A Comparative Illustration of Trip- and Activity-Based Modeling Techniques

Hayden Atchley Kamryn Mansfield Gregory S. Macfarlane

1 January 1900

This is where an abstract goes.

Table of contents

List of Figures

List of Tables

List of Acronyms

ABM activity-based model
ACS American Community Survey
BRT bus rapid transit
CRT commuter rail transit
DAP daily activity pattern
PUMA Public Use Microdata Area
PUMS Public Use Microdata Sample
TAZ transportation analysis zone
TLFD trip length frequency distribution
UDOT Utah Department of Transportation
WFH work-from-home
WFRC Wasatch Front Regional Council

Introduction

Activity-based models (ABMs) have been championed by researchers and many practitioners as being theoretically superior to the trip-based models historically used in transportation planning efforts since the 1950s.

Despite the theoretical benefits, many agencies have delayed or declined to transition to an ABM citing additional data requirements, staff training, computational resources, and related concerns. There is also not a consensus that ABMs result in better infrastructure volume forecasts, the primary—or even sole—purpose of many regions' travel demand analysis efforts.

In this research, we investigate the quality and characteristics of travel analyses enabled by an ABM. We do this by applying an array of infrastructure and behavior scenarios to both an ABM and a trip-based model representing the Wasatch Front (Salt Lake City) region of Utah, USA. The results will compare the implications of each model and help inform agencies that are considering adopting an ABM.

The document proceeds in a typical fashion: Chapter 1 provides a discussion of the differences between trip-based models and ABMs, alongside a discussion of previous studies examining the theoretical and analytical benefits of each framework. Chapter 2 first describes the model frameworks used in this research, namely the regional trip-based model and an activity-based model constructed to support research activities in the region; this section also describes three scenarios designed to test the usefulness and applicability of the different model frameworks. Chapters 3–5 describe the findings from each scenario, alongside a discussion of their limitations and implications. Chapter 6 provides a series of recommendations and opportunities for future research.

1 Literature Review

Travel demand modeling in the modern sense has its origins in the 1950's, with the Chicago Area Transportation Study (Chicago Area Transportation Study 1959) being one of the first urban planning studies to use the now-ubiquitous "four-step" modeling framework (McNally 2007). Up to this point, most urban transportation planning used existing demand or uniform-growth travel forecasts to model travel demand, but the Chicago Study used a combination of trip generation, trip distribution, modal split, and network assignment models to more accurately represent travel behavior (Weiner 1997). Since then, there have been numerous studies iterating on the "four-step" (more appropriately termed "trip-based") framework, and trip-based models are now the primary tool used in forecasting travel demand across the United States (Park et al. 2020).

These trip-based models are not without problems, however. Rasouli and Timmermans (2014) give several shortcomings of trip-based models. First, they use several sub-models that are (implicitly or explicitly) assumed independent, and this can result in a lack of consistency or integrity between sub-models; for example, the assumed value of time in the mode choice model might be radically different than the assumed value of time in the tolling assignment model. Second, they are strongly aggregated in nature, which can cause significant aggregation bias with high and low values excluded. Finally, they lack "behavioral realism"—that is, they do not have a concept of individuals making decisions, which is what travel behavior actually is.

Jones (1979) proposed an alternative to the trip-based paradigm, namely an "activity-based" framework that models daily activity patterns at an individual rather than aggregate level. An activity-based model (ABM) places the focus on "activities" rather than "trips" as the basic unit of analysis, and predicts a sequence of activities for each individual and household, with information such as activity location, start time, and duration, using a high level of temporal and spatial granularity. "Trips" are then journeys from one activity to the next (Pinjari and Bhat 2011). By adopting this activity-centric framework, ABMs provide a more consistent and comprehensive representation of travel behavior. They take into account complex dependencies and interactions within the model as a whole and at an individual level. ABMs acknowledge that travel choices are not made in isolation, but rather influenced by the preceding activities. This means that e.g. if an individual takes transit to work, they will not be able to drive home. ABMs therefore present a more conceptually accurate model of actual travel behavior than traditional trip-based models.

Despite these advantages, many agencies have yet to adopt ABMs, and instead continue to use trip-based models (Miller 2023). While ABMs are superior in certain aspects, they also have disadvantages, such as requiring more detailed input data and greater computational resources. It is also not always clear if ABMs provide substantially better forecasts than their trip-based counterparts, nor if this tradeoff is worth it for every agency. This literature review presents an overview of both modeling frameworks, and discusses the advantages and disadvantages of using an ABM.

1.1 Overview of Model Types

Trip-based models are often referred to as "four-step" models due to their four fundamental sub-models: trip generation, trip distribution, mode choice, and network assignment (National Academies 2012, 28). They can be more complicated than this, possibly including integration with a land use forecast, iteration between mode and destination choice, etc., but the "four steps" are the central component of any of these models (McNally 2007).

In a typical trip-based model, travel demand is predicted based on aggregate population data, often delineated by each transportation analysis zone (TAZ). Each sub-model relies on this aggregate data; for example, the modal split sub-model will often use average TAZ income as an input (National Academies 2012, 14). The aggregate data is also often "disaggregated" and segmented along variables such as household size and vehicle ownership. Regardless of the segmentation variables used in the first three model steps, the resulting trip matrices by mode and time of day are then assigned to a transportation network (there are various algorithms for doing so, see Ortúzar and Willumsen (2011)).

Activity-based models differ significantly from this approach. Rather than using aggregate data, ABMs use data representing an actual or synthetic population, with individual person and household data (Vovsha, Bradley, and Bowman 2005). These models use an activity or tour scheduler to assign a daily activity pattern (DAP) of zero or more tours to each individual (*n.b.* a tour is a series of trips that begin and end at home). These DAPs are restricted temporally, spatially, and modally; i.e., each person has a logical and followable sequence of trips and activities (Bowman 1998). A "drive alone" trip from work to lunch, for example, cannot be made if transit was taken to work. ABMs output a list of tours and trips by person, time, location, and type, and these can then be assigned to a transportation network in a similar manner as in a trip-based model.

