

Determining the significance of a consistent mode choice between activity-based models and  
agent-based microsimulation tools

Christopher Day

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Brigham Young, Chair  
Karl G. Maeser  
Harvey Fletcher

Department of Civil and Construction Engineering  
Brigham Young University

Copyright © 2022 Christopher Day

All Rights Reserved

## ABSTRACT

Determining the significance of a consistent mode choice between activity-based models and agent-based microsimulation tools

Christopher Day  
Department of Civil and Construction Engineering, BYU  
Master of Science

This document describes the use of a new  $\text{\LaTeX}$  template for writing masters theses at Brigham Young University. This was made necessary by the relative age and complication of prior templates. Additionally, the new `cosmodown` R package contains an implementation of this thesis in R markdown, with the help of the `thesisdown` package.

The abstract of a thesis should describe the motivation, objective, overall results, and central findings of the thesis. It may have multiple paragraphs if necessary.

Keywords: Four Step Model, Discrete Choice Model, Activity-based Model

## TABLE OF CONTENTS

Title Page .....	i
Abstract .....	ii
Table of Contents .....	iii
List of Tables .....	vi
List of Figures .....	vii
1 Outline .....	1
2 Introduction .....	3
2.1 Problem Statement .....	3
2.2 Purpose of Research .....	3
2.3 Outline of Research .....	3
3 Literature Review .....	4
3.1 Overview .....	4
3.2 Mode Choice in the Four Step Model .....	4
3.3 A Brief History of Mode Choice in Discrete Choice Models .....	5
3.4 An Extensive Look into Mode Choice in Activity-Based Models .....	6
3.4.1 Simplified main tour model .....	8
3.4.2 Two-tier nested logit model .....	9
3.4.3 Simplified main tour mode and conditional trip-level mode .....	9
3.4.4 An activity-based model with exogenous mode choice .....	10
3.4.5 Simulation-based tour-based mode choice .....	10
3.4.6 Combinatorial tour-based mode choice .....	11
3.4.7 Dynamic tour-based mode choice .....	11
3.5 The Mode Choice Models used in ActivitySim .....	12
3.5.1 ActivitySim's Tour Mode Choice Model .....	13

3.5.2	ActivitySim’s Tour Mode Choice Utility Functions .....	14
3.5.3	ActivitySim’s Trip Mode Choice Model .....	17
3.5.4	ActivitySim’s Trip Mode Choice Utility Functions .....	18
3.6	The Mode Choice Models used in Agent-based Models .....	21
3.7	The Mode Choice Model used in MATSim .....	21
3.8	The Mode Choice Model used in BEAM .....	22
3.9	Previous Attempts to Enhance Mode Choice in Agent-based Models .....	24
3.10	Mode Choice Consistency between ActivitySim and BEAM .....	26
3.11	Summary .....	27
4	Methods .....	28
4.1	Overview .....	28
4.2	Creating and Setting up the Test Scenario .....	28
4.2.1	Test Scenario Region .....	29
4.2.2	Setting up and Running ActivitySim .....	31
4.2.3	Setting up BEAM .....	33
4.3	Improved Mode Choice in BEAM .....	41
4.3.1	Purpose-based Decisions .....	41
4.3.2	Additional Attributes in Utility Equation .....	48
4.3.3	Carpool Alternatives .....	52
4.3.4	Mode Choice Algorithms .....	52
4.4	Data Validation and Calibration .....	52
4.4.1	ActivitySim Path Variable Validation .....	52
4.4.2	Mode Choice Calibration .....	52
4.5	Path, Person, and Location Analysis .....	52
4.6	Summary .....	52
5	Results .....	54
6	Discussion .....	55

7	Conclusions .....	56
8	About .....	57
8.1	Usage .....	57
8.2	Render book .....	57
8.3	Preview book .....	58
9	Cross-references .....	59
9.1	Chapters and sub-chapters .....	59
9.2	Captioned figures and tables .....	59
10	Parts .....	62
11	Footnotes and citations .....	63
11.1	Footnotes .....	63
11.2	Citations .....	63
12	Blocks .....	64
12.1	Equations .....	64
12.2	Callout blocks .....	64
13	Sharing your book .....	65
13.1	Publishing .....	65
13.2	404 pages .....	65
13.3	Metadata for sharing .....	65
	References .....	67

## LIST OF TABLES

4.1	A Subset of the Population Plans .....	34
4.2	A Subset of the Person Attributes File .....	35
4.3	A Subset of the Households File .....	36
4.4	A Subset of the Vehicles File .....	36
4.5	A Subset of the Vehicle Types Input File .....	37
4.6	A Subset of the Vehicle Types Input File .....	38
4.7	Cost coefficient values for some modes from the utility parameters csv file .....	50
9.1	Here is a nice table! .....	61

## LIST OF FIGURES

3.1	Tour Mode Choice Nested Logit Model (MTC 2012). . . . .	14
3.2	Trip Mode Availability based on Tour Mode (MTC 2012). . . . .	17
3.3	Mode choice between ActivitySim and BEAM. . . . .	24
4.1	The TAZs of each county in the region of study. . . . .	30
4.2	Transportation network for area of study. . . . .	39
4.3	Close up of the transportation network of Salt Lake City. . . . .	40
4.4	BEAM code flow chart behind the default multinomial logit choice model. . . . .	43
4.5	BEAM code flow chart behind the latent class choice model. . . . .	45
4.6	BEAM code flow chart behind the tour purpose choice model. . . . .	47
9.1	Here is a nice figure! . . . . .	60

## 1 OUTLINE

Here is my outline for the entire thesis. Hopefully this will give me a rough idea of what to write about.

1. Introduction (*5pg*)
  - a. First/Last mile transit
2. Literature Review (*20-30pg*)
  - a. Discuss Mode Choice (4-step model, ABMs, microsimulations)
  - b. Discuss ActivitySim (modechoice)
  - c. Discuss MATSim (modechoice)
  - d. Discuss BEAM (modechoice)
  - e. Figure explaining tour type and first trip difference?
  - f. Variables showing path, person, location?
3. Methods
  - a. What did we add to BEAM?
  - b. Algorithms
  - c. How did we make scenario?
  - d. What scenarios did we run? SLC + SL / Data & Calibration
  - e. What tests did we run? (path, person, location effects)



#### 4. Results

- a. Describe tests and findings

#### 5. Discussion

- a. What does it mean?
- b. Limitations --> Future research recommendation

#### 6. Conclusions

- a. Wider implications
- b. First mile last mile
- c. climate change lol

## **2 INTRODUCTION**

In the intro talk about the tscore setup, along with the desire to understand on-demand transit vehicles. Talk about how a consistent mc will help better model these on demand transit vehicles.

### **2.1 Problem Statement**

### **2.2 Purpose of Research**

### **2.3 Outline of Research**

### **3 LITERATURE REVIEW**

#### **3.1 Overview**

Mode choice models are essential components in the transportation planning process. Key investment decisions along with other transportation decisions rely on the determination of individual mode choices within a population. Whether mode choice is estimated in a four step model, activity-based model, or an agent based model, its level of significance remains the same.

#### **3.2 Mode Choice in the Four Step Model**

The four step model (FSM) has been the primary tool in person travel demand modeling since its development (McNally 2000). Although travel is theorized to be derived from activity participation, the FSM focuses on modeling with trip-based not activity-based travel methods. The methodology presented in the FSM is mostly universal. The FSM can be divided up into two stages. The first defining the traveler and land use characteristics and the second where demand is loaded onto a transportation network and travel characteristics are determined (McNally 2000).

As its name suggests, the FSM includes four different steps: trip generation, trip distribution, mode choice, and route choice. In trip generation the magnitude of total trips is estimated and in trip distribution those trips are disbursed along different directions of travel. In mode and route choice, a multitude of variables are used to determine specifically how travel occurs.

Of the four steps in the FSM, mode choice in particular is a diverse, variable, and disaggre-

gate step among different modeling applications. According to McNally (2000), mode choice is almost exclusively modeled on a disaggregate level within separate choice-based sampling. Choice probabilities are calculated on the individual trip level. McNally (2000) also mentions that with transit, carpooling vehicles, automobile tolls, and other new factors the mode choice algorithm can become extensive and difficult to determine. One common method used to estimate mode choice probabilities is a nested logit model (see Section 3.4). These mode choice models can reflect trip-maker characteristics as well as multiple performance variables (Ortuzar and G. Willumsen 1994; McNally 2000). Most applications of the FSM require multiple iterations of the mode choice and trip distribution steps in order to approximate formal convergence. Like with the nested logit model, discrete choice modeling has historically been the preferred method in estimating mode choice in the FSM as well as in other modeling applications.

### **3.3 A Brief History of Mode Choice in Discrete Choice Models**

Various choice modeling systems have been used to accurately determine mode choices. The beginning of such analyses started with the developments of McFadden et al. (1973), who outlined a general procedure for modeling population choice behavior from distributions of individual decision rules. Before then, discrete choices were simply modeled using a random utility model (RUM) framework. The interpretation assumed that each individual carried their own utility function distribution and selected one randomly; according to Manski (2001), McFadden reinterpreted the random assignment of the utility functions across the population itself rather than the individual. This yielded an entirely new view of discrete choice modeling, and a more general version of the multinomial logit model (MNL) was introduced. This new MNL model, also called the conditional logit analysis, was the beginning of modern mode choice analysis. Goulias, Davis, and Bhat (2020) describe that McFadden's MNL model, the nested logit model, and the multinomial probit model (MNP) were the main engines in mode choice and most discrete choice analyses up through the 1990s.

Although the developments of McFadden are still in effect today, mode choice analysis has

evolved greatly in the last few decades. Goulias, Davis, and Bhat (2020) explain that up through 2010, mixed multinomial logit (MMNL) models and multiple discrete-continuous (MDC) choice models were the main models in practice. In addition, Bhat (1995) estimated a disaggregate mode choice model to explain how new and updated services affected ridership and mode choice on intercity travel. His heteroscedastic extreme value model overcame the independence of irrelevant alternatives (IIA) property and allowed more flexibility among the alternatives than the nested logit model. Ben-Akiva, Mcfadden, et al. (2002) developed a hybrid choice model (HCM) that went beyond the standard RUM as it included latent variables, latent classes, and a flexible error structure. It implemented multiple facets of discrete choice modeling and more accurately represented the behaviors of the individuals. Pinjari et al. (2007) used a simultaneous mixed logit model to understand how commuters' mode choice is affected by the surrounding build environment.

According to Goulias, Davis, and Bhat (2020), from 2010 to 2015 joint models of data with mixed dependent variables arrived. Vij, Carrel, and Walker (2013), for example, used a behavioral mixtures model, composed of a latent class choice model (LCCM) and a continuous logit mixture model, to extrapolate unobserved modality styles and their effects on travel decisions. Paulssen et al. (2014) focused on improving the integrated choice and latent variable (ICLV) model of McFadden (1986) and Ben-Akiva, Walker, et al. (2002) by including a value system among the individuals. Overall, the general advancement within mode choice and discrete choice modeling is comprehensive across recent years. Similarly, the variability among mode choice analysis can specifically be seen within activity-based modeling.

### **3.4 An Extensive Look into Mode Choice in Activity-Based Models**

A significant part of activity-based modeling is the mode choice decision within the model. Bhat and Koppelman (1999) stated that originally travel demand models used individual trips as the primary unit of analysis. In recent years though, there has been a shift toward activity-based travel demand modeling. According to Eluru et al. (2010), this shift has occurred due to improved understanding of activity-based travel, global climate concerns, and advances in microsimulation

techniques. With this shift to activity-based travel modeling, FSMs have evolved immensely, and especially within the last decade (Hasnine and Nurul Habib 2021). A major maturing in activity scheduling, activity duration, activity generation, and location choice has occurred and resulted in advanced activity-based modeling. Although these components have evolved significantly, the mode choice decision has remained relatively simple and constant.

In general, the two types of mode choice models commonly applied in travel demand modeling are trip-based and tour-based models. Trip-based mode choice models select modes based on the characteristics of individual trips, independent of other existing trips. A tour is defined as a chain of trips that end in the same location as they start (John L. Bowman et al. 1999). Tour-based models are based on the characteristics of the trips' overarching tour, meaning trip modes are not independently calculated. Most activity-based models commonly implement the tour-based mode choice structure (Hasnine and Nurul Habib 2021).

The most common use of the tour-based mode choice model is the “simplified tour-based” mode choice model. The simplified tour-based model has been used in a variety of studies (Arentze and Timmermans 2000; J. L. Bowman and Ben-Akiva 2001; Ram M. Pendyala et al. 1997). This simplified model hinges on trip-based mode choice where a single mode is used to define the entire tour opposed to a string of possible mode option combinations. For example, a simplified tour-based mode choice would define the tour-mode as “public transit” whereas a true tour-based model would define the tour-mode as “public transit - ride hail - public transit”. Simplified tour-based models overlooks mobility attributes as well as the dynamics that exist between trips on a tour (Hasnine and Nurul Habib 2021).

There have been many attempts during the last few decades to include trip and tour based mode choice models into activity-based models. As a way to better understand the diversity that exists with tour-based mode choice in activity-based models, Hasnine and Nurul Habib (2021) conducted research on a variety of activity-based models from the years 1995 to 2020 and found that seven specific mode choice analyses have been used. All seven techniques are framed about the tour-based methodology further reinforcing the knowledge that “the tour-based approach is the

most relevant to the activity-based modelling framework” (Hasnine and Nurul Habib 2021). These seven categories are further discussed in the following subsections.

### **3.4.1 Simplified main tour model**

The simplified main tour model is another name for the simplified tour-based model previously discussed and implemented in most activity-based modeling applications (Arentze and Timmermans 2000; J. L. Bowman and Ben-Akiva 2001; Ram M. Pendyala et al. 1997). In the A Learning-Based Transportation Oriented Simulation System (ALBATROSS) developed by Arentze and Timmermans (2000), for example, when individuals had a work activity in their daily plan, a transport mode was determined for the work travel purpose. A singular tour mode value was selected and given to that individual. In cases where households had less cars than people, a car tour mode assigned to one individual in a household may prevent a car mode being given to another individual in that same household. In this way, the simplified main tour model keeps track of car ownership and availability in a logical way. The ALBATROSS system selected the tour mode based on a subset of rules and decision trees programmed into the model. This means that the ALBATROSS system only allowed each trip within a tour to have the same mode as the tour mode (Arentze and Timmermans 2000).

