3]	15	Y = 1.411(*Total Employment*)
[3]	16	Y = 1.467(*Total Employment*)
[8]	17	Y = 1.397(*Total Employment*)
[8]	18	Y = 1.244(*Total Employment*)
[8]	19	Y = 1.145(*Total Employment*)
[8]	20	Y = 1.171(*Total Employment*)
[8]	21	Y = 1.505(*Total Employment*)
[8]	22	Y = 1.177(*Total Employment*)
[8]	23	Y = 1.206(*Total Employment*)
[9]	24	Y = (*Total Employment*)
[1]	25	Y = 0.40(*Total Population* * (*Zonal Employment* / *Total Employment*))
[1]	26	Y = 32.94 + 0.87(*Total Employment*) + 0.0008(*Manufacturing Floor Area*)
[1]	27	Y = 1.04(*Total Employment*) - 0.20(*Manufacturing Floor Area*) + 0.40(*Public and Quasi-Public Floor Area*)

Home-based Other (HBO) Regression – Productions

References		Production Equations
[1]	1	Y = -23.82007 + 3.74414(# *of Vehicles*)
[1]	2	Y = 1.02(*Population*)
[1], [2]	3	Y = 18 + 2.7(# *of Vehicles*)
[1], [2]	4	Y = 0.96(*Population*)
[3]	5	Y = 0.54(# *of 0-Vehicle Households*) + 2.36(# *of 1-Vehicle Households*) + 3.74(# *of 2-Vehicle Households*) + 4.24(# *of 3+-Vehicle Households*)
[3]	6	Y = 0.57(# *of 0-Vehicle Households*) + 2.49(# *of 1-Vehicle Households*) + 3.94(# *of 2-Vehicle Households*) + 4.46(# *of 3+-Vehicle Households*)
[3]	7	Y = 0.62(# *of 0-Vehicle Households*) + 2.70(# *of 1-Vehicle Households*) + 4.25(# *of 2-Vehicle Households*) + 4.87(# *of 3+-Vehicle Households*)
[3]	8	Y = 0.98(# *of 0-Vehicle Households*) + 4.25(# *of 1-Vehicle Households*) + 6.64(# *of 2-Vehicle Households*) + 7.73(# *of 3+-Vehicle Households*)
[3]	9	Y = 1.14(# *of 0-Vehicle Households*) + 4.98(# *of 1-Vehicle Households*) + 7.78(# *of 2-Vehicle Households*) + 9.02(# *of 3+-Vehicle Households*)
[3]	10	Y = 0.77(# *of 1-Vehicle Households*) + 3.11(# *of 2-Vehicle Households*) + 3.88(# *of 3+-Vehicle Households*)
[3]	11	Y = 1.45(# *of 1-Vehicle Households*) + 2.81(# *of 2-Vehicle Households*) + 3.52(# *of 3+-Vehicle Households*)
[3]	12	Y = 1.80(# *of 1-Vehicle Households*) + 3.23(# *of 2-Vehicle Households*) + 4.7(# *of 3+-Vehicle Households*)
[4]	13	Y = 1.998(# *of 1-Person Households*) + 4.104(# *of 2-Person Households*) + 5.936(# *of 3-Person Households*) + 7.888(# *of 4-Person Households*) + 10.292(# *of 5+-Person Households*)

Home-based Other (HBO) Regression – Attractions

Reference		Attraction Equations
[2]	1	Y = 206 + 0.95(*# of Households*) + 0.6(*Total Employment*)
[3]	2	Y = 10(*Retail Employment*) + 0.5(*Non-Retail Employment*) + 1(*# of Households*)
[7]	3	Y = 317 + 0.26(*Population*)
[3]	4	Y = 0.608(*Total Employment*) + 0.37(*Population*)
[1]	5	Y = 88.7802 + 1.0354(*# of Vehicles*) + 11.4978(*Retail Employment*) + 0.3718(*Non-Retail Employment*) + 0.3924(*School Enrollment*)
[1]	6	Y = 0.4(*Population*) + 2.22(*Non-Industrial Employment*)
[1]	7	Y = 206 + 0.6(*Total Employment*) + 0.3(*School Enrollment*) + 0.8(*# of Households*)
[4]	8	Y = 2(*Retail Employment*) + 0.9(*# of Households*) + 1.7(*Service Employment*) + 0.5(*Basic Employment*)
[4]	9	Y = 9(*Retail Employment*) + 0.9(*# of Households*) + 1.7(*Service Employment*) + 0.5(*Basic Employment*)
[3]	10	Y = 0.3105(*Population*) + 9.0274(*Retail Employment*) + 0.3561(*Non-Retail Employment*) - 1.1917(*Basic Employment*)
[3]	11	Y = 195.14 + 0.38(*Population*) + 2.21(*Service Employment*)
[3]	12	Y = 0.88(*Retail Employment*) + 0.22(*# of Households*) + 0.09(*Office Employment*) + 0.44(*Other Non-Industrial Employment*)
[3]	13	Y = 1.12(*Retail Employment*) + 0.27(*# of Households*) + 0.17(*Office Employment*) + 0.7(*Other Non-Industrial Employment*)
[3]	14	Y = 1.5(*Retail Employment*) + 0.28(*# of Households*) + 0.11(*Office Employment*) + 1.17(*Other Non-Industrial Employment*)
[3]	15	Y = 1.57(*Retail Employment*) + 0.3(*# of Households*) + 0.27(*Office Employment*) + 1.52(*Other Non-Industrial Employment*)
[1]	16	Y = 0.96(*Total Population *(Non-Industrial Zonal Employment / Non-Industrial Total Employment)*))
[3]	17	Y = 2.04(*Retail Employment*) + 0.61(*Non-Retail Employment*) + 0.61(*# of Households*) + 0.2(*Basic Employment*)
[3]	18	Y = 2.55(*Retail Employment*) + 0.81(*Non-Retail Employment*) + 0.71(*# of Households*) + 0.26(*Basic Employment*)
[3]	19	Y = 3.05(*Retail Employment*) + 1.02(*Non-Retail Employment*) + 0.81(*# of Households*) + 0.36(*Basic Employment*)
[3]	20	Y = 7.38(*Retail Employment*) + 3.46(*Non-Retail Employment*) + 1.43(*# of Households*) + 0.71(*Basic Employment*)
[3]	21	Y = 11.71(*Retail Employment*) + 3.46(*Non-Retail Employment*) + 1.43(*# of Households*) + 0.22(*Basic Employment*)
[3]	22	Y = 13.74(*Retail Employment*) + 4.58(*Non-Retail Employment*) + 1.53(*# of Households*) + 0.22(*Basic Employment*)
[8]	23	Y = 0.664(*Retail Employment*) + 1.023(*# of Households*) + 1.268(*Service Employment*)
[8]	24	Y = 0.664(*Retail Employment*) + 1.069(*# of Households*) + 0.594(*Service Employment*)