1.2 Comparison of Modeling Frameworks

In discussing the differences between ABMs and trip-based models, there are really two comparisons that need to be made: how the population data is structured, and how travel is organized. Trip-based models generally use aggregate population data while ABMs use a synthetic population, and trip-based models organize travel into trips while ABMs organize travel into activities and tours. The following sections will explain these aspects of travel demand modeling and discuss the claimed advantages and disadvantages of each model type.

1.2.1 Population Data

The aggregate population data used in trip-based models can vary in origin and level of detail, but the basic concept is the same: the study area is organized into generally small zones, and certain demographic and socioeconomic data is known or obtained for each zone (National Academies 2012, 14). This includes data such as number of households, average household income, population, number of workers, etc. Based on this information, the zone can be segmented along arbitrary variables. For example, since households with more workers and more vehicles tend to make more work trips, it is useful to estimate the distribution of households in each zone along these two variables. Then, average trip production rates are determined for each household category (e.g. for households with each number of workers by vehicles), and the total number of trips produced in a zone is calculated based on these rates (National Academies 2012, 37).

This approach is relatively straightforward: the required input data is usually easy to obtain, the trip generation models are often simple, and it is computationally inexpensive (National Academies 2012). However, the types of analyses possible are limited by the initial segmentation of the aggregate population data. An analysis based on parents'/adults' highest received education, for example, would require determining the number of households in each TAZ with each possible combination of education level. This can theoretically be done, but more detailed and varied analyses would require more levels of segmentation, greatly increasing the number of classifications needed. Aggregation at any point precludes that segmentation from use in subsequent model steps as well as in any post-hoc analysis. Since these segmentations need to be carried through each model step, trip rates, mode choice equations, etc. need to be estimated for every classification, and while relevant real-world data may exist, sample sizes approach zero very quickly, and so the estimates have little statistical value (Moeckel et al. 2020; National Academies 2012).

This becomes a particular issue in equity analysis because it is perhaps impossible to determine equitable distribution of "winners" and "losers" of a potential policy without using demographic variables in the trip generation and destination and mode choice steps (Bills and Walker 2017). Though many studies have shown that trip production and mode choice behavior differ by ethnic group even after controlling for income (**cite?**), including such variables in trip-based models is problematic. Does coding such a variable in a mode choice model represent discrimination? Or does doing so assert that present differences resulting from unequal opportunity will persist into the future planning years? Regardless the reasons for their exclusion, these variables consequently cannot be used in a post-hoc analysis of a transportation policy because the trip matrices do not contain the adequate segmentation.

An alternative approach to population data is to use a full synthetic population. A synthetic population takes demographic and socioeconomic data at various levels of detail to create a "population" with generally the same distribution as the study area (National Academies 2012, 93). The goal is to have a population that is functionally similar to the actual population, but without the privacy concerns of using real individual household data. Castiglione et al. (2006) argue that the major advantage with this approach is that the demographic and socioeconomic data is known at the person and household level, rather than the zone level, and this data remains available throughout the modeling process. This allows, for example, an equity analysis to determine the "winners" and "losers" of a proposed development without needing to encode demographic variables into each step of the model.

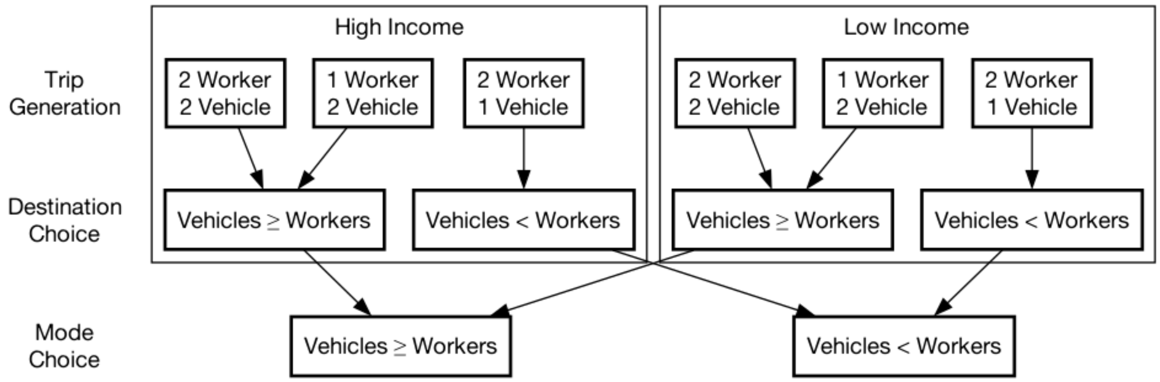
Bills and Walker (2017) used the 2000 Bay Area Travel Survey to create a synthetic population and compare the effects that certain scenarios had on high income and low income populations. With a 20% reduction in travel cost, they found that high income workers benefited more than low income workers. They did similar comparisons for scenarios involving reduced travel times for different mode choices and saw the effects each scenario had on the high and low income workers. These types of analysis, which are not possible with aggregate population data (**this is actually possible; a lot of trip-based models segment high/low income**), can be very valuable in transportation planning and policy making.

It is important to note that while many connect them only with ABMs, synthetic populations can be used in running trip-based models as well. Trip-based models using a synthetic population—often called trip-based microsimulation models—do exist (see Walker (2005) and Moeckel et al. (2020)), but these are relatively rare.