Since a singular mode value is assigned to the tour in the simplified main tour model method, no flexibility exists for trip mode selection. Once a tour mode is determined, that same mode is held constant for each trip within that tour. Although not entirely realistic, many modelers prefer this method as it is a computationally easier task than a tour-based mode choice that permits trip modes to be different than their overarching tour mode. Realistically, however, individuals often vary their modes on different trips within a tour. For example, those who walk or bike or drive to transit are unable to do so in a simplified main tour model. Using the simplified main tour mode choice model may be popular and computationally simple, but it is not dynamic or realistic in representing travel behavior (Hasnine and Nurul Habib 2021).

### **3.4.2 Two-tier nested logit model**

The two-tier nested logit model first completes an upper-level location choice followed by a lower-level mode choice which corresponds to the destination's location. One example of this approach being implemented in the literature is the Florida Activity Mobility Simulator (FAMOS) model (Ram M. Pendyala et al. 2005). The FAMOS model uses both the Household Attributes Generation System and the Prism-Constrained Activity-Travel Simulator in order to model activity patterns. The sub models within FAMOS include an activity-type choice, a joint destination and mode choice, and an activity duration choice. Additionally, FAMOS mode choice model first determines the destination location, and then selects a mode dependent on the destination location. Although it uses a two-tier nested logit model, Hasnine and Nurul Habib (2021) states that a tour-based mode choice is not “explicitly identified” in the FAMOS model. At its development, FAMOS gave promising results in the effectiveness of an activity-based modeling system (Ram M. Pendyala et al. 2005). It did not however include an extensive and powerful mode choice model.

### **3.4.3 Simplified main tour mode and conditional trip-level mode**

The simplified main tour mode and conditional trip-level mode choice model has two parts. First a tour mode is estimated. Then, conditional on the estimated tour mode and attributes of the activity, trip modes are estimated. One example of a model that uses this structure is called DaySim (John L. Bowman, Bradley, and Gibbs 2006). DaySim follows the same activity scheduling approach developed by J. L. Bowman and Ben-Akiva (2001). The DaySim model is a microsimulation model that works off of four levels: long term person/household decisions, single day activity patterns, tour-level decisions, and trip-level decisions. The mode choice structure within DaySim involves a series of logsums and approximated logsums in order to econometrically estimate individual mode choice. More specifically, first a logsum is used to determine an overarching tour mode. This mode is used for all trip modes under that tour for the first steps within the model (John L. Bowman, Bradley, and Gibbs 2006). Later, however, trip modes are updated individually based on the tour mode value, origin and destination locations, and start times. The mode choice approach in DaySim



is effective in that it relies heavily on fundamental econometric theories and allows trip modes to vary from tour modes (Hasnine and Nurul Habib 2021).

Other examples in the literature also use the simplified main tour mode and conditional trip-level mode choice structure, however only the details relating to ActivitySim's mode choice structure will be discussed further (see Section 3.5)

#### **3.4.4 An activity-based model with exogenous mode choice**

The idea behind this mode choice model is that the mode choice decision is not estimated within the activity-based model, but estimated outside of it. One example of an activity-based model with the mode choice decision estimated exogenously is the comprehensive utility-based system of activity-travel scheduling options modelling (CUSTOM). The research done by Nurul Habib (2017) using the CUSTOM system assumed an exogenous mode choice. This meant that the mode specific travel time used to estimate travel behavior was an explanatory variable in the model. Nurul Habib (2017) mentions that the exogenous mode choice was a limitation in their research and further research is needed to specify how to seamlessly align the exogenous mode choice into the CUSTOM structure. Hasnine and Nurul Habib (2021) also mentions that a limitation to this mode choice model is that the mode is not modeled endogenously with other activity attributes.

#### **3.4.5 Simulation-based tour-based mode choice**

The simulation-based tour-based mode choice model is a simulation-based approach to estimating modes when large sets of modal alternatives are present. An example of this approach in the literature is a model used by Miller, Roorda, and Carrasco (2005) called the Travel/Activity Scheduler for Household Agents (TASHA). TASHA is both agent-based and activity-based where decisions are modeled using a random utility framework. In addition, TASHA is both chain-based and trip-based meaning that individual trips have individual modes chosen, but tour modes are constructed in a chain-based manner. Utility functions are used to calculate the utility of both individual trip modes as well as chains of modes used in the tour-based mode choice model. TASHA adopts a

probit modeling structure in order to determine probabilities of modal alternatives (Miller, Roorda, and Carrasco 2005; Hasnine and Nurul Habib 2021). This simulation approach is computationally heavy but is effective in providing unique trip modes based on chain-based tour modes.

Another example of a simulation-based mode choice model is presented by Eluru et al. (2010). Eluru et al. (2010) developed a joint multiple discrete continuous extreme value (MDCEV) framework to model the individual's choices across five dimensions: activity type, time of day, mode, destination, and time use. The MDCEV framework aimed to model activity travel choices simultaneously.

### **3.4.6 Combinatorial tour-based mode choice**

The combinatorial tour-based mode choice model simulates all combinations of modes within any given tour. A utility value is then calculated for the each combination of trip modes and a logit model is used to determine which combination any given agent will select. This method is computationally heavy, but allows the possibility for any combination of modes to exist (Hasnine and Nurul Habib 2021). An example of this model being used in practice is the research conducted by Vovsha et al. (2017).

### **3.4.7 Dynamic tour-based mode choice**

Dynamic tour-based mode choice models are programmed with dynamic discrete choice modeling. The central idea is to model sequential discrete choices in order to determine current and future outcomes. Saleem, Västberg, and Karlström (2018) provides an example of implementing a dynamic tour-based mode choice model in the random utility based travel model named SCAPER. SCAPER makes decisions sequentially in time based on a random utility theory. MATSim also implements the SCAPER model by generating activity schedules, activity durations, activity locations, and mode choices. MATSim iteratively updates the mode choice based on the best modal option available, and continues to iteratively update mode choice until stochastic equilibrium is reached (Hasnine and Nurul Habib 2021). More intimate details inside the MATSim mode choice

model will be discussed later on (see Section 3.6).

Overall, the seven categories of mode choice presented by Hasnine and Nurul Habib (2021) gives a good overview on the details behind existing mode choice structures within activity-based models. It is shown that although trip-based models have been used, tour-based mode choice models are the prominent form of mode choice modeling in activity-based models. ActivitySim is the activity-based model used in this research, and so the details of its mode choice structure will be analyzed more closely in Section 3.5.

### **3.5 The Mode Choice Models used in ActivitySim**

The activity-based model used in this research to generate activity plans is ActivitySim. ActivitySim is an activity simulator used to generate plans for millions of agents each with their own demographic attributes (Gali et al. 2008). Instead of independently modeling each trip, ActivitySim simulates each individual and their daily travel diaries and schedules. Since ActivitySim forecasts travel based on a disaggregate population, more realistic travel estimates are generated than can be generated by forecasting travel using aggregate trips. Long term decisions are made first, and then shorter term decisions are calculated based on the long term decisions (“ActivitySim: An Advanced Activity-Based Travel Demand Model Built by and for Users” 2021). Overall, ActivitySim is an advanced activity-based model that is cost effective, user friendly, and efficient at forecasting travel behavior.

As discussed in Section 3.4.3, ActivitySim adopts a simplified main tour mode model and a conditional trip-level mode model for its mode choice framework. The specifics of this dual level mode choice is described further in its application by MTC (2012). The details of this specific mode choice framework is discussed further.

The mode choice model in ActivitySim is multifaceted between tours, trips, and purpose. In other words, there is a mode choice model that determines the *primary* mode for each tour and a separate mode choice model that determines the mode for each trip within each tour (MTC 2012). The tour choice is the upper-level choice whereas the trip mode choice is the lower-level choice that

is conditional upon the upper-level choice. Two levels of mode choice is a more advanced than the simplified main tour mode choice model.

This overarching mode choice structure exists separately for each purpose. This means that since there are ten purposes specified in ActivitySim, there are ten tour mode choice models and 10 trip mode choice models, making a total of 20 mode choice models!

Multiple travel statistics are calculated at the trip and the tour level. At the tour level the tour modes, stop frequencies, and stop locations are estimated. Then, at the trip level the departure times, trip modes, parking for automobiles, and vehicle assignments are determined (“ActivitySim: An Advanced Activity-Based Travel Demand Model Built by and for Users” 2021).

### **3.5.1 ActivitySim’s Tour Mode Choice Model**

The tour mode choice model assigns the primary mode that is used to get from origin to destination. This level accounts for variables that affect the entire tour. The tour mode decision also affects the alternative values that are available for each trip. The specific details that determine which mode is used for each tour have to do with whether the tour is done by private car or by public transit, walking, or biking; whether there is carpooling available; and whether the transit mode is accessible by foot or by car (MTC 2012).

The specific choice model used to determine the overarching tour’s mode is a nested logit model. It separates similar modes into the differing bins as to “more accurately model the cross-elasticities between the alternatives” (MTC 2012). Figure 3.1 shows all 18 modes within ActivitySim and how they fit within the nested logit structure.

The first decision of modes is whether it is an automobile, a non-motorized vehicle, or a transit type. In a broad sense, the second level of nesting for automobile and non-motorized has to do with vehicle occupancy. If it is an automobile, three secondary bins exist: drive-alone, shared-ride with two people, and shared-ride with three or more people. Then, in the third level of nesting, two bins exist having to do with if the mode is free or not. On the other hand, the non-motorized section is only divided up into the walk or bike mode on a secondary level, with no options on the third

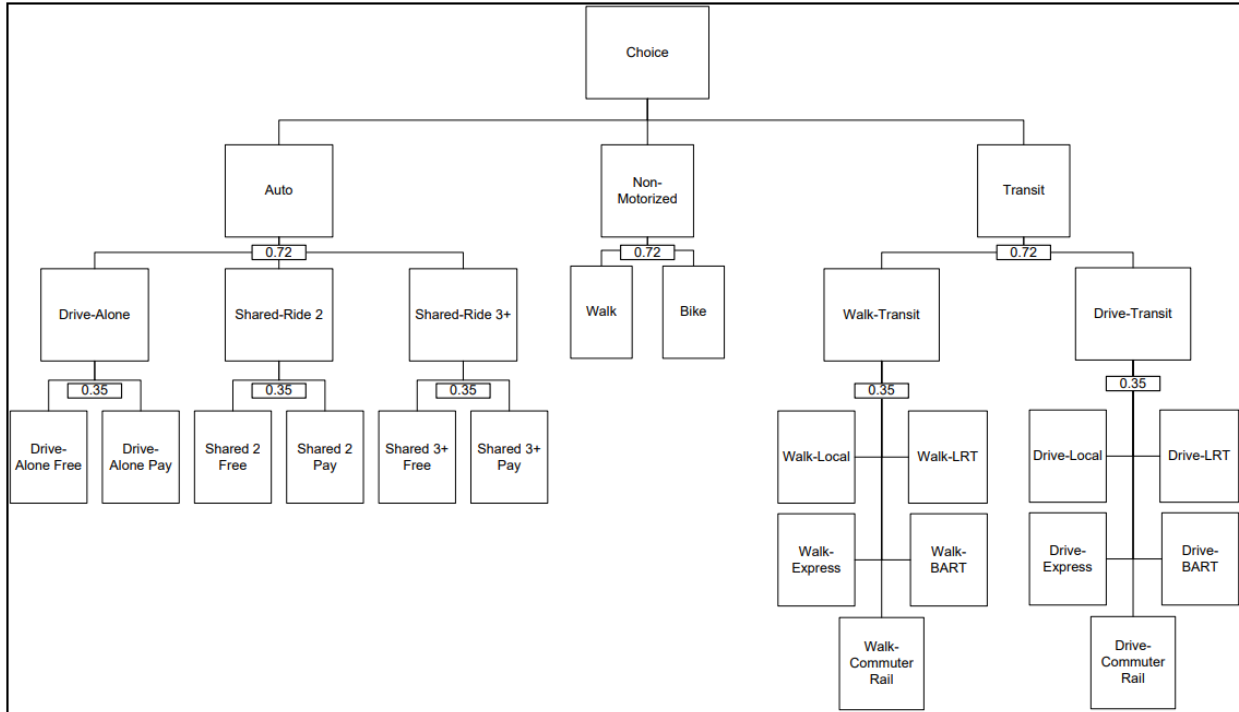


Figure 3.1: Tour Mode Choice Nested Logit Model (MTC 2012).

level. The transit nest is divided into walk-transit and drive transit on the secondary level. There are then five more bins within each one of those transit options having to do with which specific transit type is involved, like LRT, express, etc. The logsum coefficients for the first nest is 0.72 and for the second nest 0.35 (MTC 2012).

The utility functions for each mode are shown in Section 3.5.2. The primary variables in the utility functions are in-vehicle time, other travel times, cost, characteristics of the destination zone, demographics, and auto ownership of the household. Each mode has different parameter values for each one of the variables, for each tour purpose value.

### 3.5.2 ActivitySim's Tour Mode Choice Utility Functions

The utility functions for the auto, nonmotorized, and transit modes are presented in this section. All  $\beta$  and  $ASC$  values are meant to represent parameter values. All other variables are dependent on other things, like time for example. (The utility equations are deduced from information from MTC (2012)).