Home-based Other (HBO) Regression – Attractions Continued

References		Attraction Equations
[8]	25	Y = 0.664(*Retail Employment*) + 1.194(*# of Households*) + 0.393(*Service Employment*)
[8]	26	Y = 0.664(Retail Employment) + 1.1(# of Households) + 0.87(Service Employment)
[8]	27	Y = 0.664(*Retail Employment*) + 1.361(*# of Households*) + 1.189(*Service Employment*)
[8]	28	Y = 0.771(*Retail Employment*) + 1.137(*# of Households*) + 0.516(*Service Employment*)
[8]	29	Y = 0.771(*Retail Employment*) + 1.361(*# of Households*) + 1.152(*Service Employment*)
[8]	30	Y = 1.11(*Retail Employment*) + 1.138(*# of Households*) + 0.508(*Service Employment*)
[9]	31	Y = 0.1168(*Population*) + 0.8257(*Retail Employment*) + 0.0408(*Commercial and Government Employment*) + 0.0137(*Other Employment*)

Non-home-based (NHB) Regression – Productions

References		Production Equations
[2]	1	Y = 67 + 0.15(*# of HHs*) + 0.6(*Total Employment*) + 0.1(*# of Vehicles*)
[1]	2	Y = 1.57(*Total Employment*) + 0.34(*Population*)
[1]	3	Y = 80.74 + 3.66(*Retail Employment*)
[1]	4	Y = 0.08(*Total Employment*) + 0.11(*Population*) + 3.99(*Retail Employment*)
[1]	5	Y = 0.64(*Total Population**(*Zonal Employment / Total Employment*))
[1]	6	Y = 20.8651 + 0.5241(*# of Vehicles*) + 5.5973(*Retail Employment*) + 0.2986(*Non-Retail Employment*) + 0.1441(*School Enrollment*)
[1]	7	Y = 67 + 0.6(*Total Employment*) + 0.1(*# of Vehicles*) + 0.1(*School Enrollment*) + 0.1(*# of Households*)
[3]	8	Y = (*Retail Employment*) + 0.25(*# of Households*) + 0.13(*Office Employment*) + 0.26(*Industrial Employment*) + 0.28(*Other Employment*)
[3]	9	Y = 2.79(*Retail Employment*) + 0.25(*# of Households*) + 0.26(*Office Employment*) + 0.43(*Industrial Employment*) + 0.57(*Other Employment*)
[3]	10	Y = 2.79(*Retail Employment*) + 0.25(*# of Households*) + 0.32(*Office Employment*) + 0.43(*Industrial Employment*) + 0.57(*Other Employment*)
[4]	11	Y = 0.962(*# of 1-Person Households*) + 1.824(*# of 2-Person Households*) + 2.65(*# of 3-Person Households*) + 3.128(*# of 4-Person Households*) + 3.486(*# of 5+-Person Households*)

Non-home-based (NHB) Regression – Attractions

References		Attraction Equations
[1]	1	Y = 1.57(*Total Employment*) + 0.34(*Population*)
[1], [2]	2	Y = 67 + 0.4(*Total Employment*) + 0.5(*# of Households*)
[1]	3	Y = 82.31 + 3.62(*Retail Employment*)
[1]	4	Y = 0.06(*Total Employment*) + 0.12(*Population*) + 3.46(*Retail Employment*)
[3]	5	Y = 2(*Retail Employment*) + 2.5(*Non-Retail Employment*) + 0.5(*# of Households*)
[3]	6	Y = 0.439(*Total Employment*) + 0.252(*Population*)
[1]	7	Y = 0.64(*Total Population**(*Zonal Employment / Total Employment*))
[1]	8	Y = 26.2343 + 0.5987(*# of Vehicles*) + 5.3545(*Retail Employment*) + 0.2326(*Non-Retail Employment*) + 0.0935(*School Enrollment*)
[4]	9	Y = 1.4(*Retail Employment*) + 0.5(*# of Households*) + 1.2(*Service Employment*) + 0.5(*Basic Employment*)
[4]	10	Y = 4.1(*Retail Employment*) + 0.5(*# of Households*) + 1.2(*Service Employment*) + 0.5(*Basic Employment*)
[3]	11	Y = 0.2231(*Total Employment*) + 5.996(*Retail Employment*) + 0.0213(*Non-Retail Employment*) - 0.7044(*Basic Employment*)
[3]	12	Y = 161.41 + 0.07(*Total Employment*) + 0.19(*Population*) + 4.68(*Retail Employment*) +4.54(*Wholesale Employment*)
[3]	13	Y = (*Retail Employment*) + 0.25(*# of Households*) + 0.13(*Office Employment*) + 0.26(*Industrial Employment*) + 0.28(*Other Non-Industrial Employment*)
[3]	14	Y = 2.79(*Retail Employment*) + 0.25(*# of Households*) + 0.26(*Office Employment*) + 0.43(*Industrial Employment*) + 0.57(*Other Non-Industrial Employment*)
[3]	15	Y = 2.79(*Retail Employment*) + 0.25(*# of Households*) + 0.32(*Office Employment*) + 0.43(*Industrial Employment*) + 0.57(*Other Non-Industrial Employment*)
[3]	16	Y = 0.65(*Retail Employment*) + 0.32(*Non-Retail Employment*) + 0.22(*# of Households*) + 0.11(*Basic Employment*)
[3]	17	Y = 1.30(*Retail Employment*) + 0.43(*Non-Retail Employment*) + 0.27(*# of Households*) + 0.16(*Basic Employment*)
[3]	18	Y = 1.95(*Retail Employment*) + 0.65(*Non-Retail Employment*) + 0.32(*# of Households*) + 0.22(*Basic Employment*)
[3]	19	Y = 4.00(*Retail Employment*) + 1.08(*Non-Retail Employment*) + 0.54(*# of Households*) + 0.32(*Basic Employment*)
[3]	20	Y = 4.65(*Retail Employment*) + 1.30(*Non-Retail Employment*) + 0.54(*# of Households*) + 0.22(*Basic Employment*)
[8]	21	Y = 8.479(*Retail Employment*) + 1.327(*# of Households*) + 0.518(*Service Employment*)
[8]	22	Y = 7.45(*Retail Employment*) + 0.823(*# of Households*) + 1.285(*Service Employment*)
[8]	23	Y = 6.471(*Retail Employment*) + 0.725(*# of Households*) + 0.833(*Service Employment*)
[8]	24	Y = 6.842(*Retail Employment*) + 0.908(*# of Households*) + 0.891(*Service Employment*)