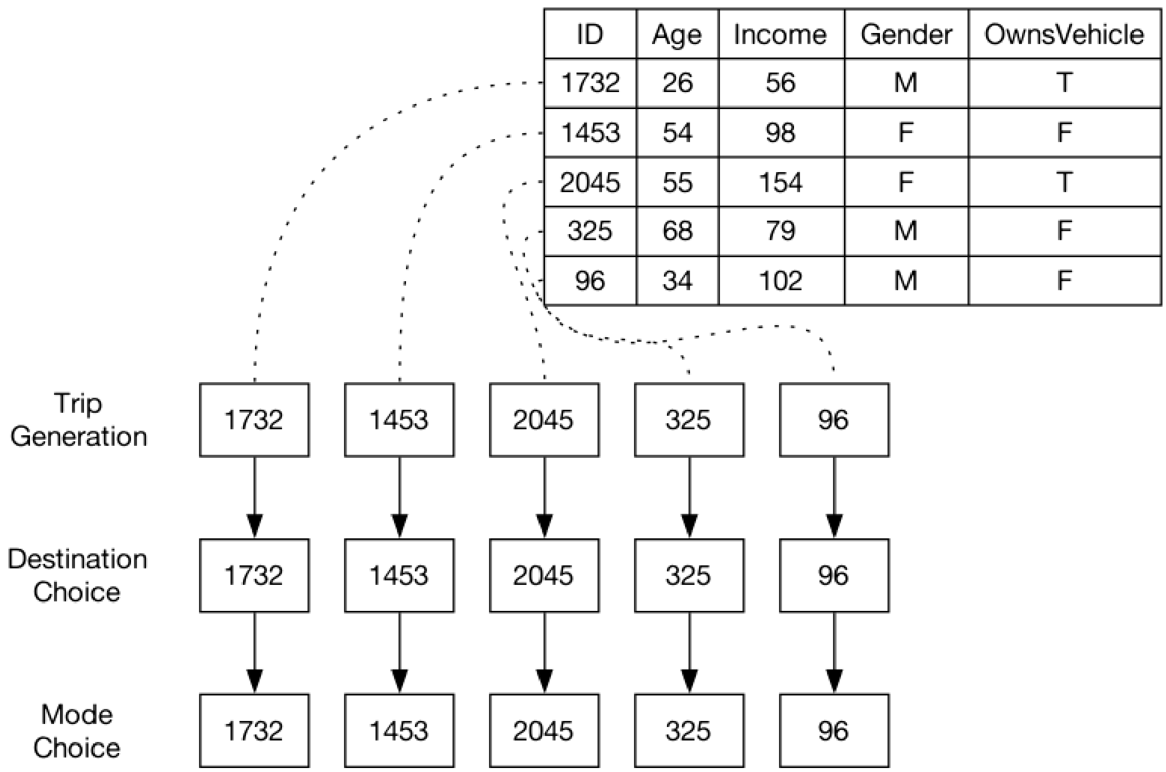
Figure 1.1 gives a visualization of an example "information pipeline" for a trip-based model using aggregate data and an ABM using a synthetic population. In the aggregate data model, it is impossible to know which trips are made by e.g. 2-worker, 1-vehicle, low-income households; it only describes which trips are made by households with fewer vehicles than workers. With a synthetic population, however, *individuals* are being modeled, and so each trip can be traced to a specific person. All information is known at each point in the model regardless of which data is used in previous steps.

1.2.2 Travel Behavior

The other primary difference between trip-based models and ABMs—and the main difference from trip-based microsimulation models—is that ABMs organize travel into "tours", a sequence of trips that begin and end at the home, rather than just trips. It should be noted that Miller (2023) argues that many current "activity-based" models ought to be labeled "tour-based" due to this focus on building tours. This is contrasted with "activity scheduling" models, in which activity participation is modeled explicitly and trips emerge as the means to get from one activity to the next. However, in practice there are few true "activity scheduling" models, and the term "activity-based" is used to refer to both activity scheduling and tour-based models.



(a) Trip-based



(b) Activity-based

Figure 1.1: Example "information pipeline" for a trip-based model and an ABM.

In a typical trip-based model, trips are forecasted based on empirical trip rates, usually by trip purpose and by household type (for example, low-income, 1-vehicle households make a certain number of "home-based work" trips) (McNally 2007). These trips are then assigned an origin and destination, mode, and often a time of day (peak/off-peak, etc.), resulting in a list of trips between each zone by mode and purpose. A trip-based microsimulation model may use regression models rather than aggregate data for some of the model steps (Moeckel et al. 2020), but the end result is similar: a list of trips by person, noting mode and purpose. However, this trip list may be inconsistent, and the forecasted trips may not be physically possible to complete in any sequence, as there is no sense of "trip-chaining". The hope, though, is that over an area the inconsistencies would cancel out, leaving an overall accurate forecast.

ABMs, on the other hand, model *tours* rather than trips. This attempts to create consistency in trip origins/destinations, mode choice, and time of day: since each trip is a part of a "chain" (tour), the trips within a tour are dependent on each other (Rasouli and Timmermans 2014). The open-source ABM ActivitySim (Association of Metropolitan Planning Organizations 2023), for example, has a tour-scheduling model that determines the number of "mandatory" (work, school, etc.) and "discretionary" tours each individual will make, and performs tour-level mode and destination choice for each tour. After the tour-level decisions are made, trip-level mode/destination choice is done for each trip in the tour, including the possible addition of subtours (see Vovsha, Bradley, and Bowman (2005), fig. 18.1).

Figures **?@fig-network-aggregate** and **?@fig-network-synth** show an example of the trips assigned to a network in the various model forms. **?@fig-network-aggregate** depicts network assignment in a typical trip-based model where the total number of trips between each zone is given. With these results, the mode and purpose of each trip is known, but, with aggregate data, there is no way of telling who made which trips other than the segmentation in the previous steps (see Figure 1.1a). It is also not possible to construct a coherent daily list of trips for individuals.