## Auto

$$V_{DriveAlone} = IVTT(\beta_{IVTT}) + ASC_{DA1619}$$

$$V_{Shared2} = IVTT(\beta_{IVTT}) + OVTT_w(\beta_{OVTT_w}) + ASC_{DA16P} + ASC_{HHS} + ASC_{ITC} + ASC_{JTC}$$

$$V_{Shared3P} = IVTT(\beta_{IVTT}) + OVTT_w(\beta_{OVTT_w}) + ASC_{DA16P} + ASC_{HHS} + ASC_{ITC} + ASC_{JTC}$$

- $IVTT$  is the in-vehicle travel time (min),
- $OVTT_w$  is the initial wait time (min);  $\beta_{OVTT_w}$  is different depending on if  $OVTT_w$  is above or below 10 minutes,
- $ASC_{DA1619}$  is a demographic variable for individuals between 16-19 years old,
- $ASC_{DA16P}$  is a demographic variable for individuals 16 years old or older,
- $ASC_{HH}$  is a demographic variable for household size; different depending on size,
- $ASC_{ITC}$  is an individual tour constant; different depending on ratio of number of cars to number of workers, and
- $ASC_{JTC}$  is a joint tour constant; different depending on ratio of number of cars to number of workers

## NonMotorized

$$V_{Walk} = DIS(\beta_{DIS}) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o}) + ASC_{ITC} + ASC_{JTC}$$

$$V_{Bike} = DIS(\beta_{DIS}) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o}) + ASC_{ITC} + ASC_{JTC}$$

- $DIS$  is the distance traveled (miles);  $\beta_{DIS}$  is different depending on if  $DIS$  is above or below 1.5 miles for walk and 6 miles for bike,
- $ZTI$  is the zonal topography index at the destination,
- $ZDI_d$  is the zonal density index at the destination,
- $ZDI_o$  is the zonal density index at the origin,
- $ASC_{ITC}$  is an individual tour constant; different depending on ratio of number of cars to number of workers, and
- $ASC_{JTC}$  is a joint tour constant; different depending on ratio of number of cars to number of workers

## Transit

$$V_{WalkTransit} = IVTT(\beta_{IVTT}) + [OVTT_w(\beta_{OVTT_w}) + OVTT_t(\beta_{OVTT_t}) + ASC_{OVTT_d} + ASC_{OVTT_o}] \\ + T(\beta_T) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o}) + ASC_{T10} + ASC_{CBD} + ASC_{ITC} \\ + ASC_{JTC} + ASC_{TLH}$$

$$V_{DriveTransit} = IVTT(\beta_{IVTT}) + [OVTT_w(\beta_{OVTT_w}) + OVTT_t(\beta_{OVTT_t}) + OVTT_{dr}(\beta_{OVTT_{dr}})] \\ + T(\beta_T) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + DIS_{15}(\beta_{DIS_{15}}) + ASC_{T10} + ASC_{CBD} + ASC_{ITC} \\ + ASC_{JTC} + ASC_{TLH}$$

- $IVTT$  is the in-vehicle travel time (min);  $\beta_{IVTT}$  is different depending on which transit mode is selected,
- $OVTT_w$  is the initial wait time (min);  $\beta_{OVTT_w}$  is different depending on if  $OVTT_w$  is above or below 10 minutes,
- $OVTT_{dr}$  is the drive to transit time (min),
- $OVTT_t$  is the transfer wait time travel time (min),
- $ASC_{OVTT_d}$  is the destination walk time constant; different depending on short or long walk to destination,
- $ASC_{OVTT_o}$  is the origin walk time constant; different depending on short or long walk from origin,
- $T$  is the number of transfers,
- $ZTI$  is the zonal topography index at the destination,
- $ZDI_d$  is the zonal density index at the destination,
- $ZDI_o$  is the zonal density index at the origin,
- $DIS_{15}$  is the number of miles less than 15,
- $ASC_{T10}$  is a demographic variable for individuals under 10 years old,
- $ASC_{CBD}$  is a constant used if the destination is in CBD for area types 0 and 1,
- $ASC_{ITC}$  is an individual tour constant; different depending on ratio of number of cars to number of workers,

Tour Mode	Trip Mode												
	Drive Alone (Free)	Drive Alone (Pay)	Shared Ride 2 (Free)	Shared Ride 2 (Pay)	Shared Ride 3+ (Free)	Shared Ride 3+ (Pay)	Walk	Bike	Local	LRT/Ferry	Express	Heavy	Commuter
Drive Alone (Free)	X						X						
Drive Alone (Pay)	X	X					X						
Shared Ride 2 (Free)	X		X				X						
Shared Ride 2 (Pay)	X	X	X	X			X						
Shared Ride 3+ (Free)	X		X		X		X						
Shared Ride 3+ (Pay)	X	X	X	X	X	X	X						
Walk							X						
Bike							X	X					
Walk-Local			X		X		X		Walk				
Walk-LRT/Ferry			X		X		X		Walk	Walk			
Walk-Express			X		X		X		Walk	Walk	Walk		
Walk-Heavy			X		X		X		Walk	Walk	Walk	Walk	
Walk-Commuter			X		X		X		Walk	Walk	Walk	Walk	Walk
Drive-Local									Drive				
Drive-LRT/Ferry									Drive	Drive			
Drive-Express									Drive	Drive	Drive		
Drive-Heavy									Drive	Drive	Drive	Drive	
Drive-Commuter									Drive	Drive	Drive	Drive	Drive

Figure 3.2: Trip Mode Availability based on Tour Mode (MTC 2012).

- $ASC_{JTC}$  is a joint tour constant; different depending on ratio of number of cars to number of workers, and
- $ASC_{TLH}$  is the transit line-hail mode constant; different depending on which transit mode is selected.

### 3.5.3 ActivitySim's Trip Mode Choice Model

The trip mode choice model assigns a specific mode to each trip on a given tour. It is similar to the tour mode choice model, except only certain trips are allowed depending on the tour mode. Figure 3.2 displays a detailed explanation of which trip modes are allowed according to the tour mode that was selected in the tour mode choice.

The rules that define the correspondence between trips and tours are defined as the following: (MTC 2012)

1. Pay trip modes are only allowed under pay tour modes. For example, Shared Ride 2 (Pay) is only an available trip under the tour mode Shared Ride 2 (Pay).
2. The auto occupancy for the tour mode is determined by the maximum auto occupancy among all the auto trips within the tour. Therefore, the tour mode's auto occupancy is the trip with the largest auto occupancy within the tour.



3. Transit tours do allow auto-shared (carpooling) trips for particular legs. For example, if someone takes an auto-shared trip to work, the mode choice structures allows the individual to take transit back home.
4. The walk mode is an allowed mode for any given trip, no matter the tour.
5. “The availability of transit line-haul submodes on transit tours depends on the skimming and tour mode choice hierarchy” (MTC 2012). Albeit a low probability, free shared-ride trip modes are allowed in walk-transit tours. Paid shared-ride trip modes are not allowed, however, on transit tours because no evidence promotes this mode selection

The utility function equations used for the trip mode choice model are shown in Section 3.5.4. These parameters are very similar to the ones used in the tour mode choice model. In most cases, the in-vehicle times for coefficients in the trip mode choice are identical with the in-vehicle times in the tour mode choice. For coefficients of variables that apply in the trip mode choice to multiple legs, but only apply once in the tour mode choice, the trip mode choice in-vehicle time is half that of the tour mode choice in-vehicle time.

The trip is determined using a nested logit model structure, just like the tour mode choice model. There is one for each tour purpose as well. Notice how it is based on *tour* purpose and not each *trip* purpose.

### 3.5.4 ActivitySim’s Trip Mode Choice Utility Functions

The utility functions for the auto, nonmotorized, and transit modes are presented in this section. All  $\beta$  and  $ASC$  values are meant to represent parameter values. All other variables are dependent on other things, like time for example. (The utility equations are deduced from information from MTC (2012)).

#### **Auto**

$$V_{DriveAlone} = IVTT(\beta_{IVTT})$$

$$V_{Shared2} = IVTT(\beta_{IVTT}) + OVTT_w(\beta_{OVTT_w}) + ASC_{HHS} + ASC_{ITC}[TM_{S2}, TM_{S3P}, TM_T]$$

$$V_{Shared3P} = IVTT(\beta_{IVTT}) + OVTT_w(\beta_{OVTT_w}) + ASC_{HHS} + ASC_{ITC}[TM_{S3P}, TM_T]$$

- $IVTT$  is the in-vehicle travel time (min),
- $OVTT_w$  is the initial wait time (min);  $\beta_{OVTT_w}$  is different depending on if  $OVTT_w$  is above or below 10 minutes,
- $ASC_{HH}$  is a demographic variable for household size; only used if the household size is 1,
- $ASC_{ITC}$  is an individual tour constant,
- $TM_{S2}$  is a tour mode indicator for the Shared2 mode (1 if the tour mode is Shared2 and 0 if not),
- $TM_{S3P}$  is a tour mode indicator for the Shared3P mode (1 if the tour mode is Shared3P and 0 if not), and
- $TM_T$  is a tour mode indicator for a Transit mode (1 if the tour mode is Transit and 0 if not)

### NonMotorized

$$V_{Walk} = DIS(\beta_{DIS}) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o}) + ASC_{ITC}[TM_{DA}, TM_{S2}, TM_{S3P}, TM_B, TM_T]$$

$$V_{Bike} = DIS(\beta_{DIS}) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o})$$

- $DIS$  is the distance traveled (miles);  $\beta_{DIS}$  is different depending on if  $DIS$  is above or below 1.5 miles for walk and 6 miles for bike,
- $ZTI$  is the zonal topography index at the destination,
- $ZDI_d$  is the zonal density index at the destination,
- $ZDI_o$  is the zonal density index at the origin,
- $ASC_{ITC}$  is an individual tour constant,
- $TM_{DA}$  is a tour mode indicator for the DriveAlone mode (1 if the tour mode is DriveAlone and 0 if not),
- $TM_{S2}$  is a tour mode indicator for the Shared2 mode (1 if the tour mode is Shared2 and 0 if not),

- $TM_{S3P}$  is a tour mode indicator for the Shared3P mode(1 if the tour mode is Shared3P and 0 if not), and
- $TM_B$  is a tour mode indicator for the Bike mode (1 if the tour mode is Bike and 0 if not), and
- $TM_T$  is a tour mode indicator for a Transit mode(1 if the tour mode is Transit and 0 if not)

### Transit

$$V_{WalkTransit} = IVTT(\beta_{IVTT}) + [OVTT_w(\beta_{OVTT_w}) + OVTT_t(\beta_{OVTT_t}) + ASC_{OVTT_d} + ASC_{OVTT_o}] + T(\beta_T) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ZDI_o(\beta_{ZDI_o}) + ASC_{ITC}[TM_T]$$

$$V_{DriveTransit} = IVTT(\beta_{IVTT}) + [OVTT_w(\beta_{OVTT_w}) + OVTT_t(\beta_{OVTT_t}) + OVTT_{dr}(\beta_{OVTT_{dr}})] + T(\beta_T) + ZTI(\beta_{ZTI}) + ZDI_d(\beta_{ZDI_d}) + ASC_{ITC}[TM_T]$$

- $IVTT$  is the in-vehicle travel time (min);  $\beta_{IVTT}$  is different depending on which transit mode is selected,
- $OVTT_w$  is the initial wait time (min);  $\beta_{OVTT_w}$  is different depending on if  $OVTT_w$  is above or below 10 minutes,
- $OVTT_{dr}$  is the drive to transit time (min),
- $OVTT_t$  is the transfer wait time travel time (min),
- $ASC_{OVTT_d}$  is the destination walk time constant; different depending on short or long walk to destination,
- $ASC_{OVTT_o}$  is the origin walk time constant; different depending on short or long walk from origin,
- $T$  is the number of transfers,
- $ZTI$  is the zonal topography index at the destination,
- $ZDI_d$  is the zonal density index at the destination,
- $ZDI_o$  is the zonal density index at the origin,
- $ASC_{ITC}$  is an individual tour constant; different depending which transit mode is selected, and
- $TM_T$  is a tour mode indicator for a Transit mode(1 if the tour mode is Transit and 0 if not)

### 3.6 The Mode Choice Models used in Agent-based Models

Thus far, an extensive review of mode choice in activity-based models as well as the tour and trip mode choice models in Activitysim has been conducted. It is clear that an overarching tour mode value is essential to mode choice trip estimation in activity-based models. In ActivitySim, a multitude of person, path, and location variables are also needed to estimate accurate mode choice decisions.

As a comparison, we will now review the mechanics of mode choice in agent-based models / microsimulation tools. In Section 3.7 the mode choice model of the agent-based model MATSim is discussed. Then, in Section 3.8 the mode choice model of the agent-based model BEAM is discussed. The BEAM microsimulation tool is the agent-based mode used extensively this research.

### 3.7 The Mode Choice Model used in MATSim

Microsimulation is a modeling tool used to simulate individual vehicle movement across a network for the purpose of evaluating the traffic performance of the entire system. MATSim, as mentioned in Section @ref(lit3.7), is an example of a microsimulation tool used today and implements a unique mode choice model.

W Axhausen, Horni, and Nagel (2016) explains that mode choice in MATSim, along with time choice, destination choice, and route assignment, are chosen through an iterative cycle where every agent optimizes its daily activity pattern by picking the best option each cycle. After multiple iterations, the optimal choice is selected. Specifically, it uses scoring functions (the Charypar-Nagel utility function) to help determine the optimal choice for each iteration. Equation (3.1) shows the basic form of the Charypar-Nagel Function used in MATSim (W Axhausen, Horni, and Nagel 2016).

$$S_{plan} = \sum_{q=0}^{N-1} S_{act,q} + \sum_{q=0}^{N-1} S_{trav,mode(q)} \quad (3.1)$$

In the Charypar-Nagel utility function,  $S_{plan}$  represents the utility of a plan,  $S_{act,q}$  represents the sum of all the activity utilities, and  $S_{trav,mode(q)}$  represents the sum of all the travel “(dis)utilities” (W Axhausen, Horni, and Nagel 2016). In other words, the plan’s utility is computed from adding the activity utilities and subtracting the mode and travel utilities.

During each iteration, a new plan for the agent is calculated, resulting in different travel and mode values selected. After each iteration, the plan with the best overall “score” or utility is selected probabilistically among the previous and current iteration utility plans. As discussed in Section 3.4.7, MATSim continues to iteratively update mode choice until stochastic equilibrium is reached (Hasnine and Nurul Habib 2021). This action of determining the best modal alternative over a simulation process proves that the mode choice model is a mix of a simulation-based model (Section 3.4.5) and a dynamic tour-based model (Section 3.4.7).

MATSim’s mode choice program contrasts greatly from ActivitySim’s mode choice model. Various efforts have been proposed to create a more advanced mode choice model system in MATSim. These efforts will be further explained in Section 3.9. First, Section 3.8 explores the mode choice model of a different microsimulation tool named BEAM.

### **3.8 The Mode Choice Model used in BEAM**

Another microsimulation tool, similar to MATSim, is BEAM. BEAM stands for Behavior, Energy, Autonomy, and Mobility and is being developed at the Lawrence Berkely National Laboratory. According to El Zarwi, Vij, and Walker (2017), BEAM’s framework is designed to model and forecast the adoption and diffusion of new transportation services. Originally, BEAM implemented a Latent Class Choice Model (LCCM) as the mode choice structure to more accurately model early adopters, imitators, and non-adopters of the new transportation services. BEAM’s user guide, however, explains that the LCCM model is not currently being used, and instead something similar to a nested multinomial logit model is in effect. The observed utility function for the multinomial logit model includes parameters relating to travel time, cost, and a constant term.