Non-home-based (NHB) Regression – Attractions Continued

References	Attraction Equations	
[8]	25	Y = 1.007(*Retail Employment*) + 0.802(*# of Households*) + 3.347(*Service Employment*)
[8]	26	Y = 7.009(*Retail Employment*) + 0.682(*# of Households*) + 0.974(*Service Employment*)
[8]	27	Y = 6.525(*Retail Employment*) + 0.733(*# of Households*) + 1.041(*Service Employment*)
[9]	28	Y = 0.0575(*Population*) + 0.7593(*Retail Employment*) + 0.0846(*Commercial and Government Employment*) + 0.1053(*Other Employment*)

Home-based Shopping (HB-Shop) Regression – Productions

References	Production Equations	
[1]	1	Y = 0.86(*# of Vehicles*)
[1]	2	Y = 0.52(*Labor Force*)
[3]	3	Y = 0.59(*# of 1-Vehicle Households*) + 1.02(*# of 2-Vehicle Households*) + 1.44(*# of 3-Vehicle Households*)
[3]	4	Y = 0.71(*# of 1-Vehicle Households*) + 1.31(*# of 2-Vehicle Households*) + 1.68(*# of 3-Vehicle Households*)
[3]	5	Y = 0.59(*# of 1-Vehicle Households*) + 1.42(*# of 2-Vehicle Households*) + 1.76(*# of 3-Vehicle Households*)

Home-based Shopping (HB-Shop) Regression – Attractions

References	Attraction Equations	
[1]	1	Y = 33.61 + 3.07(*Retail Employment*)
[1]	2	Y = 2.99(*Retail Employment*) - 1.55(*Local Retail Commercial Area*)
[3]	3	Y = 0.2814(*Population*) + 5.8735(*Retail Employment*) - 0.6476(*Non-Retail Employment*) - 0.9724(*Basic Employment*)
[3]	4	Y = 41.96 + 0.04(*Total Employment*) + 0.10(*Population*) + 5.06(*Retail Employment*) + 4.98(*Wholesale Employment*)
[3]	5	Y = 0.33(*Retail Employment*)
[3]	6	Y = 2.31(*Retail Employment*)
[3]	7	Y = 4.04(*Retail Employment*)
[3]	8	Y = 6.04(*Retail Employment*)
[8]	9	Y = 7.80(*Retail Employment*)
[8]	10	Y = 7.78(*Retail Employment*)
[8]	11	Y = 8.97(*Retail Employment*)
[8]	12	Y = 6 67(*Retail Employment*)
[8]	13	Y = 6.78(*Retail Employment*)
[8]	14	Y = 8.42(*Retail Employment*)
[8]	15	Y = 8.17(*Retail Employment*)

Total Trips Generated Regression

Reference		Total Trips Equations
[10]	1	Y = 0.5(# of Households)*[0.30804 + 9.0858(Average # of Vehicles per Household) + 4.77048(Average Household Size)]
[10]	2	Y = 0.5(# of Households)*[-3.59856 + 9.4332(Average # of Vehicles per Household) + 2.2514(Average Household Size)]
[10]	3	Y = 0.5(# of Households)*[-1.97 + 8.86018(Average # of Vehicles per Household) + 1.199186(Average Household Size)]
[10]	4	Y = 0.5(# of Households)*[-5.832 + 10.0882(Average # of Vehicles per Household) + 2.74013(Average Household Size)]

Tanner Function General Form

$$FF_{ij} = a * Impedance^{b} * e^{c*Impedance}$$

Home-based Work (HBW) Tanner Functions

Reference		a	b	c
[1]	1	28,507	0	0
[1]	2	100	-0.02	-0.125
[1]	3	100	0	0
[2]	4	50,000	-0.0174	-0.0425
[3]	5	1	-2	0
[4]	6	1,064,931	-0.452	-0.063
[4]	7	1,046,313	0	0
[4]	8	997,214	-0.904	0.003
[4]	9	1,023,345	0	0
[5]	10	100,000,000	0	-0.0231

Home-based Other (HBO) Tanner Functions

Reference		a	b	c
[1]	1	139,173	-1	0
[1]	2	100	-1.30	-0.100
[1]	3	100	-1	0
[2]	4	150,000	-0.2560	-0.0886
[3]	5	1	2	0
[4]	6	1,072,144	-1.646	-0.070