?@fig-network-synth, on the other hand, depicts visual representations of an *individual's* travel made possible by the use of a synthetic population. **?@fig-network-synth-1** depicts the trip list that could be given for an individual in a trip-based microsimulation model. Though each individual's trips are known, there is no guarantee of consistency between trips. For example, it could predict that the individual takes transit to work but then drives home or that the individual makes two recreational trips without ever making a return trip. The activity-based approach, depicted in **?@fig-network-synth-2**, attempts to add this consistency by modeling tours, and only generating trips consistent with each tour.

In addition to intra-person dependencies, Rasouli and Timmermans (2014) note that ABMs can model dependencies between members of a household as well. A vehicle can't be used by multiple people in the same household at the same time to travel to different destinations. Because the people within the household will have travel patterns that depend on the patterns of others in the household, a policy affecting one person in the household can affect everyone in the household no matter how directly the policy connects to them (Vovsha, Bradley, and Bowman 2005). These effects aren't possible to forecast in a trip-based model.

Another advantage of organizing travel into tours comes regarding accessibility analyses. Dong et al. (2006) note that when trip-based models are used to analyze accessibility, each zone must be analyzed independently of travel behavior. This approach only analyzes zones' proximity to each other and does not take into account individual travel patterns. They argue that this is a limited view of accessibility, and discuss the "activity-based accessibility measure" (ABA), which is evaluated based on all trips in a day rather than particular trips. As an example, if an individual doesn't live within a 20-minute drive of a grocery store, traditional measures might rate this as poor accessibility. However, if they travel to a park with a nearby grocery store as part of their DAP, then in reality the accessibility should be rated much higher. This proximity may even influence *which* park is chosen. Overall, they found that the ABA predicts more reasonable accessibility outcomes compared to traditional measures.

1.3 Lack of ABM Adoption

Though ABMs have many clear advantages over trip-based models, adoption among agencies has been relatively slow. Many ABMs are implemented in proprietary software, which creates difficulty in maintaining and iterating on the model, Miller (2023) argues. Even in an open-source model like ActivitySim (Association of Metropolitan Planning Organizations 2023), Miller notes several disadvantages of ABMs:

- Computational inefficiency and complicated program design: ABMs take more time, more computing power, and more money to run. This is because the synthetic population needed to run an ABM uses much more data. In areas with thousands of TAZs and millions of people, a super computer is needed, and it will cost much more than what is spent to run trip-based models. If a city can see similar results using a trip-based model, they may decide not to invest in an ABM.
- Absence of a standard model system: The modeling systems are often designed with different approaches and for specific areas making it hard to transfer from one urban area to another. This also makes it difficult for agencies to determine which approach is the best and decide which to implement. In relation to this, Miller also states that the pressures of publishing unique and ground-breaking research in academia can deter researchers from converging towards best theories and methods.
- Lack of resources: Most of these models were developed in academic settings which often lack resources, and possibly desire, to put them into practice. This leaves it up to governments and consultants to put the models into practice, but they can be hesitant to promote software development and to invest in new systems.

For these reasons, as well as the inertia of current practices, the majority of agencies and organizations in the US remain using trip-based models for demand forecasting and policy analysis.

1.4 Research Gap

Although there has been much research on ABMs and their theoretical advantages, practical comparisons of the model frameworks have been limited. It is often taken as given that ABMs are unilaterally superior to traditional trip-based models due to their better theoretical foundation, but it is not clear if that better foundation always yields better results in terms of analytical flexibility or policy outcomes. Ferdous et al. (2012) compared the trip- and activity-based model frameworks of the Mid-Ohio Regional Planning Commission and found that the ABM performed slightly better at the region level, but about the same at the project level. This is not an overwhelming victory for ABMs, and so based on this an agency may reasonably decide that the increase in data, computational, training, and potentially financial requirements outweigh the potential gains of switching to an ABM.

Ferdous et al. (2012), however, mainly compared the *accuracy* of the two frameworks, but did not address the methodological differences between them. What types of data collection/synthesis are needed for each model type? Are there analyses that can only be done through (or that are made easier by) one of the model types? What would an agency need in order to transition from a trip-based model to an ABM? Are certain types of scenarios suited to one model type? Though some of these questions have been discussed (see e.g. Lemp, McWethy, and Kockelman 2007), a holistic methodological comparison is lacking. Additionally, the answers in the current literature are mainly theoretical, with little use to an agency considering the transition.

This research aims to answer these questions by providing a side-by-side comparison of a potential trip-based and activity-based modeling methodology. Several "proposed development" scenarios are run in each model, and the strengths and weaknesses of each approach are compared. It is important to note that this paper is not focused on model accuracy, as in any model type this can be adjusted dramatically through calibration efforts. Rather, the focus is on the methodological differences between the approaches, and the types of analyses that can be done with each model type.

2 Methodology

This paper seeks to compare methodological differences between trip- and activity-based modeling frameworks. The Wasatch Front Regional Council (WFRC) travel demand model is used as a representative trip-based model, and an ActivitySim implementation in the same study area is used as a representative activity-based model (ABM). Note that the focus is not on comparing model accuracy or performance, but rather on comparing the process of using each model, including the types of analyses that can be performed. There are therefore few direct comparisons of model outputs between each type. Instead, this research highlights the strengths and weaknesses of each model type in planning and policy analysis, and illustrates these differences.

The following sections discuss the specific models in more detail.