BEAM’s multinomial logit model takes in certain parameters and inputs and uses them to

probabilistically estimate the likely mode choice for each user in their respective situation. Mathematically, the observable utility equation for the simple multinomial logit model in BEAM is shown in Equation (3.2).

$$V_j = ASC_j + \beta_{cost}(cost) + \beta_{time}(time) + \beta_{xfer}(numtransfers) \quad (3.2)$$

where

- $j$  is the modal alternative,
- $V_j$  is the observable portion of the utility equation,
- $ASC_j$  is the alternative specific constant, and
- $\beta_{cost}$ ,  $\beta_{time}$ , and  $\beta_{xfer}$  are generic coefficients for the cost, time, and number of transfer parameters.

All things considered, the observable utility equation for BEAM's multinomial logit model is pretty simple. There are only three  $\beta$  coefficients along with one alternative specific constant. Also, in BEAM, utility equations are specific to different mode options. For example, the utility for car, walk, bike, etc. will all be different. Yes Equation (3.2) shows the overarching utility equations, but the alternative specific constant differs between different modes.

A total of nine different modal alternatives are available in BEAM. Unlike ActivitySim, no tour-mode structure exists. BEAM's mode choice is designed on a trip-based system. Figure 3.3 highlights how BEAM's trip-based mode choice works and summarizes the difference between ActivitySim's mode choice and BEAM's mode choice.

ActivitySim provides mode choices off of the primary tour activity. For example, referring to in Figure 3.3, if a person wants to go to their work activity by train, ActivitySim knows that they can either walk or take transit to their gym activity. Contrastingly, BEAM bases mode choice on the first trip of the day. This means that BEAM will determine the optimal mode to go to the gym in, and then figure out how to get to work based on the first trip mode. In addition, ActivitySim knows that when a person leaves for lunch from their work activity, they will be returning back to their

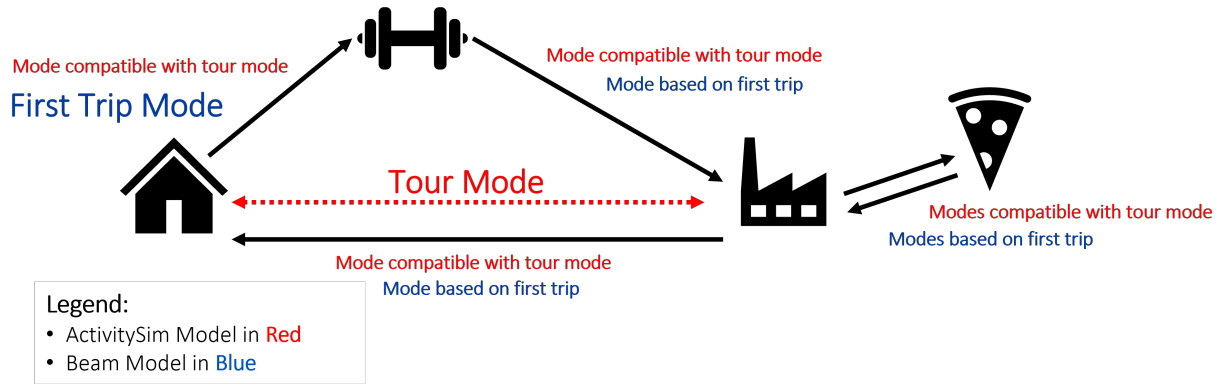


Figure 3.3: Mode choice between ActivitySim and BEAM.

work activity after lunch. This means that the person is able to leave their car at work and possibly walk or take transit to and from their lunch activity. In BEAM though, if someone goes to lunch, they will most likely take their car as to not abandon their vehicle for the rest of their day. It is clear that BEAM's mode choice is less representative of how people actually transport themselves during the day. ActivitySim provides a better representation of mode choice. As a result, it is important to improve the mode choice model in BEAM (See Section 3.9). In addition, in this research we explore the importance of creating consistency between the mode choice models between BEAM and ActivitySim (See Section 3.10).

### 3.9 Previous Attempts to Enhance Mode Choice in Agent-based Models

In recent years, suggestions have been made on how to improve MATSim's mode choice structure. Axhausen, Balmer, and Ciari (2008) proposed introducing discrete choice, in the form of multinomial logit models, on the subtour level of mode choice. Currently, MATSim's mode choice is performed at the tour level, meaning subtour mode choice is limited. This new introduction would allow variation in mode on all subtours. Similarly, Hörl, Balac, and Axhausen (2018) proposed updating upon a subtour mode choice model to eliminate heavily randomized processes and to increase the use of realistic subtour modal options.

In addition to improving the subtour mode choice strategy within MATSim, Hörl, Balac, and Axhausen (2018) also proposed introducing a chain-based modal selection model, similar to the

type of model discussed in Section 3.4.6. Hörl, Balac, and Axhausen (2018) describes this chain-based model with, “If one iterates over all possible combinations of modes on these trips, one arrives at a set of available mode chains  $C$  of size  $M^N$ , where  $M$  is the number of possible modes and  $N$  is the number of trips.” The resulting set of feasible chains  $C_f \in C$  could look as follows:

$$C_f = \{(car, car, car), (transit, transit, transit), (transit, walk, transit), \dots\} \quad (3.3)$$

Equation (3.3) provides a great example of all the chain-based modal alternatives that would exist in the choice set. In order to model the chain selection, Hörl, Balac, and Axhausen (2018) suggest incorporating “an out-of-the-box discrete mode choice model into MATSim”. They propose that this discrete choice model would involve two sampling approaches. The first approach involves calculating the total utility value for the chain-based modes and the other would be based on individual trip probabilities. In order to calculate the utility however, various variables like travel time and travel costs would be involved. The downside to this model, however, would be that analyzing the chain-based modes and calculating individual utilities would cause significant computational overhead compared to MATSim’s default mode choice model.

In the work done by Hörl, Balać, and Axhausen (2019), a discrete choice model is actually inserted into MATSim and tested using a case study of the city of Zurich. In this model, both a trip based and a tour based mode choice model are used. The tour based model is identical to the chain-based model described in the work done by Hörl, Balac, and Axhausen (2018). A multinomial logit model is used to probabilistically select between the utility values of each modal alternative. With the introduction of discrete choice and chain-based tour mode choice into MATSim, Hörl, Balać, and Axhausen (2019) states that there is better convergence behavior. Since mode choice is not done at random and no irrational modal decisions are being made, this newer model arrives at convergence with fewer iterations. However, since more explanatory variables are needed to estimate mode choice, extensive calibration is required in order to ensure the mode choice model



performs in an unbiased manner.

Hörl, Balać, and Axhausen (2019) found that the advanced trip-based model resulted in a more accurate mode share by time of day than the simple trip-based model. It was also determined that run time was higher with the advanced trip-based mode. Overall, the introduction of discrete choice and chain-based tour mode choice into MATSim could result in scenarios with quicker convergence, more accurate modal shares, and higher run times.

Since BEAM is a new agent-based model that is still being developed regularly, not many proposals have been made to improve its mode choice model. However, one example in the literature of a mode choice improvement to BEAM is done by Barth et al. (2020). Barth et al. (2020) proposed that BEAM implements a fundamental influencing factor (FIF) mode choice model instead of the original LCCM or the current MNL. The primary goal of the FIF model would be to calibrate the model against data on people's actual travel decisions while simultaneously allowing the introduction of any hypothetical mode. The FIF model would still implement an LCCM by grouping individuals into different classes based on a logit function. The logit utility function, however, would include seven parameters; this makes the influencing factors more than twice that of BEAM's current function. Overall, this new model would make BEAM more effective at estimating the usage of new transportation services, but the structure would be vastly different than many activity-based models.

The research we present is another idea of how to improve the mode choice model in the agent-based tool BEAM. This proposed improvement is discussed in Section 3.10 and revolves around the idea of creating consistency between activity-based models and microsimulation tools.

### **3.10 Mode Choice Consistency between ActivitySim and BEAM**

Oftentimes, the outputs of an activity-based model will serve as the inputs to a microsimulation tool. However, as discussed in Section 3.4 activity-based models can include a wide range of mode choice modeling. Usually the mode choice models within activity-based models rely on a tour-mode choice first, followed by trip-mode choice. Mode choice models within activity-based

models are usually more thorough than mode choice models in microsimulation tools. As discussed in Section 3.7 MATSim simply uses a scoring function to determine mode decisions; after each iteration a new option is explored. In Section 3.8, BEAM's mode choice model was explained as a simply multinomial logit model with only a few parameters. Overall, the mode choice structures within activity-based models and microsimulation tools differ between each other.

Currently, there exists a gap in the literature with understanding the effects of a consistent mode choice model between an activity-based model and a microsimulation tool. Since its possible for the outputs of an activity-based model to be used as the inputs of a microsimulation tool, it would make intuitive sense that the structure for estimating mode choice would be held constant. However, this is not the case; mode choice models differ between different travel forecasting tools.

We propose to make a consistent mode choice model between an activity-based model and a microsimulation tool. Specifically, we propose to align BEAM's current multinomial logit mode choice model with ActivitySim's mode choice model. This will involve calculating the utility of each modal alternative in BEAM with the same path, person, and location variables that ActivitySim uses. In other words, the utility function for each modal alternative will be the same between both models. In addition, we will add a few new mode options in BEAM to allow for the ability to carpool. This new model will include basing utility parameters off of tour purpose values. Lastly, it would be ideal to develop a tour mode choice structure within BEAM; however, a tour-based mode choice model will not be discussed within this project because of the complexity of adding this in the BEAM code framework.

We aim to better understand the effects of a consistent mode choice model between an activity-based model and a microsimulation tool. We will do this by upgrading the mode choice model in BEAM to be more closely aligned with that of ActivitySim. We will then explore the outputs and compare the effects of the new mode choice model and the original mode choice model. Chapter 4 explains the methodology of how we accomplished this task.

### **3.11 Summary**

## **4 METHODS**

### **4.1 Overview**

A systematic approach to determine the effects of a consistent mode choice model between an activity-based model and a microsimulation tool requires four two main steps. The first step involves creating a test scenario that runs within the activity-based model ActivitySim and the microsimulation tool BEAM. Section 4.2 explains this step including how the test scenario was created and how ActivitySim and BEAM were configured. The second step involves adjusting the internal code of BEAM to align the mode choice model with that of ActivitySim's. Section 4.3 explains the details behind the changes made to BEAM's default mode choice model to align more closely with that of ActivitySim's. The third step is to calibrate the mode choice utility values within BEAM. Section 4.4 explains the details behind the calibration of the mode choice parameters. The last step is to compare the modal distributions between different mode choice models, which is explained in Section 4.5.

### **4.2 Creating and Setting up the Test Scenario**

Designing a test scenario to use within ActivitySim and BEAM was essential to understanding and testing mode choice between models. This section explains the region that was used to model the test scenario. This section also explains how the input files needed to run both ActivitySim and BEAM were generated for the test scenario.

#### **4.2.1 Test Scenario Region**

The test scenario used in this research includes the approximate 2.5 million agents of the Salt Lake City, Utah, USA region. This region includes the Box Elder, Davis, Salt Lake, Utah, and Weber counties. Figure 4.1 shows the 2881 Traffic Analysis Zones (TAZs) along with the five counties that make up the region of study.

This extensive region was selected for the following reasons:

1. There is substantial data that exists for this region. Having sufficient data points for each of the TAZs was essential to creating the inputs to ActivitySim and BEAM. The majority of the data used for this research was gathered and shared by Wasatch Front Regional Council (WFRC), the Metropolitan Planning Organization (MPO) of Salt Lake county.
2. There currently exists a need to analyze this region with a microsimulation tool, such as BEAM. Specifically, the Utah Transit Authority (UTA) wishes to understand what specific sub regions in the Salk Lake area will best serve new on-demand transit vehicles. In some future research, this specific question will be answered. Along this this question, having a microsimulation model of a region will prove to be an effective method to answering multiple transportation service questions in the future. To date, only one previous analysis of this region has been conducted with the BEAM software (Lant 2021).
3. The size and population within the study area are comparable to other regions around the United States. This means that the work completed in this research may be useful to other cities besides Salt Lake City.
4. The project researchers live in this region, and therefore have real life knowledge of how transportation works in the region. This allows the researchers to compare the the results of the models with their own experiences, hopefully providing extra verification.

Overall, the Salt Lake City, Utah region proved to be an effective location for this mode choice analysis research. After selecting the region of study, the next steps involved seting up and running ActivitySim.

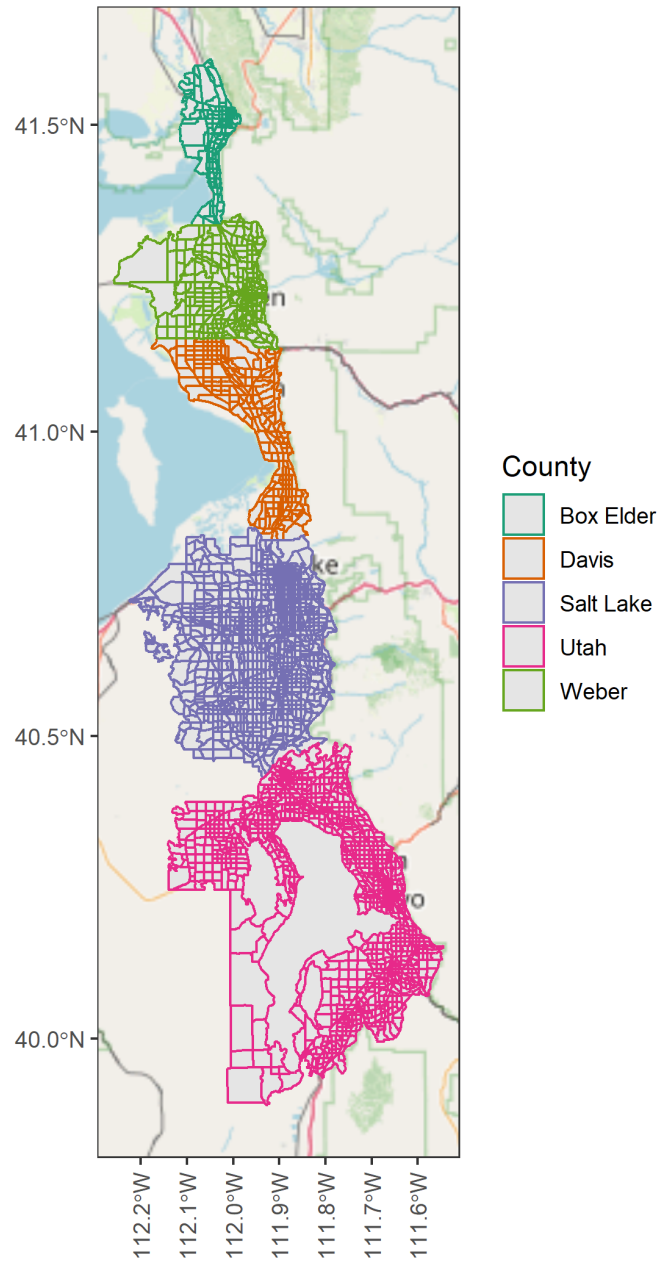


Figure 4.1: The TAZs of each county in the region of study.