Non-home-based (NHB) Tanner Functions

Reference		a	b	c
[1]	1	219,113	-1	0
[1]	2	100	-1.35	-0.100
[2]	3	100,000	0	0
[3]	4	1	1.9000	-0.1000
[4]	5	967,402	-3	0

Home-based Shopping (HB-Shop) Tanner Functions

Reference		a	b	c
[2]	1	200,000	0	0
[3]	2	1,185,375	-1.64	-0.170

Synthetic Friction Factors

Reference	[1]	[2]	[1]	[2]	[1]
Impedance (Minutes)	HBW		HBO		NHB
1	25,214	10,340	126,632	48,508	198,293
2	21,990	9,720	47,295	29,609	71,303
3	19,291	9,140	25,562	21,214	37,607
4	16,963	8,480	16,072	16,229	23,203
5	14,936	7,780	10,979	12,870	15,601
6	13,161	7,165	7,904	10,441	11,075
7	11,605	6,590	5,900	8,604	8,163
8	10,236	5,970	4,522	7,172	6,184
9	9,032	5,310	3,537	6,031	4,784
10	7,972	4,900	2,811	5,107	3,763
11	7,037	4,490	2,263	4,349	2,999
12	6,213	4,120	1,841	3,720	2,417
13	5,486	3,870	1,511	3,195	1,966
14	4,845	3,540	1,250	2,753	1,612
15	4,280	3,380	1,041	2,379	1,331
16	3,780	3,010	872	2,060	1,105
17	3,339	2,880	734	1,789	923
18	2,950	2,590	620	1,556	774
19	2,607	2,470	527	1,356	652
20	2,303	2,335	449	1,183	551
21	2,035	2,138	383	1,034	467
22	1,798	1,956	329	905	397
23	1,589	1,789	282	792	339
24	1,404	1,635	243	695	290
25	1,241	1,494	210	610	248
26	1,097	1,365	182	536	213
27	969	1,246	158	471	184
28	857	1,137	137	415	158
29	757	1,038	119	365	137
30	669	947	104	322	118
31	592		90		102
32	523		79		89
33	462		69		77
34	409		60		67
35	361		53		58
36	319		47		51
37	282		41		44
38	249		36		39
39	221		32		34
40	195		28		30
41	172		25		26
42	152		22		23
43	135		19		20

Synthetic Friction Factors Continued

Reference	[1] [2]	[1] [2]	[1]
Impedance (Minutes)	**HBW**	**HBO**	**NHB**
44	119	17	18
45	105	15	15
46	93	13	14
47	82	12	12
48	73	10	10
49	64	9	9
50	57	8	8
51	50	7	7
52	44	6	6
53	39	6	6
54	35	5	5
55	31	4	4
56	27	4	4
57	24	4	3
58	21	3	3
59	19	3	3
60	17	3	2

Link Performance Function General Form

$$t = t_0\left(\delta + \alpha\left(\frac{v}{c}\right)^{\beta}\right)$$

Single Purpose Link Performance Functions

Reference		A	β	δ
[1]	1	0.15	4	1
[1]	2	1	4	1
[1]	3	1	10	1
[1]	4	0.15	6	1
[1]	5	0.075	7	1
[2]	6	0.84	5.5	1
[1]	7	0.13	4	0.87
[2]	8	0.13	3	0.87

Freeway Link Performance Functions

Reference		A	β	δ
[3]	1	0.7	8.5	1
[1]	2	0	8	1
[1]	3	0.4374	-3	0.25
[1]	4	0.15	7	1
[2]	5	0.88	9.8	1
[2]	6	0.83	5.5	1
[2]	7	0.56	3.6	1

Arterial Link Performance Functions

Reference		α	β	δ
[3]	1	0.76	5.1	1
[1]	2	0	7	1
[1]	3	0.5184	-3	0.25
[1]	4	0.15	4	1

Multi-lane Link Performance Functions

Reference		α	β	δ
[3]	1	0.73	2.1	1
[2]	2	1	5	1
[2]	3	0.83	2.7	1
[2]	4	0.71	2	1

Signalized Link Performance Function

Reference	α	β	δ
[4]	1	0.05	10	1

Un-signalized Link Performance Function

Reference	α	β	δ
[4]	1	0.2	10	1

Appendix B: Input Data Availability

The typical lag, smallest geographic level, and most recent release date values are reported as of availability on September 24, 2008 for the CAMPO study area. For more populated areas, smaller geographic levels are available. Floor areas, basic, service, and industrial employment are not available for the studied data sources.

Typical Lag in Years Between Data Releases	10	1	5	1	Varies 1, 5, or 10	Varies 5 to 7	1/12	Varies 1/12, 1/4, 1, or 10	Varies 1/12, 1/4, 1, or 10
Input Variable	**Decennial Census**	**ACS**	**Economic Census**	**Population Estimates**	**USA Counties**	**NHTS**	**Demographics NOW**	**STATS Indiana**	**Hoosier Data.com**
Household Counts									
Total Population	Block (2000)	Unified School District (2007)		County (2007)	County (2007)		Block (August 2008)	Township (2007)	Metro (2007)
Group Quarters Population	Block (2000)	Unified School District (2007)			County (2000)		Block (August 2008)		
Average Number of Children Under Age 5	Block (2000)	Unified School District (2007)			County (2006)		Block (August 2008)	County (2006)	Metro (2006)
Number of Households	Block (2000)	Unified School District (2007)		County (2007)	County (2000)	Census Tract (2001)	Block (August 2008)	Metro (2007)	Metro (2000)
Number of 1-Person Households	Block (2000)	Unified School District (2007)			County (2006)	Census Tract (2001)	Block (August 2008)	Metro (2000)	Metro (2000)
Number of 2-Person Households	Block (2000)	Unified School District (2007)				Census Tract (2001)	Block (August 2008)		
Number of 3-Person Households	Block (2000)	Unified School District (2007)				Census Tract (2001)	Block (August 2008)		
Number of 4-Person Households	Block (2000)	Unified School District (2007)				Census Tract (2001)	Block (August 2008)		
Number of 5+-Person Households	Block (2000)	Unified School District (2007)				Census Tract (2001)	Block (August 2008)		