2.1 WFRC Model

The WFRC model is implemented in the CUBE software by Bentley (Bentley Systems n.d.), and is currently used by WFRC for modeling travel in the Salt Lake City, Utah area. WFRC provided the model directly, including land use forecasts and the current long-range transportation plan. The model is taken essentially as-is, with no changes other than those noted in Chapters 3–5.

The WFRC model, like many trip-based models, requires the following inputs:

- Land use data, including information about population, employment, and socioeconomic variables such as income, delineated by transportation analysis zone (TAZ). This is provided by WFRC directly, as an output of their land use forecasting model(s).
- Travel skims, detailing travel time, cost, etc. between each origin-destination pair of TAZs. The WFRC model uses an iterative process of assigning volumes to the transportation network and recalculating the skims, which are used in the mode and destination choice model steps.
- Transportation networks, including highway, transit, etc. networks which connect the TAZs to each other. These networks contain information such as link speed and capacity. Though the WFRC model assigns travel volumes to the network, this paper does not compare the model's network assignment results. However, the network volumes are still used to calculate the loaded network skims.

- Lookup tables, used in many model steps for information such as trip rates by household type. These are taken directly from the WFRC model without modification.
- Model constants and coefficients, which some model steps such as mode choice require for calibration. These are also taken directly from the WFRC model.

Like many trip-based models, the WFRC model follows the "four-step" approach and has main steps of trip generation, trip distribution, mode choice, and network assignment. The model also includes a disaggregation step at the beginning where the TAZ-level data is used via lookup tables to estimate the number of households by size, income group, number of workers, and auto ownership. The trip generation step uses lookup tables for each household type to determine the number of trips produced by purpose. Trip attractions are determined based on the number of jobs in each TAZ, with differing coefficients by job type. Trip distribution uses a gravity model of the form

$$T_{ij} = P_i \times \frac{A_j F_{ij}}{\sum_J A_j F_{ij}},$$

where T_{ij} is the number of trips from zone i to j, P_i is the productions at i, A_j is the attractions at j, F_{ij} is the cost term/function from i to j, and J is the set of all zones trips from i can be attracted to. Mode choice uses a choice model to assign a percentage of trips of each purpose to each mode. Network assignment is done via an iterative process to equalize travel time between potential routes.

The WFRC model outputs include trip tables by purpose, mode, and time of day, as well as loaded network skims.

2.2 ActivitySim

ActivitySim is an open-source ABM led by a consortium of transportation planning agencies. ActivitySim is highly configurable, and many agencies have their own bespoke implementation. This paper uses an ActivitySim implementation based on (**macfarlane2021?**), which is in turn based on the prototype configuration for the Metropolitan Transportation Commission serving the San Francisco area (**cite?**).

ActivitySim requires similar inputs to the WFRC model, though it does not assign traffic and so does not require any transportation networks. Additionally, ActivitySim requires population data at an individual level, including information such as age, household income, and home location. Due to privacy concerns, real data is rarely used for this purpose, and a synthetic population representative of the study area is used instead. Section 2.2.1 discusses the population used in more detail.

ActivitySim, like all ABMs, simulates transportation decisions on an individual level. ActivitySim has a hierarchical decision tree, where long-term decisions (such as auto ownership and telecommute frequency) are made first, followed by daily and tour- and trip-level decisions such as scheduling and mode choice (see Figure 2.1). Each of these steps determines information that will be used in subsequent steps, and many steps can be turned on or off depending on what is needed for the model implementation.

The steps can broadly be categorized into five groups, as shown in Figure 2.1: aggregate, household/personal, daily, tour-level, and trip-level steps. The aggregate steps mainly involve determining impedance measures between each pair of zones (travel time, distance, cost, etc.). In this case, these impedances are supplied directly as network skims, output from the WFRC model.

The household/personal steps relate to long-term decisions that are unlikely to change quickly based on daily transportation conditions. These steps include determining remote work status, work/school location, auto ownership, transit pass ownership, and free parking availability at work. Much of this information can be supplied directly or explicitly modeled. This ActivitySim implementation does not supply any of this information directly, and explicitly models remote work status, work/school location, auto ownership, and free parking availability. Transit pass ownership is not modeled.

The daily decisions primarily concern an individual's DAP. ActivitySim contains a step to assign mandatory, non-mandatory, and joint tours based on personal and household information (joint tours combine both mandatory and non-mandatory activities). For example, full-time workers are more likely to make a mandatory tour than part-time workers, all else being equal.

Once a DAP is chosen, ActivitySim creates tours for each major activity in the day. Additionally, ActivitySim determines if an individual makes an "at-work" tour (e.g. leaving for lunch and returning to the workplace). Each tour is scheduled and assigned a primary mode, as well as a primary destination for non-mandatory and joint tours. The tours are then populated with trips, and ActivitySim assigns each trip a purpose, destination, time of day, and mode compatible with the tour-level assignment.

The final steps of ActivitySim are writing output trip matrices and other tables, including information on land use, persons, households, tours, and trips.

2.2.1 PopulationSim

This paper uses PopulationSim (**cite?**) to create a synthetic population for ActivitySim. The synthetic population aims to be representative of the study area while maintaining privacy. PopulationSim takes as input a "seed" of individuals and households, and populates the area with copies of these to match given control totals.