#### 4.2.2 Setting up and Running ActivitySim

Setting up and running ActivitySim for the Salt Lake region was an arduous task. The entire process for setting up the activity-based model for this region is explained in the research done by Lant (2021). The research done by Lant (2021) was the basis to the work completed in this project.

However, a brief explanation of the setup of the activity-based model is explained in Section 4.2.2. The two main steps behind creating an activity-based mode for the Salt Lake region was creating corresponding input files and calibrating and validating the results.

##### *Creating Input Files*

ActivitySim requires the following three input files in order to run:

- A synthetic population of the agents within the study area.
- A zonal socioeconomic data file describing the characteristics of each zone.
- A set of skims that describe the cost and travel times of all modes between all zones.

The synthetic population is used to generate an entire population with specific individual details for each agent (like gender, age, income, vehicle ownership). Gathering individual information for all the agents in the region may be considered an invasion of privacy. So instead a synthetic population is used. The synthetic population is a generated population with specific individual attributes that add up to the regional characteristics as a whole. The synthetic population was generated using a software called PopulationSim (PopulationSim 2021). To run PopulationSim a “seed” table and a set of “targets” were used. The “seed” table represented information about a subset of the population and the set of “targets” represented demographic data for smaller areas within the region (Lant 2021). Together, PopulationSim was able to simulate individual attributes for the approximate 2.5 million agents in the Salt Lake region.

The zonal socioeconomic data file stores zonal characteristics regarding household information, worker information, and other activity type information. This file specifically gives information on what types and how many types of activities may be found within a certain zonal region. This file

was created using data from WFRC, Utah Automated Geographic Reference Center (AGRC 2021), and the synthetic population when necessary.

The third input to ActivitySim is detailed travel skims. Skims are large matrices showing travel times and costs between every set of zones within the area of study. Included in these skims were further details regarding differences in modes, distances, wait times, etc. (Lant 2021). The existing skims from WFRC were used along with some slight adjustments to fit the ActivitySim model.

### *Calibration and Validation*

The calibration and validation of ActivitySim was conducted by Lant (2021). The purpose of the calibration and validation was to ensure the outputs generated by ActivitySim matched target regional values. Specifically trip productions, trip distributions, and mode choices were tested to match the given target values from WFRC's four-step model.

To validate the trip productions, total trips per county were compared between ActivitySim and WFRC target values. These trips were close to identical, with only a variation of 1.0 percent. The trip distributions were compared as well, with about a 4.0 percent margin of error. Overall, Lant (2021) shows that the trip productions and distributions of the ActivitySim model were a good representation of travel behavior, requiring no calibration.

The mode choice validation was difficult as the mode choice categories between WFRC and ActivitySim are vastly different. ActivitySim considers five trip mode choice alternatives whereas the WFRC model considers 10 trip mode choice alternatives. However, ActivitySim does include 10 tour mode choice alternatives. When comparing mode choice distributions, both the trip and tour distributions from ActivitySim were compared with the trip distributions of WFRC. Lant (2021) performs an in depth comparison between these values and found that ActivitySim underrepresents non-motorized university trips and overrepresents non-motorized work trips. In addition ActivitySim underrepresents shared ride trips for all tour purposes and does not accurately represent commuter rail and local bus trips for all tour purposes. overall, the mode choice model did not match closely with the WFRC targets, and required further calibration.

Lant (2021) performed an in depth calibration across ActivitySim modes to minimize error with the WFRC target shares. The calibration process will not be explained in this research paper however. Upon calibration, trip productions and distributions remained relatively unchanged, and therefore needed no calibration. Mode choice distributions were closer to target values as well. Overall, after some calibration the results of ActivitySim produced sufficiently accurate information to move forward with the research.

### **4.2.3 Setting up BEAM**

After setting up, calibrating, and running the ActivitySim scenario for the Salt Lake region, the next step was setting up BEAM using the results of the ActivitySim model. BEAM requires a variety of input files, most of which are as follows:

- A person population and corresponding person attributes files.
- A household population with household attributes file and a vehicle fleet file.
- Vehicle types files explaining different person vehicle types, public transit vehicle types, and ride hailing vehicle information.
- Transportation services including a mapable street network and a file including General Transit Feed Specification (GTFS) transit information.
- Other minor input files.

The BEAM input files were directly creating using the `generate_beam_plans.py` python script ActivitySim add on. This script was developed by the developers at Lawrence Berkeley National Lab, and allows easy csv formatted input files to BEAM. The details behind these input files along with example tables of these files are discussed and shown in the following subsections.

#### *Population File*

The population file was created from the ActivitySim generated data. ActivitySim's main purpose is to generate daily activity patterns (DAPs) for each agent from the synthetic population. The



Table 4.1: A Subset of the Population Plans

ID	Mode	Type	End Time	Purpose	Household ID	LocaitonX	LocaitonY
50	NA	Home	25668.0	escort	6	412577.3	4606340
50	hov3	NA	NA	escort	6	NA	NA
50	NA	escort	28170.0	escort	6	413589.0	4593487
50	hov3	NA	NA	escort	6	NA	NA
50	NA	home	NA	escort	6	412577.3	4606340
51	NA	Home	36079.2	work	6	412092.6	4606274
51	car	NA	NA	work	6	NA	NA
51	NA	escort	38109.6	work	6	418892.5	4567448
51	car	NA	NA	work	6	NA	NA
51	NA	work	64998.0	work	6	424344.6	4537656
51	car	NA	NA	work	6	NA	NA
51	NA	escort	70578.0	work	6	409491.7	4552587
51	car	NA	NA	work	6	NA	NA
51	NA	home	NA	work	6	412092.6	4606274
52	NA	Home	67010.4	othdiscr	6	412437.5	4606445

DAPs generated by ActivitySim consists of all the activities that each agent within the population embark on during the day. Using the activity types, activity locations, activity durations, mode choices, departure and arrival times, and overall activity data, BEAM generated a population file along with a person attributes file.

The population file generated by the BEAM ActivitySim add-on code created a csv file describing all the events and choices each agent embarks on during the plan. For example, the plans file defines each agents home location, their activity locations, their travel and activity times, their mode choices for each trip, etc. A subsection of this file is shown in Table 4.1 to provide an example of some of its contents.

The person attributes file describes individual characteristics such as income, gender, age, vehicle ownership, etc. A subsection of this file is shown in Table 4.2 to provide an example of some of its contents.

Table 4.2: A Subset of the Person Attributes File

ID	Household ID	Age	Sex	VOT	TAZ	Income	Household Size	Auto Work Ratio
50	6	50	2	9.518956	9	103337.40	10	auto_sufficient
51	6	56	1	9.518956	9	103337.40	10	auto_sufficient
52	6	35	2	9.518956	9	103337.40	10	auto_sufficient
53	6	23	2	9.518956	9	103337.40	10	auto_sufficient
54	6	15	2	6.349144	9	103337.40	10	auto_sufficient
55	6	11	1	6.349144	9	103337.40	10	auto_sufficient
56	6	5	1	6.349144	9	103337.40	10	auto_sufficient
57	6	2	1	6.349144	9	103337.40	10	auto_sufficient
58	6	1	1	6.349144	9	103337.40	10	auto_sufficient
59	6	43	1	9.518956	9	103337.40	10	auto_sufficient
599	58	71	1	8.431196	5	52461.37	9	auto_sufficient
600	58	63	2	8.431196	5	52461.37	9	auto_sufficient
601	58	35	2	8.431196	5	52461.37	9	auto_sufficient
602	58	23	2	8.431196	5	52461.37	9	auto_sufficient
603	58	16	1	5.623608	5	52461.37	9	auto_sufficient

### *Household and Vehicle Files*

The household and vehicle files were generated using the same code to generate the population files. ActivitySim generates three specific files: a households, persons, and trips file. The persons and trips files were used to construct the population plans file. The households file, however, was used to generate both the household and vehicle BEAM input files.

Within the BEAM households file, information such as household income, vehicle ownership, and the IDs of the agents and vehicles that correspond to each household ID is provided. Unfortunately, ActivitySim does not generate household coordinate locations, but does provide the TAZ in which the house presides. Therefore, the `generate_beam_plans.py` code designed by the BEAM developers assigns a random coordinate within the TAZ to each house. This household coordinate is then used to simulate starting and ending locations for each agent within their population plans file. An example of a subset of the households file that is used as an input to BEAM is shown in Table 4.3.

The vehicles file is quite simple in that it only includes information on the vehicle ID, vehicle

Table 4.3: A Subset of the Households File

Household ID	TAZ	Income	Household Size	LocationX	LocationY	Auto Work Ratio
304379	1281	37362.029	8	419080.7	4498686	auto_sufficient
270209	1104	163050.420	4	426367.4	4506262	auto_sufficient
68811	555	17291.518	3	428645.4	4524432	auto_sufficient
83245	553	35756.388	2	426902.8	4523207	no_auto
373813	1516	54036.582	2	430560.8	4491969	auto_sufficient
574302	341	68375.208	4	418986.0	4559249	no_auto
253735	1197	192125.910	2	431933.2	4504105	auto_sufficient
53111	473	21239.420	2	420837.4	4542079	auto_sufficient
366644	1481	13778.640	2	429353.8	4494406	no_auto
101334	666	9777.942	2	422532.8	4514347	auto_sufficient

Table 4.4: A Subset of the Vehicles File

Vehicle ID	Vehicle Type	Household ID
1	CAR	6
2	CAR	6
3	CAR	6
4	CAR	6
5	CAR	30
6	CAR	30
7	CAR	30
8	CAR	30

type, and the household the vehicle belongs to. An example of a subset of the vehicles file that is used as an input to BEAM is shown in Table 4.4.

#### *Vehicle Types Files*

In addition to the household and vehicle files, two vehicle type files are also needed to run BEAM. The first describes attributes that relate to each vehicle type. For example, some attributes that are included in the vehicle types file is the fuel efficiency, seating capacity, vehicle length, etc. An example of a subset of the vehicle types file that is used as an input to BEAM is shown in Table 4.5.

A secondary vehicle types file exists that describes the nature of the ride hail fleet. One

Table 4.5: A Subset of the Vehicle Types Input File

ID	Seats	Standing Room	Length (m)	Fuel	Vehicle Type
BODY-TYPE-DEFAULT	0	0	0.5	Food	Body
BIKE-DEFAULT	2	0	1.5	gasoline	Bike
FAST-BIKE	2	0	1.5	gasoline	Bike
Car	4	0	4.5	gasoline	Car
CAR	4	0	4.5	gasoline	Car
Car-rh-only	4	0	4.5	gasoline	Car
WAV	6	0	4.5	gasoline	Car
CAV	4	0	4.5	gasoline	Car
BEV	4	0	4.5	electricity	Car
PHEV	4	0	4.5	electricity	Car
BUS-DEFAULT	37	20	12.1	diesel	MediumDutyPassenger
RAIL-DEFAULT	801	2052	129.6	diesel	MediumDutyPassenger
FERRY-DEFAULT	149	0	26.0	diesel	MediumDutyPassenger
SUBWAY-DEFAULT	336	864	129.6	electricity	MediumDutyPassenger
CABLE_CAR-DEFAULT	29	31	8.4	electricity	MediumDutyPassenger
TRAM-DEFAULT	58	19	14.0	electricity	MediumDutyPassenger

particular advantage to the microsimulation tool BEAM is its ability to model on-demand transit vehicles in a realistic manner. Part of creating realistic ride-hailing services is using an input file describing the nature of the fleet of ride-hail vehicles. An example of a subset of the ride-hail vehicle attributes file that is used as an input to BEAM is shown in Table 4.6. For this research project, the ride-hail file used was taken from an example BEAM scenario. In the future however, a ride hailing fleet file that accurately depicts the nature of ride hail vehicles in the Salt Lake region will be used. For this research it was not necessary to include, and so a default file was used instead.

### *Transportation Services*

BEAM requires two main files in order to provide transportation services to agents within the simulation. These two files are a GTFS file and a routeable transportation network.

GTFS data is a common format that houses information relating to transportation schedules and transportation locations. Many public agencies publish transportation data in the GTFS format. The GTFS data used as the input to BEAM was obtained from UTA's mobility data feed from 2019

Table 4.6: A Subset of the Vehicle Types Input File

ID	Vehicle Type	Starting Point X	Starting Point Y	Shift Times
rideHailVehicle-033000	Car	414692.0	4498566	{10:25200}
rideHailVehicle-012301	Car	418969.2	4495377	{10:25200}
rideHailVehicle-035201	Car	422864.0	4527397	{10:25200}
rideHailVehicle-031400	Car	430458.0	4508789	{10:25200}
rideHailVehicle-030800	Car	417063.9	4568028	{10:25200}
rideHailVehicle-016300	Car	415844.3	4544000	{10:25200}
rideHailVehicle-033100	Car	408380.1	4550982	{10:25200}
rideHailVehicle-061500	Car	436159.1	4474228	{10:25200}
rideHailVehicle-032601	Car	420785.5	4484536	{10:25200}
rideHailVehicle-015500	Car	444235.6	4452609	{10:25200}
rideHailVehicle-016400	Car	410876.3	4506167	{10:25200}
rideHailVehicle-061400	Car	414579.1	4485811	{10:25200}
rideHailVehicle-017601	Car	411338.4	4548806	{10:25200}
rideHailVehicle-020700	Car	446697.0	4437390	{10:25200}
rideHailVehicle-010500	Car	425560.0	4483818	{10:25200}

(UTA 2021). BEAM maps the GTFS data onto the highway network allowing agents to use both transportation and highway services interchangeably (Lant 2021).

A detailed routeable road network for the study area was obtained from WFRC’s travel demand model. After some manipulation, the network was cleaned of misconnections and filtered to include the correct attributes. Within the network, various attributes such as capacity, free flow speed, number of lanes, functional classification, direction, whether it is a one-way street, etc. were included. The network was constructed from a series of link and node combinations. The nodes represent intersections and activity locations whereas the links represent the roadways that connect each node. The links include the attributes of the network whereas the nodes include the coordinates. A detailed highway network of links and nodes allows BEAM to simulate traffic conditions in a more realistic manner.