Input Variable	Decennial Census	ACS	Economic Census	Population Estimates	USA Counties	NHTS	Demographics NOW	STATS Indiana	Hoosier Data.com
Household Counts Continued									
Average Household Size	Block (2000)	Unified School District (2007)			County (2000)	Census Tract (2001)	Block (August 2008)	Metro (2007)	
Number of Vehicles Available	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)	Metro (2005)	
Number of 0-Vehicle Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Number of 1-Vehicle Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Number of 2-Vehicle Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Number of 3+-Vehicle Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Average Number of Vehicles per Household	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)	Metro (2005)	
Number of 0-Worker Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Number of 1-Worker Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		

Input Variable	Decennial Census	ACS	Economic Census	Population Estimates	USA Counties	NHTS	Demographics NOW	STATS Indiana	Hoosier Data.com
Household Counts Continued									
Number of 2-Worker Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Number of 3+-Worker Households	Block Group (2000)	Unified School District (2007)				Census Tract (2001)	Block Group (August 2008)		
Employment Counts									
Total Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (July 2008)
Average Number of Employees per Households	Block Group (2000)	Unified School District (2007)					Block Group (August 2008)	Metro (July 2008)	
Labor Force	Block Group (2000)	Unified School District (2007)			County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Average Number of Labor Force per Household	Block Group (2000)	Unified School District (2007)			County (2000)		Block Group (August 2008)	Metro (July 2008)	County (2000)
Retail Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (July 2008)
Construction Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Transportation Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)

Input Variable	Decennial Census	ACS	Economic Census	Population Estimates	USA Counties	NHTS	Demographics NOW	STATS Indiana	Hoosier Data.com
Employment Counts Continued									
Manufacturing Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Wholesale Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Public Administration Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Educational, Health, & Social Services Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Arts, Entertainment, Recreation, Accomodation, and Food Services Employment	Block Group (2000)	Unified School District (2007)	Metro (2002)		County (2006)		Block Group (August 2008)	Metro (July 2008)	Metro (2007 Q3)
Educational Enrollment									
Total School Enrollment	Block Group (2000)	Unified School District (2007)			County (2000)		Block Group (August 2008)	Metro (2007)	Unified School District (2006)
Primary School Enrollment	Block Group (2000)	Unified School District (2007)			County (2000)		Block Group (August 2008)		
Secondary School Enrollment	Block Group (2000)	Unified School District (2007)			County (2000)		Block Group (August 2008)		
College Enrollment	Block Group (2000)	Unified School District (2007)			County (2000)		Block Group (August 2008)		

Appendix C: Recommended CAMPO Travel Demand Model Parameters

Trip Generation

By Regression

HBW P =	Y = -2.90486 + 0.75377(*# of Vehicles*)
HBW A =	Y = 34.51357 + 1.47557(*Retail Employment*) + 0.4897(*Non-Retail Employment*)
HBO P =	Y = 18 + 2.7(*# of Vehicles*)
HBO A =	Y = 10(*Retail Employment*) + 0.5(*Non-Retail Employment*) + 1(*# of Households*)
NHB P =	Y = 0.08(*Total Employment*) + 0.11(*Population*) + 3.99(*Retail Employment*)
NHB A =	Y = 1.57(*Total Employment*) + 0.34(*Population*)

Trip Distribution

	Tanner Function Parameter		
	A	b	c
HBW	1	-2.1	-0.1
HBO	1	1.95	-0.1
NHB	967,402	-2.824	0.033

Impedance =	FFTT
Peak Hour =	17 to 18 (5 to 6 pm)
% AADT for Peak Hour =	8.5%

	Vehicle Occupancy Rate
HBW	1.25
HBO	1.82
NHB	1.43

Trip Assignment

	Link Performance Function Parameter
α	1
β	10
δ	1

Appendix D: Current CAMPO Input Data

Input data has been provided by the new CAMPO director Kent Anderson in 2007. Revisions to the CAMPO input data since this time are not reflected in this research.

Traffic Analysis Zone	Population	Total Employment	Retail Employment	Non-Retail Employment	Number of Vehicles	Number of 0-Vehicle Households	Number of 1-Vehicle Households	Number of 2-Vehicle Households	Number of 3+-Vehicle Households	Number of Households	Median Household Income ($1999)
1	1,200	149	11	138	1,097	9	73	147	215	456	$48,235
2	1,090	977	15	962	1,063	0	64	158	201	420	$55,924
3	1,626	89	21	68	1,477	13	107	253	241	631	$53,548
4	2,063	1,724	52	1,672	1,340	59	158	385	132	721	$46,947
5	1,434	94	36	58	1,151	3	139	153	216	510	$60,368
6	895	139	18	121	793	0	45	173	119	346	$81,595
7	163	73	0	73	157	0	15	21	29	68	$53,333
8	1,127	547	253	294	969	48	178	205	109	518	$50,577
9	1,711	1,723	170	1,553	1,590	15	81	270	272	612	$57,554
10	1,819	104	0	104	1,375	7	242	338	133	726	$78,318
11	3,564	100	3	97	2,362	45	380	614	238	1,287	$36,746
12	940	294	32	262	724	8	93	165	92	365	$51,977
13	601	365	36	329	476	11	80	97	55	240	$40,625
14	711	851	504	347	523	18	0	170	47	248	$74,412
15	1,840	2,377	954	1,423	1,480	9	256	291	178	733	$40,075
16	925	367	15	352	628	0	25	166	80	255	$49,917
17	1,074	165	51	114	849	10	102	173	120	407	$36,563
18	856	870	125	745	593	62	101	135	67	380	$30,625
19	776	55	2	53	518	42	105	116	53	307	$28,750
20	1,540	101	21	80	1,460	10	51	220	261	557	$51,042
21	1,028	171	11	160	767	24	149	106	113	379	$36,974
22	1,380	302	18	284	896	73	212	131	130	538	$29,038
23	552	436	67	369	421	16	55	80	61	217	$42,391
24	637	36,615	2,755	33,860	249	83	127	61	0	276	$16,304
25	688	124	19	105	536	0	31	137	75	241	$67,292