The seed sample comes from the 2019 American Community Survey Public Use Microdata Sample (**pums?**), which contains a sample of actual (anonymized) individuals and households at the Public Use Microdata Area (PUMA) geography (PUMAs partition the United States

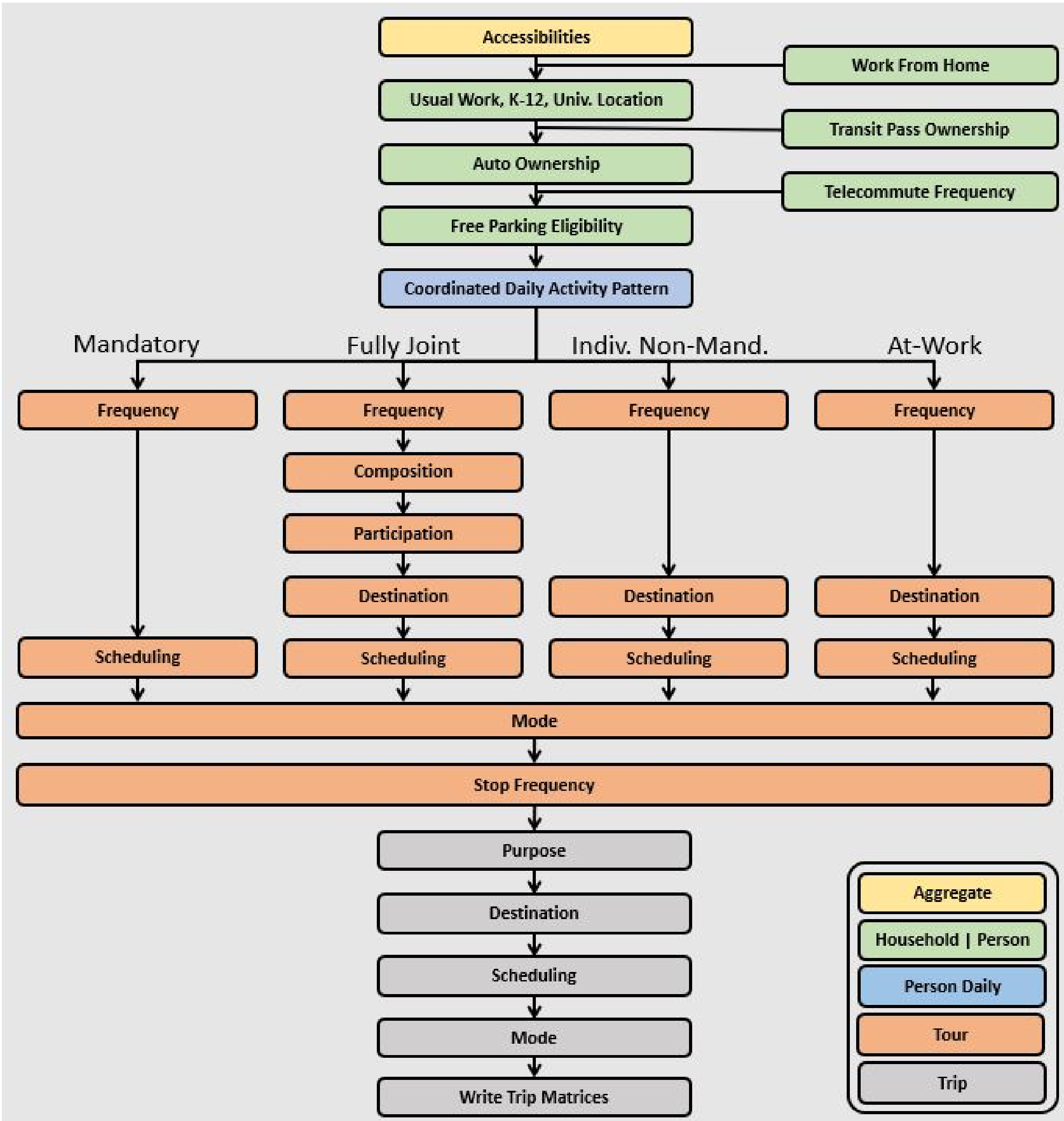


Figure 2.1: ActivitySim sub-model flowchart. Long-term decisions are made first, followed by more granular ones.

Table 2.1: PopulationSim Control Totals by Geography and Source

Control	Geography	Source	Weight
Population	Entire Region	Census	5,000
Number of Households	TAZ	WFRC Model	1,000,000,000
Household Size	Census Tract	Census	10,000
Persons by Age Group	Census Tract	Census	10,000
Households by Income Group	Census Tract	Census	500
Workers per Household	Census Tract	Census	1,000

into areas of around 100,000 people each (**puma?**)). The control totals come from two different sources: the U.S. Census and the WFRC model. Table 2.1 shows these controls as well as their geographic level and source. PopulationSim also allows setting different weights to each control, and Table 2.1 gives this information as well.

Most of these controls come from Census data, with only the number of households per TAZ coming from the WFRC model data. Note also that there are many personal and household variables that are not accounted for in these controls, such as sex, vehicle ownership, internet access, etc. These variables are not controlled for and are dependent on which seed persons or households are copied in controlling for the other variables. However, this process is assumed to still give a representative enough estimate for the uncontrolled variables without needing to model them explicitly.

The outputs of PopulationSim include a persons and households table comprising the synthetic population, as well as summary tables.

2.3 Initial Model Comparison/Calibration

While this research does not directly compare the outputs of ActivitySim to those of the WFRC model, it is important to ensure similar performance between the two models for meaningful analyses. As such, a baseline scenario in both models is used in order to calibrate the ActivitySim implementation to the WFRC model. This baseline scenario uses the 2019 WFRC model as-is. For ActivitySim, the baseline scenario uses 2019 Census and WFRC data to create the synthetic population, and the choice models use land use data and network skims from the baseline WFRC scenario.

2.3.1 Verification of the Synthetic Population

The controls for PopulationSim mostly come from the Census, as can be seen in Table 2.1. However, the WFRC model contains TAZ-level data including population and median income.

The WFRC model also has a disaggregation step that estimates the number of households by size and income group.
This section compares the output of PopulationSim to the WFRC model on each of these variables at the TAZ level.