Figure 4.2 displays the roadway network of the region of study. Figure 4.3 displays a close up of roadway network system within Salt Lake City, Utah to show significant detail.

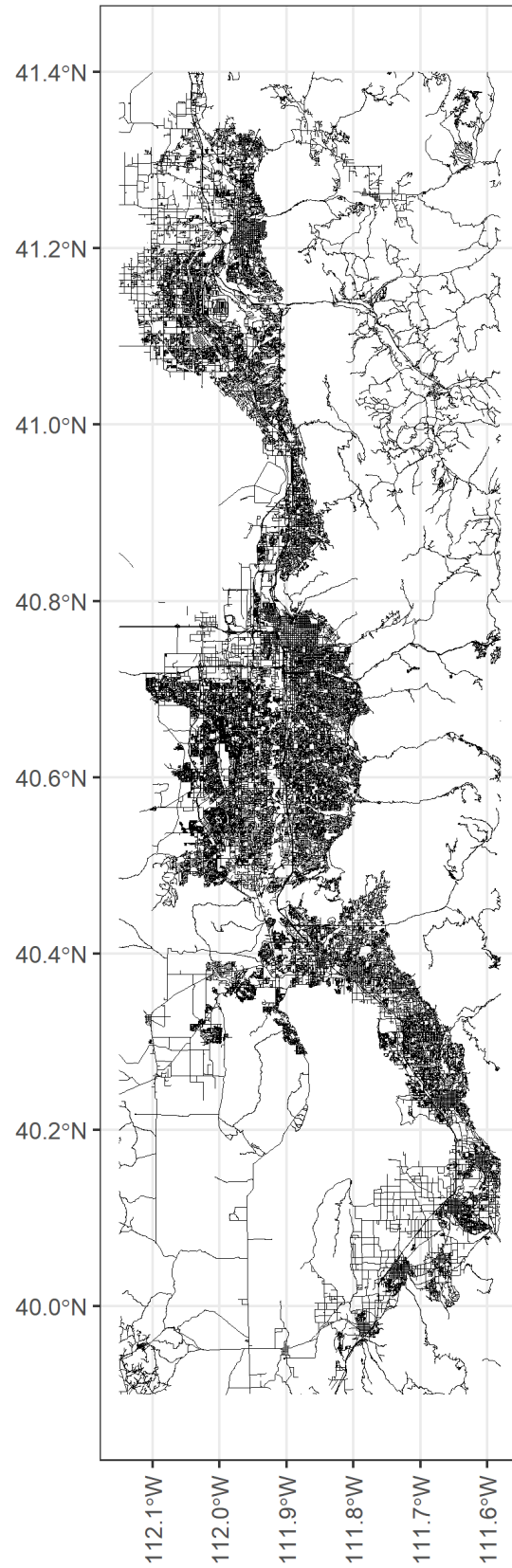


Figure 4.2: Transportation network for area of study.

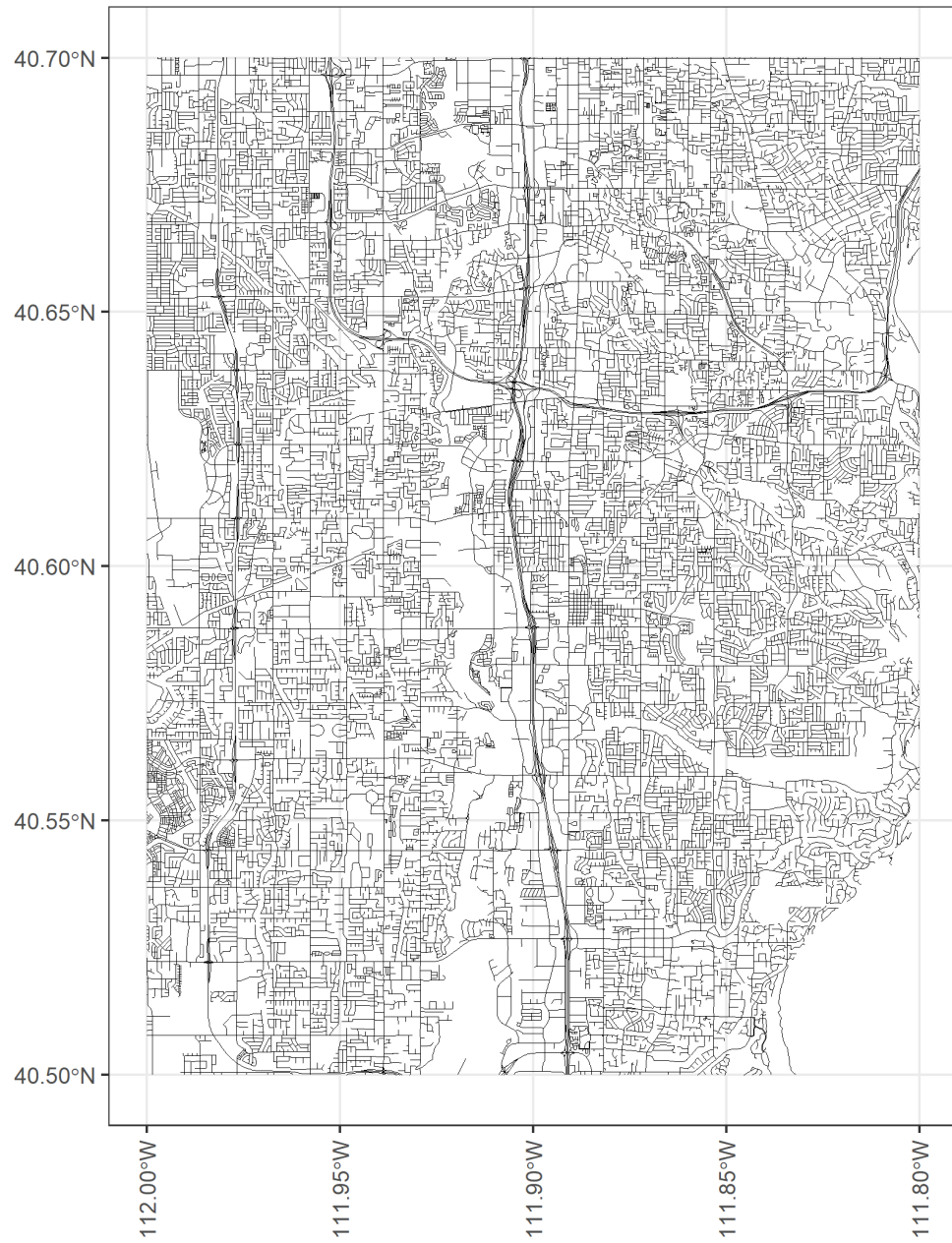


Figure 4.3: Close up of the transportation network of Salt Lake City.

### *Other Input Files*

A variety of additional input files can be included to run the BEAM simulation. Not all the files will be mentioned, but a few of them include a configuration file, a parking file, a toll prices file, TAZ specific files, and mode choice utility parameter files. Besides for the mode choice utility parameter files, details about these additional files will not be provided. The details behind the mode choice utility parameter file will be discussed in Section 4.3.

## **4.3 Improved Mode Choice in BEAM**

After creating the basic input files in order to run BEAM for the Salt Lake region, the next step was adjusting the mode choice model within BEAM itself. As discussed in Section 3.8 BEAM's default mode choice model is a simple multinomial logit model calculated with minimal travel attributes. In Section 3.5 it is discussed that ActivitySim's mode choice model is comprehensive.

Changing the mode choice model within BEAM to become more consistent with that of ActivitySim's required some major adjustments to BEAM's internal code. Specifically, in this research BEAM's code was changed in three significant ways. Section 4.3.1 describes how choices were changed to be made using activity purpose values, Section 4.3.2 describes how additional attributes were inserted to the utility equation, and Section 4.3.3 describes how new modal alternatives were added to BEAM. Then, in Section 4.3.4 algorithms explaining BEAM's improved mode choice model is presented.

### **4.3.1 Purpose-based Decisions**

The first major addition to BEAM's mode choice model to improve consistency with ActivitySim was to base mode choice decisions off of activity purposes. An activity purpose describes the primary activity of trips on a tour. Primary purposes are selected on a tour-based level. For this reason, the primary purpose is also referred to as the tour purpose. For example, if an agent were to go to the gym, then to work, and then return home, each trip on the work tour would have a primary



purpose of “work”. In this paper, tour purpose, primary purpose, activity purpose, and tour type are all used interchangeably to mean the same thing.

ActivitySim’s mode choice coefficients vary between ten different tour purposes. These tour purposes are as follows:

- Work
- University
- School
- Escort
- Shopping
- Eating Out
- Social
- Other Maintenance
- Other Discretionary
- At Work

To ensure the ability to distinguish between tour purpose values, it was necessary to adjust the population plans file to include a tour purpose trip-level attribute. Since ActivitySim bases decisions on activity purpose, it was simple to attach the attribute to the BEAM input plans file. As shown in Table 4.1 a purpose attribute exists. After assigning this attribute to the plans, the next step was adjusting the BEAM code to read in a new attribute. By adjusting BEAM’s CSV scenario reader script, along with other scripts relating to reading inputs, the primary purpose was able to be read into the BEAM software.

Once primary purpose values were present in the population plans, the next step was to design a new mode choice model within BEAM. In order to better understand the new code that was inserted into BEAM, three flow charts were constructed (See Figure 4.4, Figure 4.5, and Figure 4.6). In these flow charts inputs are displayed as folder icons, Java classes and methods are displayed as rectangles, and the result is displayed as an oval. These flow charts help display the code behind the BEAM’s mode choice models. Understanding BEAM’s default mode choice model and its

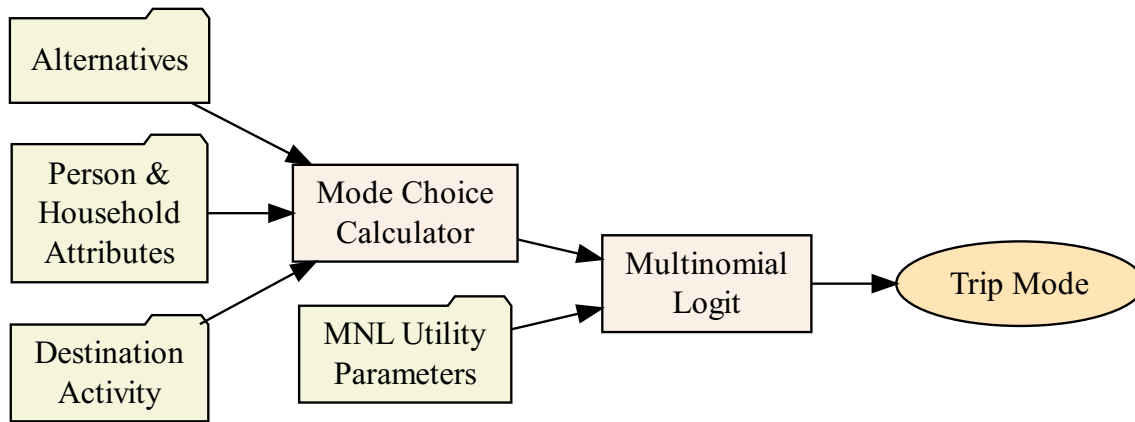


Figure 4.4: BEAM code flow chart behind the default multinomial logit choice model.

obsolete Latent Class Choice Model (LCCM) are useful toward understanding how the new Tour Purpose Choice Model (TPCM) was constructed in the BEAM code.

### *Multinomial Logit Choice Model*

Figure 4.4 shows the process behind BEAM’s default mode choice model. As described in Section 3.8, by default BEAM calculates mode choices using a multinomial logit function with only a few attribute variables. In Figure 4.4 the flow of inputs and BEAM functions are shown. As shown, the inputs to the Mode Choice Calculator are the alternatives, person and household attributes, and the destination activity data.

The alternatives input represents a set of detailed travel itineraries that the agent can choose from. Within the alternatives variable, travel statistics, travel paths, travel modes, travel times,

etc. are included. So as a result, the alternatives variable represents the modes in which the agent can choose from, while also including all the travel statistics needed to calculate the probability of selecting such mode. The next input is the person and household attributes variable. This variable includes all the information from the person attributes input file (See Table 4.2) and the household attributes file (See Table 4.3). Within the household attributes file some information from the vehicle inputs file is also included. The person and household attributes variable is also needed to calculate the probability of selecting such a mode. The last input is the destination activity data. This data simply includes addition information relating to the current trip and the destination activity. It is also useful to calculating the final mode choice.

The alternatives, person and household attributes, and destination activity are fed into the Mode Choice Calculator Java class. The purpose of this class is to organize the mode choice input data and organize it into a manner in which the choice model can process it efficiently. The Mode Choice Calculator class also houses some useful functions in the code used to calculate generalized travel time and the all day utility value.

The organized data from the Mode Choice Calculator along with the multinomial logit utility parameters are then sent to the multinomial logit choice model. The utility parameters in the default BEAM code are simply expressed in the configuration file. For each mode, an intercept, transit occupancy level, and transfer intercept/multiplier is specified. Then, within the multinomial logit choice model the utility value is calculated for each modal alternative. Those utility values are then sampled and one final mode is chosen!

### *Latent Class Choice Model*

BEAM has an obsolete LCCM within its code structure. When BEAM was first being developed, the LCCM was how modal decisions were made. Part of creating the new Tour Purpose Choice Model was first to recreate and reboot the LCCM and second to change its structure slightly to base decisions off of tour purpose. For this reason, the structure behind the LCCM is presented. Figure 4.5 shows the flow of code for the LCCM.