Traffic Analysis Zone	Population	Total Employment	Retail Employment	Non-Retail Employment	Number of Vehicles	Number of 0-Vehicle Households	Number of 1-Vehicle Households	Number of 2-Vehicle Households	Number of 3+-Vehicle Households	Number of Households	Median Household Income ($1999)
26	920	43	4	39	861	0	28	152	139	321	$54,886
27	778	2,875	124	2,751	449	55	104	115	28	304	$33,672
28	957	22	0	22	556	93	174	132	35	422	$31,058
29	1,534	733	221	512	888	85	283	185	55	597	$23,836
30	1,888	886	132	754	1,192	71	305	322	81	764	$27,072
31	1,263	446	74	372	688	117	370	123	24	657	$17,366
32	681	37	31	6	699	7	65	156	92	319	$70,156
33	1,756	385	199	186	1,279	52	354	284	110	792	$39,135
34	880	385	92	293	721	0	81	175	90	345	$59,286
35	752	12	0	12	696	0	15	150	109	291	$125,548
36	2,112	118	11	107	1,558	0	59	361	237	666	$85,954
37	1,024	3,570	1,995	1,575	696	21	207	171	43	452	$34,926
38	1,004	503	73	430	748	18	142	146	88	381	$34,288
39	898	329	138	191	635	23	130	132	70	350	$29,667
40	1,109	1,297	86	1,211	767	44	150	184	79	472	$33,462
41	1,226	3,176	248	2,928	913	46	205	257	57	578	$38,580
42	830	758	224	534	422	17	181	65	35	296	$32,955
43	823	106	15	91	681	0	120	197	50	341	$45,714
44	762	142	17	125	592	6	42	154	74	289	$62,000
45	2,186	1,063	87	976	1,394	35	406	383	72	907	$36,802
46	1,743	798	343	455	1,217	10	169	342	113	610	$61,840
47	495	862	727	135	325	50	105	98	8	268	$25,682
48	1,130	43	4	39	925	15	114	186	131	424	$65,313
49	1,372	164	12	152	812	176	284	206	33	670	$22,105
50	801	442	319	123	705	8	48	192	75	327	$59,091

Traffic Analysis Zone	Population	Total Employment	Retail Employment	Non-Retail Employment	Number of Vehicles	Number of 0-Vehicle Households	Number of 1-Vehicle Households	Number of 2-Vehicle Households	Number of 3+-Vehicle Households	Number of Households	Median Household Income ($1999)
51	1,631	24	12	12	1,269	9	103	338	148	604	$60,144
52	851	188	80	108	576	6	165	146	37	359	$34,145
53	1,248	130	3	127	752	83	147	214	59	513	$50,022
54	1,292	16	2	14	878	26	187	245	67	519	$44,231
55	1,077	158	23	135	863	20	113	233	89	480	$42,463
56	684	59	2	57	587	2	52	139	76	280	$51,328
57	980	42	14	28	908	0	65	166	123	339	$55,000
58	1,227	74	13	61	961	0	27	270	116	405	$51,447
59	1,206	67	14	53	1,107	14	73	176	189	461	$43,125
60	1,150	89	23	66	874	22	135	188	113	449	$30,833
61	883	184	2	182	808	0	73	145	124	335	$42,944
62	1,330	247	10	237	1,096	0	80	251	148	482	$53,170
63	603	72	5	67	555	0	30	111	86	225	$51,131
64	1,369	440	95	345	923	19	99	231	110	467	$37,721
65	800	74	25	49	706	0	31	116	120	281	$64,375
66	1,172	32	18	14	1,018	6	69	204	151	422	$53,214
67	1,319	312	48	264	998	15	116	183	160	478	$39,028
Σ =	77,986	70,215	10,754	59,461	58,792	1,736	8,575	12,989	7,081	30,380	--
Mean =	1,164	1,048	161	887	877	26	128	194	106	453	$47,371

Appendix E: Risk Assessment Results - Top 10 Most Congested Links and VHT Values by Model run with Random Normal Inputs

Ten model runs using randomly selected values pulled from normal distributions were applied to the CAMPO model. The top ten links as sorted by v/c ratio were recorded for each model run. The VHT values for each of these links, along with network values were also recorded for each model run.

Model Run

1				2			
Rank	Link ID	Street Name	v/c	Rank	Link ID	Street Name	v/c
1	664902	S.R. 46	1.20	1	10871959	Lincoln Park Dr.	1.36
2	672163	Washington St.	1.20	2	672163	Washington St.	1.24
3	10871342	S.R. 46	1.19	3	10873324	Lowell Rd.	1.21
4	10871959	Lincoln Park Dr.	1.18	3	10873327	Lowell Rd.	1.21
5	10873324	Lowell Rd.	1.17	5	664902	S.R. 46	1.21
5	10873327	Lowell Rd.	1.17	6	705467	Wildwood Pl.	1.20
7	665547	S.R. 46	1.16	7	10871342	S.R. 46	1.19
8	665041	Franklin St.	1.16	8	672206	Washington St.	1.18
9	10873321	Indianapolis Rd.	1.16	9	10873321	Indianapolis Rd.	1.17
9	10873328	Indianapolis Rd.	1.16	9	10873328	Indianapolis Rd.	1.17
9	10873325	Indianapolis Rd.	1.16	9	10873325	Indianapolis Rd.	1.17
9	665831	Indianapolis Rd.	1.16	9	665831	Indianapolis Rd.	1.17