Figure 2.2 shows the difference in TAZ population between PopulationSim and the WFRC data. It is worth noting that since the number of households was controlled at the TAZ level from the WFRC data with an extremely high weight, the number of households per TAZ in the synthetic population match exactly to the WFRC data. The average household size will therefore follow a similar error distribution to the one shown in Figure 2.2.

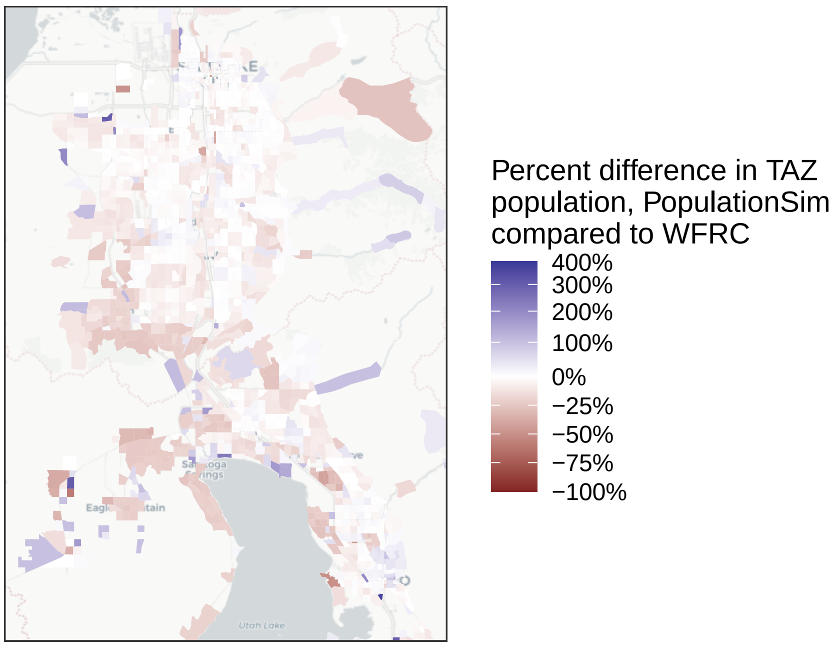


Figure 2.2: population comparison

The population per TAZ is similar to the WFRC data in most places, though there are some discrepancies especially near Herriman and Lehi. Since total population is a region-level control, but number of households is a TAZ-level control, this shows PopulationSim is predicting a smaller average household size in Herriman and Lehi than the WFRC data suggests.

Income is also an important factor in travel behavior (**cite?**), and Figure 2.3 shows a TAZ-level comparison of median income between the synthetic population and the WFRC data. The synthetic population does have a lower median income than the WFRC data in many TAZs, but the error is in most cases fairly small, especially in more populated areas. However, both the WFRC model and ActivitySim use household income *groups* rather than individual household income to inform travel decisions. These groups are taken from the WFRC model

Table 2.2: income groups

Income Group	Income Range
1	$45,000
2	$45,000–$75,000
3	$75,000–$125,000
4	$125,000

(see Table 2.2), and the groups in PopulationSim and ActivitySim were adjusted to match. Figure 2.4 shows the difference in number of households by income group, and this figure shows a similar trend of PopulationSim over-predicting low-income households.

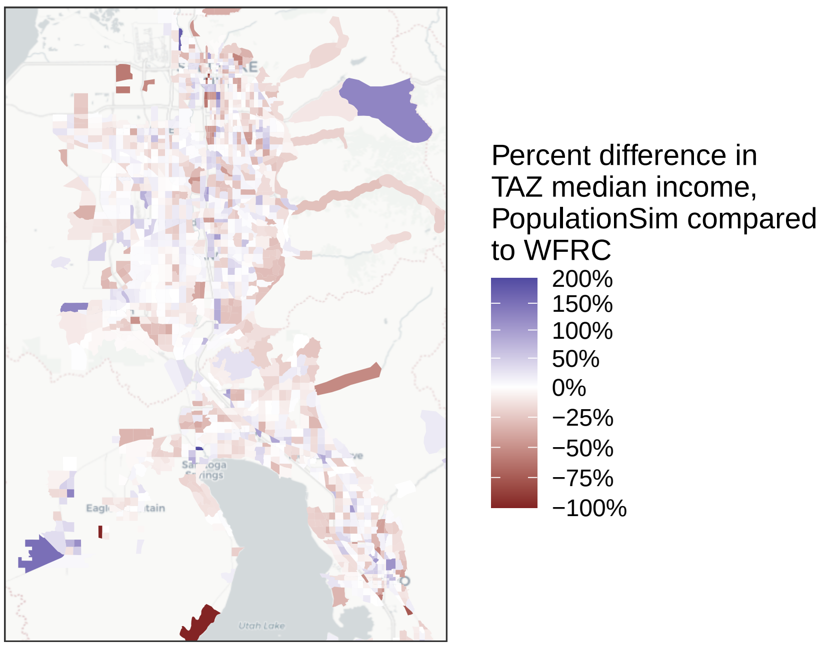


Figure 2.3: med inc

Income was not weighted very heavily as a control in PopulationSim (see Table 2.1), and this is in part why there are discrepancies between the models. However, the overall distribution of income is similar between the models, as Figure 2.5 shows.