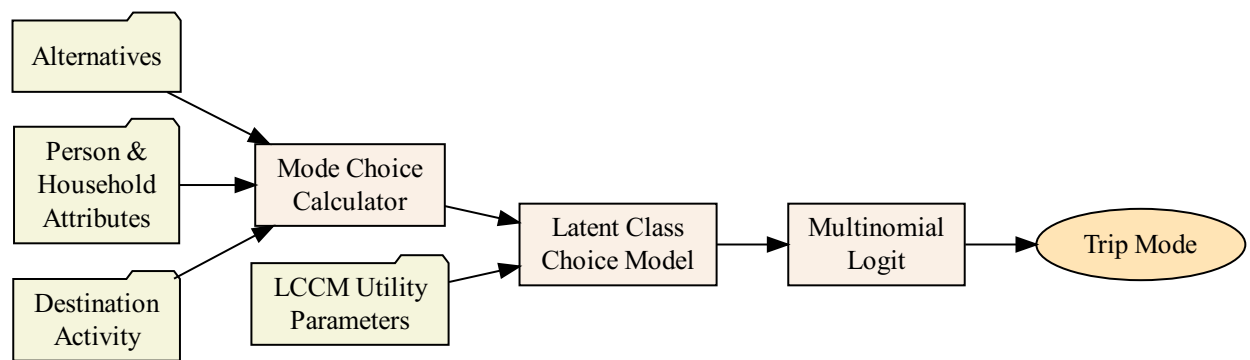


Figure 4.5: BEAM code flow chart behind the latent class choice model.

Just like with the multinomial logit choice model, the same alternatives, person and household attributes, and destination activity inputs are fed into the mode choice calculator. The mode choice calculator then acts in a similar way to organize the data for further analysis. However, an additional step is taken where the information is sent to a Latent Class Choice Model Java class before it is sent to the Multinomial Logit Java class. Within the Latent Class Choice Model Java class, a few things happen. First, the LCCM Utility Parameters data file is read into the code and parsed. After parsing the data, the utility parameters are attached to each alternative and stored.

The LCCM Utility Parameters file includes mode choice coefficients relating to number of cars, household size, income, cost, time, and also alternative specific constants. The interesting part to the LCCM is that it calculates modes based off of which *modality style* a particular agent is a part of. The *modality style* represents the travel behavior of the individual. The LCCM code divides individuals up into one of the six groups or classes based on their *modality style* travel behavior. The LCCM BEAM code provides two incredibly useful structures that are directly related to creating the TPCM. First, it provides a code structure that reads in a csv file storing specific and attribute related utility parameters. Second, it provides a code structure for storing these utility parameters into each alternative based on person/household attributes, travel statistics, and destination activity.

Interestingly enough, after the Latent Class Choice Model Java class reads in the utility information and stores it accordingly, it is sent to the same Multinomial Logit Java class. The overall utility value is calculated for each alternative and then are sampled down to a final mode choice!

### *Tour Purpose Choice Model*

As mentioned previously, the TPCM was constructed off of the LCCM within BEAM. This means the first step to creating this model was fixing the LCCM and making the LCCM run correctly. With a few minimal code adjustments, and a lot of debugging, the LCCM was brought back to life. The final flow of code for the TPCM is presented in Figure 4.6.

The flow chart shows a very similar choice process as the LCCM, except that an additional input is sent to the Mode Choice Calculator. Obviously, the tour purpose value is necessary for the

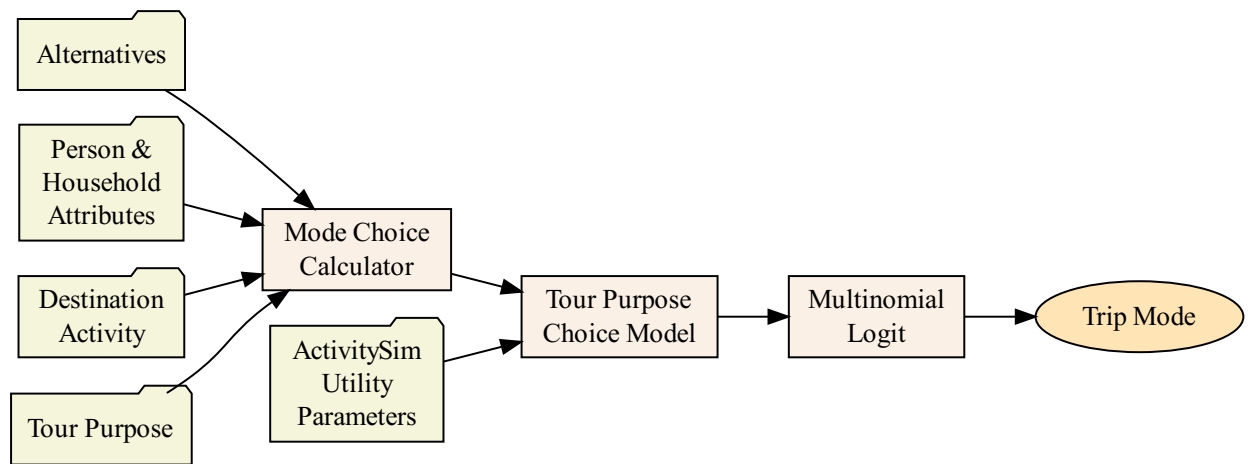


Figure 4.6: BEAM code flow chart behind the tour purpose choice model.

choice model and is therefore also sent to the Mode Choice Calculator. This input is only available, however, because it was attached to the population plans as explained previously.

In order to create the TPCM, two main steps were taken. First, a Tour Purpose Choice Model Java class was developed, mimicking how the code worked within the Latent Class Choice Model Java class. The TPCM was much more advanced than the LCCM, however, because instead of six classes to choose from there were 10 activity purposes to choose from. In addition, instead of only needing a few travel statistics to do the mode choice, the TPCM included 25 different path, person, and location type variables. The second step toward creating the TPCM was to develop an input csv file that included all the utility parameter values from ActivitySim. These additional attributes used in the utility equation are discussed in Section 4.3.2.

Overall, the Tour Purpose Choice Model Java class performed similarly to the Latent Class Choice Model Java class. It was able to parse the input utility parameters, attach those values to the alternatives, as well as calculate the travel statistics for all 25 travel variables. To do this, an additional class was built to calculate each of the travel variables. The final mode choice was determined in the same manner as well. These alternatives were sent to the Multinomial Logit Java class where their modal utilities were calculated and sampled. The sampling then determined a final mode choice for each individual.

#### **4.3.2 Additional Attributes in Utility Equation**

Equally important to developing the tour purpose choice model code structure within BEAM was to develop a consistent mode choice utility function. As explained in Section 3.5 ActivitySim bases its mode choice off of an abundance of path, person, and location type variables. Therefore, most of these path, person, and location type variables were inserted into the BEAM structure and used to calculate mode choice. The travel statistics for each variable were calculated within the TPCM code and the coefficient and intercept values were included using a csv input file.

The specific values for the coefficients used in the csv input file were taken from MTC (2012). Below, a list of the path, person, and location variables used to calculate the modal utility is shown.

- *Path Variables*: Number of Transfers, Vehicle Time, Egress Time, Wait Time, Origin to Transit Proximity, Destination to Transit Proximity, Walking Distance, Biking Distance, and Driving Distance
- *Person Variables*: Cost, Age, Household Size, and Vehicle to Worker Ratio
- *Location Variables*: Zonal Destination Index and Central Business District

In addition, Table 4.7 shows a subsection of this input file with only the cost variable coefficients provided. The purpose of showing this table is to help understand how extensive the choice model was. Within the table is clear to see that variation exists between tour purpose, modal alternative, and variable type. Only three modal alternatives are provided in the table.

The observable utility equation for the new tour purpose choice model can be divided up into three sub-equations. Equation (4.1) shows the Path type variable part of the TPCM utility equation. Equation (4.2) shows the Person type variable part of the TPCM utility equation. Equation (4.3) shows the Location type variable part of the the TPCM utility equation. The full TPCM utility equation, however, is the combination of path, person, and location type variables, as displayed in Equation (4.4).

$$V_j = \beta_{t_v}(t_v) + \beta_{t_w}(t_w) + \beta_{t_e}(t_e) + \beta_{tr_p}(tr_p) + \beta_{xfer}(xfer) + \beta_{w_{dis}}(w_{dis}) + \beta_{b_{dis}}(b_{dis}) + \beta_{d_{dis}}(d_{dis}) \quad (4.1)$$

where

- $j$  is the modal alternative,
- $t_v$  is the in vehicle travel time (mins),
- $t_w$  is the wait time (mins),
- $t_e$  is the egress time (mins),
- $tr_p$  is the proximity to transit (miles),
- $xfer$  is the number of transfers,
- $w_{dis}$  is the walk distance (miles),



Table 4.7: Cost coefficient values for some modes from the utility parameters csv file

Variable Type	Variable	Alternative	Units	Tour Purpose	Value
Person	cost	car	util/min	work	-0.026
Person	cost	car	util/min	univ	-0.044
Person	cost	car	util/min	school	-0.044
Person	cost	car	util/min	escort	-0.036
Person	cost	car	util/min	shopping	-0.036
Person	cost	car	util/min	eatout	-0.036
Person	cost	car	util/min	othmaint	-0.036
Person	cost	car	util/min	social	-0.036
Person	cost	car	util/min	othdiscr	-0.036
Person	cost	car	util/min	atwork	-0.038
Person	cost	ride_hail	util/min	work	-0.026
Person	cost	ride_hail	util/min	univ	-0.044
Person	cost	ride_hail	util/min	school	-0.044
Person	cost	ride_hail	util/min	escort	-0.036
Person	cost	ride_hail	util/min	shopping	-0.036
Person	cost	ride_hail	util/min	eatout	-0.036
Person	cost	ride_hail	util/min	othmaint	-0.036
Person	cost	ride_hail	util/min	social	-0.036
Person	cost	ride_hail	util/min	othdiscr	-0.036
Person	cost	ride_hail	util/min	atwork	-0.038
Person	cost	walk_transit	util/min	work	-0.026
Person	cost	walk_transit	util/min	univ	-0.044
Person	cost	walk_transit	util/min	school	-0.044
Person	cost	walk_transit	util/min	escort	-0.036
Person	cost	walk_transit	util/min	shopping	-0.036
Person	cost	walk_transit	util/min	eatout	-0.036
Person	cost	walk_transit	util/min	othmaint	-0.036
Person	cost	walk_transit	util/min	social	-0.036
Person	cost	walk_transit	util/min	othdiscr	-0.036
Person	cost	walk_transit	util/min	atwork	-0.038

- $b_{dis}$  is the bike distance (miles),
- $d_{dis}$  is the drive distance (miles),
- $\beta_{tr_p}$  differs between origin/destination and length, and
- $\beta_{w_{dis}}$ ,  $\beta_{b_{dis}}$ , and  $\beta_{d_{dis}}$  differ between lengths.
- *Note:* All  $\beta$  values differ between mode and tour purpose.

$$V_j = ASC_{auto} + \beta_c(c) + \beta_{ag}(ag) \quad (4.2)$$

where

- $j$  is the modal alternative,
- $ASC_{auto}$  is the alternative specific constant that differs between modal alternative and auto ownership dependency,
- $c$  is the cost, and
- $ag$  is the age grouping (if the person is between 0-10 or 16-19 years old).
- *Note:* All  $\beta$  values differ between mode and tour purpose.

$$V_j = \beta_{zdi}(zdi) + \beta_{cbd}(cbd) \quad (4.3)$$

where

- $j$  is the modal alternative
- $zdi$  is the zonal density index,
- $cbd$  is a classifier for zones labeled as central business district, and
- $\beta_{zdi}$  differs between origin/destination.
- *Note:* All  $\beta$  values differ between mode and tour purpose.

$$V_j = Eq : 2 + Eq : 3 + Eq : 4 \quad (4.4)$$

where

- $j$  is the modal alternative.

When comparing Equation (4.4) with BEAM's default equation in Equation (??), it is clear that multiple more travel variables are used to calculate the utility value. When comparing this Equation (4.4) with the ActivitySim equations in Section 3.5, most of the same variable are used. By using most of the path, person, and location type variables to calculate the mode choice, the TPCM calculates utility in a consistent manner with ActivitySim. In addition, all utility values differ by tour purpose, making it almost identical to ActivitySim's trip-based mode choice.

It is important to note that BEAM's TPCM was based heavily on ActivitySim's *trip-based* mode choice model. BEAM's internal code is built on a trip-based level. It would have been outside the scope of this project to adjust BEAM's structure to include a tour mode choice model. As a result, the mode choice consistency only exists on the trip-based mode choice level. Hopefully, in the future a consistent tour-based mode choice can be inserted into BEAM as well.

### **4.3.3 Carpool Alternatives**

### **4.3.4 Mode Choice Algorithms**

## **4.4 Data Validation and Calibration**

### **4.4.1 ActivitySim Path Variable Validation**

[graphs that compare asim, utah statewide, wfrc, and nchrp]

### **4.4.2 Mode Choice Calibration**

## **4.5 Path, Person, and Location Analysis**

## **4.6 Summary**

---

**Algorithm 1** Algorithm for Selecting Final Modal Alternative in BEAM

---

**Require:**

- 1:  $i$  : *origin*
  - 2:  $j$  : *destination*
  - 3:  $n$  : *agent*
  - 4:  $N$  : *population*
  - 5:  $t$  : *trip*
  - 6:  $P$  : *plan*
  - 7:  $\vec{A}$  : *attributes of agent*
  - 8:  $a$  : *attribute value*
  - 9:  $\vec{M}(i, j)$  : *Modal alternatives*
  - 10:  $m$  : *alternative*  $\in M(i, j)$
  - 11:  $\vec{U}(\vec{M}(i, j), \vec{A})$  : *Utilities for alternatives*
  - 12:  $u$  : *utility*  $\in \vec{U}(\vec{M}(i, j), \vec{A})$
  - 13:  $\vec{c}$  : *attribute coefficients*
  - 14:  $\mathbb{P}$  : *probability*
  - 15:  $Mode$  : *chosen mode for agent (n) on trip (t)*
  - 16:  $f(\vec{X})$  : This function takes a vector of modes and their probabilities of being chosen. With those probabilities it builds them into a cumulative distribution function, generates a random number and then drops the mode with the closest probability. This process continues until only one mode is left.
- 
- 17:  $\vec{M} \equiv \vec{M}(i, j)$
  - 18:  $\vec{U} \equiv \vec{U}(\vec{M}, \vec{A})$
  - 19: **for**  $n \in N$  **do**
  - 20:     **for**  $t \in P$  **do**
  - 21:         **procedure** DETERMINEFINALMODALALTERNATIVE( $\vec{M}, \vec{A}, \vec{c}$ )
  - 22:             **for**  $m \in \vec{M}$  **do**
  - 23:                  $u \leftarrow \sum_{a \in \vec{A}} a \times c_a$
  - 24:                  $\vec{U}_+ = [m, u]$
  - 25:             **end for**
  - 26:              $S \leftarrow \sum_{u \in \vec{U}} e^u$
  - 27:             **for**  $u \in \vec{U}$  **do**
  - 28:                  $\mathbb{P}(u) \leftarrow e^u / S$
  - 29:                  $\vec{B}_+ = [m, \mathbb{P}(u)]$
  - 30:             **end for**
  - 31:              $Mode \leftarrow f(\vec{B})$
  - 32:         **end procedure**
  - 33:     **end for**
  - 34: **end for**
-

## **5 RESULTS**

## **6 DISCUSSION**

## **7 CONCLUSIONS**

## 8 ABOUT

This is a *sample* book written in **Markdown**. You can use anything that Pandoc’s Markdown supports; for example, a math equation  $a^2 + b^2 = c^2$ .