Model Run

3				4			
Rank	Link ID	Street Name	v/c	Rank	Link ID	Street Name	v/c
1	10871959	Lincoln Park Dr.	1.49	1	10871959	Lincoln Park Dr.	1.77
2	705467	Wildwood Pl.	1.40	2	705386	Maple St.	1.64
3	705386	Maple St.	1.35	2	705492	Maple St.	1.64
3	705492	Maple St.	1.35	2	705552	Tipton Ln.	1.64
3	705552	Tipton Ln.	1.35	5	705467	Wildwood Pl.	1.60
6	665015	Washington St.	1.23	6	10873324	Lowell Rd.	1.28
7	665537	S.R. 46	1.23	6	10873327	Lowell Rd.	1.28
8	664902	S.R. 46	1.22	8	664902	S.R. 46	1.24
9	10873324	Lowell Rd.	1.21	9	10871342	S.R. 46	1.23
9	10873327	Lowell Rd.	1.21	10	672163	Washington St.	1.23

Model Run

5				6			
Rank	Link ID	Street Name	v/c	Rank	Link ID	Street Name	v/c
1	10873324	Lowell Rd.	1.30	1	10873324	Lowell Rd.	1.25
1	10873327	Lowell Rd.	1.30	1	10873327	Lowell Rd.	1.25
3	665041	Franklin St.	1.22	3	10871342	S.R. 46	1.21
4	10871342	S.R. 46	1.21	4	664902	S.R. 46	1.21
5	664902	S.R. 46	1.21	5	665547	S.R. 46	1.20
6	10873328	Indianapolis Rd.	1.20	6	672163	Washington St.	1.20
6	665831	Indianapolis Rd.	1.20	7	665041	Franklin St.	1.20
8	10873321	Indianapolis Rd.	1.20	8	10871959	Lincoln Park Dr.	1.20
8	10873325	Indianapolis Rd.	1.20	9	10873321	Indianapolis Rd.	1.18
10	665537	S.R. 46	1.17	9	10873328	Indianapolis Rd.	1.18
				9	10873325	Indianapolis Rd.	1.18
				9	665831	Indianapolis Rd.	1.18

Model Run

7				8			
Rank	Link ID	Street Name	v/c	Rank	Link ID	Street Name	v/c
1	10871959	Lincoln Park Dr.	1.38	1	10871959	Lincoln Park Dr.	1.36
2	705386	Maple St.	1.29	2	10873324	Lowell Rd.	1.25
2	705492	Tipton Ln.	1.29	2	10873327	Lowell Rd.	1.25
2	705552	Tipton Ln.	1.29	4	664902	S.R. 46	1.22
5	705467	Wildwood Pl.	1.25	5	665041	Franklin St.	1.21
6	672163	Washington St.	1.22	6	672163	Washington St.	1.21
7	10873324	Lowell Rd.	1.21	7	10871342	S.R. 46	1.21
7	10873327	Lowell Rd.	1.21	8	705467	Wildwood Pl.	1.19
9	664902	S.R. 46	1.20	9	665434	1st St.	1.18
10	10871342	S.R. 46	1.20	9	665528	Lafayette Ave.	1.18

Model Run

9				*10*			
Rank	Link ID	Street Name	v/c	Rank	Link ID	Street Name	v/c
1	10871959	Lincoln Park Dr.	1.53	1	664902	S.R. 46	1.23
2	705467	Wildwood Pl.	1.44	2	665015	Washington St.	1.23
3	705386	Maple St.	1.36	3	10873324	Lowell Rd.	1.22
3	705492	Maple St.	1.36	3	10873327	Lowell Rd.	1.22
3	705552	Tipton Ln.	1.36	5	10871342	S.R. 46	1.20
6	10873324	Lowell Rd.	1.26	6	665537	S.R. 46	1.19
7	10873327	Lowell Rd.	1.26	7	10873321	Indianapolis Rd.	1.17
8	665041	Franklin St.	1.23	7	10873328	Indianapolis Rd.	1.17
9	664902	S.R. 46	1.23	7	10873325	Indianapolis Rd.	1.17
10	10871342	S.R. 46	1.21	7	665831	Indianapolis Rd.	1.17

Model Run 1

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
664902	3,673	2,155	1,518	208,186	205,734	2,452
672163	560	128	433	208,186	208,319	-134
10871342	2,211	2,089	122	208,186	206,824	1,362
10871959	283	64	219	208,186	208,805	-619
10873324	617	165	452	208,186	208,143	42
10873327	187	34	153	208,186	208,324	-138
665547	395	218	177	208,186	208,913	-728
665041	296	100	195	208,186	208,337	-151
10873321	389	75	314	208,186	208,118	68
10873328	62	11	50	208,186	208,093	93
10873325	50	9	41	208,186	208,095	91
665831	698	201	497	208,186	207,866	320

Model Run 2

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	964	81	883	203,035	201,296	1,739
672163	717	74	643	203,035	179,640	23,395
10873324	861	189	672	203,035	203,153	-118
10873327	262	34	227	203,035	203,266	-231
664902	3,782	2,398	1,384	203,035	199,683	3,352
705467	903	493	410	203,035	201,701	1,334
10871342	2,214	2,103	111	203,035	202,372	663
672206	601	153	448	203,035	203,526	-491
10873321	424	75	349	203,035	203,646	-611
10873328	67	12	56	203,035	203,905	-869
10873325	55	9	45	203,035	203,899	-864
665831	778	209	570	203,035	203,987	-951

Model Run 3

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	2,304	133	2,171	220,316	213,620	6,696
705467	3,898	784	3,114	220,316	212,768	7,548
705386	253	20	232	220,316	219,999	317
705492	1,191	100	1,091	220,316	218,307	2,009
705552	556	46	510	220,316	219,616	700
665015	474	124	350	220,316	219,617	699
665537	657	406	251	220,316	220,503	-187
664902	4,039	2,482	1,557	220,316	216,549	3,767
10873324	824	200	623	220,316	219,848	468
10873327	250	35	215	220,316	220,258	58

Model Run 4

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	14,643	536	14,107	273,034	228,346	44,688
705386	2,177	28	2,149	273,034	270,803	2,231
705492	10,258	145	10,113	273,034	259,693	13,341
705552	4,787	62	4,725	273,034	267,860	5,174
705467	15,652	2,195	13,457	273,034	224,543	48,492
10873324	1,468	202	1,265	273,034	272,159	875
10873327	446	35	411	273,034	272,796	238
664902	5,061	3,229	1,832	273,034	267,616	5,419
10871342	3,187	2,839	348	273,034	270,491	2,543
672163	810	165	645	273,034	271,714	1,320

Model Run 5

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10873324	1,679	196	1,483	215,847	214,928	919
10873327	501	33	468	215,847	215,651	195
665041	476	191	285	215,847	216,475	-628
10871342	2,663	2,520	143	215,847	214,062	1,785
664902	3,872	2,913	959	215,847	211,667	4,179
10873328	87	12	75	215,847	216,087	-240
665831	1,050	221	829	215,847	215,375	472
10873321	535	76	459	215,847	215,224	623
10873325	69	10	59	215,847	216,168	-321
665537	410	308	102	215,847	216,029	-182

Model Run 6

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10873324	1,108	161	947	207,270	207,000	270
10873327	337	34	302	207,270	207,518	-248
10871342	2,699	2,333	366	207,270	205,926	1,344
664902	3,752	2,529	1,223	207,270	203,897	3,373
665547	542	266	276	207,270	207,285	-15
672163	563	142	421	207,270	206,854	416
665041	389	136	253	207,270	207,813	-543
10871959	348	66	282	207,270	207,150	120
10873321	441	76	366	207,270	206,744	526
10873328	70	12	58	207,270	207,613	-343
10873325	57	9	48	207,270	207,591	-321
665831	832	890	-58	207,270	207,813	-543

Model Run 7

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	1,052	90	963	214,277	212,001	2,276
705386	163	17	145	214,277	215,287	-1,010
705492	766	87	680	214,277	213,653	624
705552	358	40	318	214,277	214,419	-141
705467	1,348	580	768	214,277	210,767	3,510
672163	660	136	524	214,277	214,119	158
10873324	857	171	686	214,277	213,760	517
10873327	260	33	227	214,277	213,994	284
664902	3,581	2,379	1,203	214,277	210,715	3,562
10871342	2,432	2,301	131	214,277	212,280	1,997

Model Run 8

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	942	80	862	214,585	213,053	1,531
10873324	1,179	225	954	214,585	214,039	546
10873327	358	37	321	214,585	214,848	-264
664902	4,095	3,139	956	214,585	210,367	4,218
665041	441	122	319	214,585	214,647	-62
672163	622	146	475	214,585	214,612	-27
10871342	2,582	2,323	258	214,585	213,020	1,565
705467	854	554	300	214,585	213,680	905
665434	467	108	358	214,585	215,078	-493
665528	487	118	369	214,585	215,032	-447

Model Run 9

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
10871959	2,975	149	2,826	226,714	218,101	8,613
705467	5,172	931	4,241	226,714	216,474	10,240
705386	277	21	256	226,714	225,478	1,235
705492	1,306	101	1,205	226,714	224,182	2,532
705552	609	45	564	226,714	225,972	742
10873324	1,216	169	1,047	226,714	225,707	1,007
10873327	369	36	333	226,714	226,296	418
665041	492	98	394	226,714	226,296	418
664902	4,379	2,832	1,547	226,714	222,376	4,338
10871342	2,651	2,414	237	226,714	224,602	2,112

Model Run 10

Link ID	Link VHT w/o Project	Link VHT w/ Project	Link VHT Savings	Network VHT w/o Project	Network VHT w/ Project	Network VHT Savings
664902	4,512	2,128	2,384	205,492	201,484	4,008
665015	476	133	342	205,492	204,749	742
10873324	895	210	685	205,492	205,196	295
10873327	272	34	238	205,492	204,314	1,178
10871342	2,349	2,153	197	205,492	203,927	1,565
665537	498	217	281	205,492	204,653	838
10873321	418	75	343	205,492	205,048	443
10873328	66	12	55	205,492	205,446	45
10873325	54	9	45	205,492	204,482	1,010
665831	774	208	566	205,492	205,240	252

Appendix F: Change in Inputs

TAZ #	Total Employment	Retail Employment	Non-Retail Employment
1	-48	-3	-45
2	210	255	-45
3	-21	1	-22
4	1,856	189	1,667
5	35	-1	36
6	-15	-7	-8
7	-73	0	-73
8	150	142	8
9	-12	104	-116
10	-75	0	-75
11	-23	-1	-22
12	-3	64	-67
13	-48	13	-61
14	340	196	144
15	-224	127	-351
16	756	20	736
17	28	12	16
18	-9	56	-65
19	42	-2	44
20	-32	-5	-27
21	96	45	51
22	-1	-3	2
23	26	13	13
24	-30,667	-1,568	-29,099
25	-59	0	-59
26	-4	12	-16
27	-220	321	-541
28	0	0	0
29	1,587	568	1,019
30	0	10	-10
31	0	-40	40
32	0	-23	23
33	24	-51	75
34	20	-12	32
35	0	3	-3
36	150	0	150
37	-209	146	-355
38	676	201	475

TAZ #	Total Employment	Retail Employment	Non-Retail Employment
39	-191	-98	-93
40	-99	-50	-49
41	650	217	433
42	0	-38	38
43	0	-13	13
44	0	-2	2
45	0	-16	16
46	1	-39	40
47	-63	-151	88
48	5	0	5
49	0	0	0
50	0	-22	22
51	0	6	-6
52	255	41	214
53	0	4	-4
54	4	0	4
55	0	-11	11
56	-5	0	-5
57	-8	-1	-7
58	-13	-4	-9
59	24	19	5
60	1	8	-7
61	-180	2	-182
62	-41	-10	-31
63	-13	1	-14
64	0	-9	9
65	-14	25	-39
66	4	0	4
67	29	18	11
Σ =	-25,401	659	-26,060
Mean =	-379	10	-389