### 8.1 Usage

Each **bookdown** chapter is an .Rmd file, and each .Rmd file can contain one (and only one) chapter. A chapter *must* start with a first-level heading: `# A good chapter`, and can contain one (and only one) first-level heading.

Use second-level and higher headings within chapters like: `## A short section` or `### An even shorter section`.

The `index.Rmd` file is required, and is also your first book chapter. It will be the homepage when you render the book.

### 8.2 Render book

You can render the HTML version of this example book without changing anything:

1. Find the **Build** pane in the RStudio IDE, and
2. Click on **Build Book**, then select your output format, or select “All formats” if you’d like to use multiple formats from the same book source files.

Or build the book from the R console:



```
bookdown::render_book()
```

To render this example to PDF as a `bookdown::pdf_book`, you'll need to install XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): <https://yihui.org/tinytex/>.

### 8.3 Preview book

As you work, you may start a local server to live preview this HTML book. This preview will update as you edit the book when you save individual .Rmd files. You can start the server in a work session by using the RStudio add-in “Preview book”, or from the R console:

```
bookdown::serve_book()
```

## 9 CROSS-REFERENCES

Cross-references make it easier for your readers to find and link to elements in your book.

### 9.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
  - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world = # Hello world {#hello-world}`.
  - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter 9.
  - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

### 9.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure 9.1.

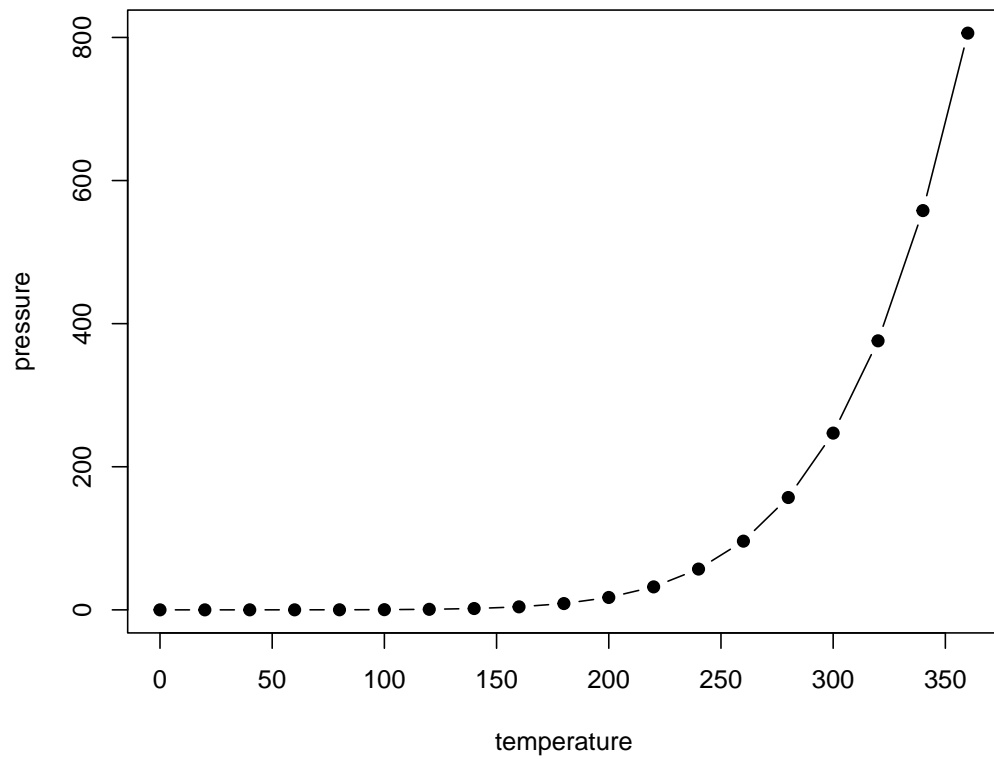


Figure 9.1: Here is a nice figure!

```
par(mar = c(4, 4, .1, .1))  
plot(pressure, type = 'b', pch = 19)
```

Don't miss Table 9.1.

```
knitr::kable(  
  head(pressure, 10), caption = 'Here is a nice table!',  
  booktabs = TRUE  
)
```

Table 9.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

## 10 PARTS

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

## 11 FOOTNOTES AND CITATIONS

### 11.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one <sup>1</sup>.

### 11.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package (Xie 2021) (check out the last code chunk in `index.Rmd` to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** (Xie 2015) (this citation was added manually in an external file `book.bib`). Note that the `.bib` files need to be listed in the `index.Rmd` with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: <https://rstudio.github.io/visual-markdown-editing/#/citations>

---

<sup>1</sup>This is a footnote.

## 12 BLOCKS

### 12.1 Equations

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (12.1)$$

You may refer to using `\@ref{eq:binom}`, like see Equation (12.1).

### 12.2 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: <https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html>

## 13 SHARING YOUR BOOK

### 13.1 Publishing

HTML books can be published online, see: <https://bookdown.org/yihui/bookdown/publishing.html>

### 13.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

### 13.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book- all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's



source file.

Read more about the features of this output format here:

<https://pkgs.rstudio.com/bookdown/reference/gitbook.html>

Or use:

```
?bookdown::gitbook
```

## REFERENCES

- “ActivitySim: An Advanced Activity-Based Travel Demand Model Built by and for Users.” 2021. RSG. <https://rsginc.com/activitysim-white-paper/#advanced-activity-based-travel-demand-model>.
- AGRC. 2021. *Automated Geographic Reference Center*. <https://gis.utah.gov/>.
- Arentze, Theo, and Harry Timmermans. 2000. *Albatross: A Learning Based Transportation Oriented Simulation System*. Citeseer.
- Axhausen, K, M Balmer, and F Ciari. 2008. “A New Mode Choice Model for a Multi-Agent Transport Simulation.” In *8th Swiss Transport Research Conference*. *ETH, Eidgenössische Technische Hochschule*.
- Barth, Matthew, Peng Hao, Guoyuan Wu, Shams Tanvir, Chao Wang, Jeff Gonder, Jacob Holden, Andrew Devall, and Bingrong Sun. 2020. “Evaluating Energy Efficiency Opportunities from Connected and Automated Vehicle Deployments Coupled with Shared Mobility in California.” University of California-Riverside.
- Ben-Akiva, Moshe, Daniel Mcfadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, et al. 2002. “Hybrid Choice Models: Progress and Challenges.” *Marketing Letters* 13 (3): 163–75. <https://doi.org/10.1023/a:1020254301302>.
- Ben-Akiva, Moshe, Joan Walker, Adriana T. Bernardino, Dinesh A. Gopinath, Taka Morikawa, and Amalia Polydoropoulou. 2002. “Integration of Choice and Latent Variable Models.” In *Perpetual Motion*, 431–70. <https://doi.org/10.1016/b978-008044044-6/50022-x>.
- Bhat, Chandra R. 1995. “A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice.” *Transportation Research Part B: Methodological* 29 (6): 471–83.
- Bhat, Chandra R, and Frank S Koppelman. 1999. “Activity-Based Modeling of Travel Demand.” In *Handbook of Transportation Science*, 35–61. Springer.
- Bowman, J. L, and M. E Ben-Akiva. 2001. “Activity-Based Disaggregate Travel Demand Model System with Activity Schedules.” *Transportation Research Part A: Policy and Practice* 35 (1): 1–28. [https://doi.org/10.1016/s0965-8564\(99\)00043-9](https://doi.org/10.1016/s0965-8564(99)00043-9).

- Bowman, John L., Mark A. Bradley, and John Gibbs. 2006. "The Sacramento Activity-Based Travel Demand Model: Estimation and Validation Results." *In European Transport Conference*.
- Bowman, John L, Mark Bradley, Yoram Shiftan, T Keith Lawton, and Moshe Ben-Akiva. 1999. "Demonstration of an Activity-Based Model for Portland." *In World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research* World Conference on Transport Research Society. Volume 3.
- El Zarwi, Feras, Akshay Vij, and Joan L Walker. 2017. "A Discrete Choice Framework for Modeling and Forecasting the Adoption and Diffusion of New Transportation Services." *Transportation Research Part C: Emerging Technologies* 79: 207–23.
- Eluru, Naveen, Abdul Pinjari, Ram Pendyala, and Chandra Bhat. 2010. "An Econometric Multi-Dimensional Choice Model of Activity-Travel Behavior." *Transportation Letters* 2 (4): 217–30. <https://doi.org/10.3328/tl.2010.02.04.217-230>.
- Gali, Emmanuel, Stephan Eidenbenz, Sue Mniszewski, Leticia Cuellar, and Christof Teuscher. 2008. "ActivitySim: Large-Scale Agent Based Activity Generation for Infrastructure Simulation," January. <https://www.osti.gov/biblio/960770>.
- Goulias, Konstadinos G., Adam Wilkinson Davis, and Chandra R. Bhat. 2020. "Chapter 5 - Consumer Choice Modeling: The Promises and the Cautions." *In Mapping the Travel Behavior Genome*, 63–80. Elsevier.
- Hasnine, Md Sami, and Khandker Nurul Habib. 2021. "Tour-Based Mode Choice Modelling as the Core of an Activity-Based Travel Demand Modelling Framework: A Review of State-of-the-Art." *Transport Reviews* 41 (1): 5–26. <https://doi.org/10.1080/01441647.2020.1780648>.
- Hörl, Sebastian, Milos Balac, and Kay W Axhausen. 2018. "A First Look at Bridging Discrete Choice Modeling and Agent-Based Microsimulation in MATSim." *Procedia Computer Science* 130: 900–907.
- Hörl, Sebastian, Miloš Balać, and Kay W Axhausen. 2019. "Pairing Discrete Mode Choice Models and Agent-Based Transport Simulation with MATSim." *In 2019 TRB Annual Meeting Online*, 19–02409. Transportation Research Board.

- Lant, Nathan John. 2021. “Estimation and Simulation of Daily Activity Patterns for Individuals Using Wheelchairs.” PhD thesis, Brigham Young University.
- Manski, Charles F. 2001. “Daniel McFadden and the Econometric Analysis of Discrete Choice.” *The Scandinavian Journal of Economics* 103 (2): 217–29.
- McFadden, Daniel et al. 1973. “Conditional Logit Analysis of Qualitative Choice Behavior.”
- McFadden, Daniel. 1986. “The Choice Theory Approach to Market Research.” *Marketing Science* 5 (4): 275–97.
- McNally, Michael G. 2000. “The Four Step Model.” In *Handbook of Transport Modelling*. Emerald Group Publishing Limited.
- Miller, Eric J., Matthew J. Roorda, and Juan Antonio Carrasco. 2005. “A Tour-Based Model of Travel Mode Choice.” *Transportation* 32 (4): 399–422. <https://doi.org/10.1007/s11116-004-7962-3>.
- MTC. 2012. *Travel Model Development: Calibration and Validation Technical Report*. Metropolitan Transportation Commission with Parsons Brinckerhoff, Inc.
- Nurul Habib, Khandker. 2017. “A Comprehensive Utility-Based System of Activity-Travel Scheduling Options Modelling (Custom) for Worker’s Daily Activity Scheduling Processes.” *Transportmetrica A: Transport Science* 14 (4): 292–315. <https://doi.org/10.1080/23249935.2017.1385656>.
- Ortuzar, J.de D., and L. G. Willumsen. 1994. Wiley.
- Paulssen, Marcel, Dirk Temme, Akshay Vij, and Joan L. Walker. 2014. “Values, Attitudes and Travel Behavior: A Hierarchical Latent Variable Mixed Logit Model of Travel Mode Choice.” *Transportation* 41 (4): 873–88. <https://doi.org/10.1007/s11116-013-9504-3>.
- Pendyala, Ram M., Ryuichi Kitamura, Akira Kikuchi, Toshiyuki Yamamoto, and Satoshi Fujii. 2005. “Florida Activity Mobility Simulator.” *Transportation Research Record: Journal of the Transportation Research Board* 1921 (1): 123–30. <https://doi.org/10.1177/0361198105192100114>.
- Pendyala, Ram M, Ryuichi Kitamura, Cynthia Chen, and Eric I Pas. 1997. “An Activity-Based Microsimulation Analysis of Transportation Control Measures.” *Transport Policy* 4 (3): 183–92.

- Pinjari, Abdul Rawoof, Ram M. Pendyala, Chandra R. Bhat, and Paul A. Waddell. 2007. “Modeling Residential Sorting Effects to Understand the Impact of the Built Environment on Commute Mode Choice.” *Transportation* 34 (5): 557–73. <https://doi.org/10.1007/s11116-007-9127-7>.
- PopulationSim. 2021. *An Open Platform for Population Synthesis* (version Version: 0.5). <https://activitysim.github.io/populationsim/>.
- Saleem, Mohammad, Oskar Blom Västberg, and Anders Karlström. 2018. “An Activity Based Demand Model for Large Scale Simulations.” *Procedia Computer Science* 130: 920–25.
- UTA. 2021. “OpenMobilityData - GTFS.” Utah Transit Authority. <https://transitfeeds.com/p/utah-transportation-authority/59>.
- Vij, Akshay, André Carrel, and Joan L. Walker. 2013. “Incorporating the Influence of Latent Modal Preferences on Travel Mode Choice Behavior.” *Transportation Research Part A: Policy and Practice* 54: 164–78. <https://doi.org/10.1016/j.tra.2013.07.008>.
- Vovsha, P., J. E. Hicks, G. Vyas, V. Livshits, K. Jeon, R. Anderson, and G. Giaimo. 2017. “Combinatorial Tour Mode Choice.” In *Presented at He 96th Annual Meeting of Transportation Research Board*.
- W Axhausen, Kay, Andreas Horni, and Kai Nagel. 2016. *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press.
- Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <http://yihui.org/knitr/>.
- . 2021. *Bookdown: Authoring Books and Technical Documents with r Markdown*. <https://CRAN.R-project.org/package=bookdown>.