2.3.2 Verification and Calibration of ActivitySim

This section compares the outputs of both models to verify that trip patterns roughly agree between them. There are three comparisons of interest between the outputs of the two models:

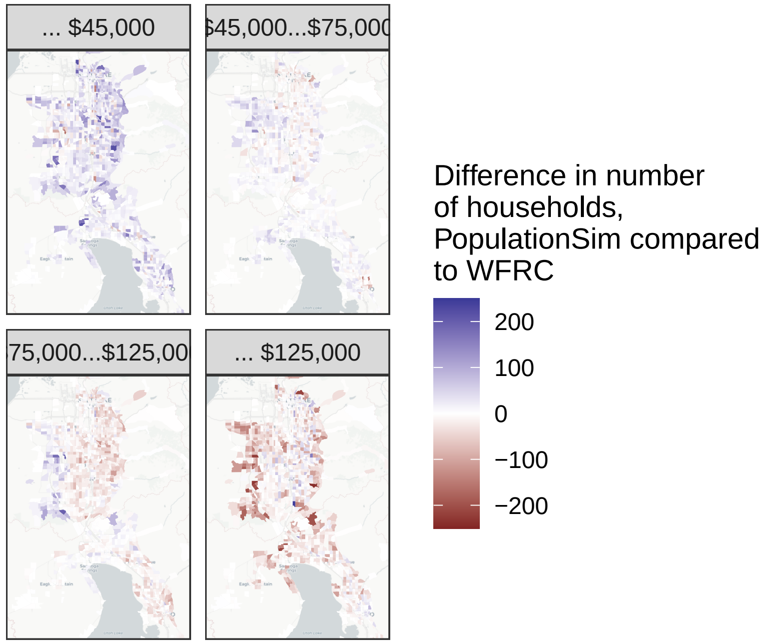


Figure 2.4: inc groups

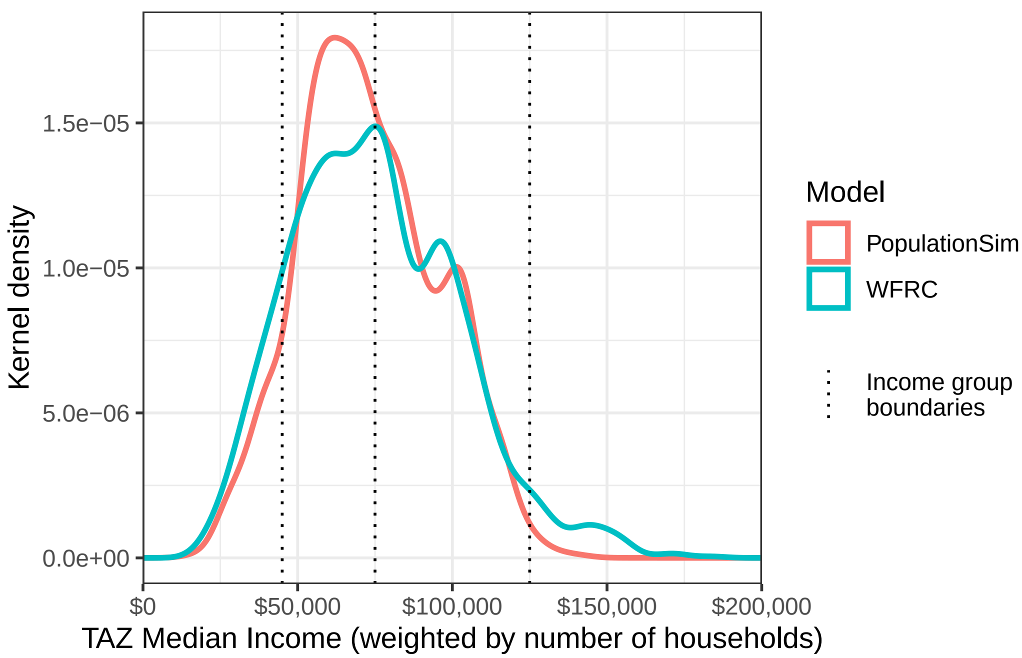


Figure 2.5: inc density

Table 2.3: mode split

Purpose	Mode	ActivitySim predicted trips	WFRC predicted trips	% Error	% Error (scaled to WFRC predicted trips)
All	All	7441918	8737729	-14.8%	0.0%
All	Auto	6515379	7905927	-17.6%	-3.2%
All	Non-motorized	768463	732250	4.9%	23.2%
All	Transit	158076	99551	58.8%	86.4%
Home-based Other	All	3696056	4641724	-20.4%	-6.5%
Home-based Other	Auto	3092740	4094887	-24.5%	-11.3%
Home-based Other	Non-motorized	531497	509491	4.3%	22.5%
Home-based Other	Transit	71819	37346	92.3%	125.8%
Home-based Work	All	1759501	1711578	2.8%	20.7%
Home-based Work	Auto	1617449	1586413	2.0%	19.7%
Home-based Work	Non-motorized	85430	76413	11.8%	31.3%
Home-based Work	Transit	56622	48752	16.1%	36.4%
Non–home-based	All	1986361	2384427	-16.7%	-2.2%
Non–home-based	Auto	1805190	2224628	-18.9%	-4.7%
Non–home-based	Non-motorized	151536	146346	3.5%	21.6%
Non–home-based	Transit	29635	13453	120.3%	158.6%

mode split, trip length frequency distribution, and remote work.

The initial baseline ActivitySim scenario predicted a mode split significantly different to that from the WFRC model, and so calibration efforts were needed. We adjusted the alternative-specific constants in ActivitySim's mode choice model to better match the mode split in the WFRC model. Table 2.3 compares the mode split of both models after several iterations of this calibration. Overall, the calibration resulted in a reasonably similar mode split between the two models, though there are still discrepancies (e.g. ActivitySim is predicting about twice as many transit trips as the WFRC model). However, further adjustment of the constants would cause their values to become unreasonably large, as ActivitySim's mode split begins converging at this point regardless of the calibration constant values (see Figure 2.6). This is likely due to the mode choice coefficients being unrepresentative of the study area. This ActivitySim configuration is ultimately based on the San Francisco area, and so coefficients on variables such as travel time and income are calibrated for that area (hence in part why there are so many more transit trips).

Figure 2.7 compares the trip length frequency distribution of the two models by mode and purpose. Both ActivitySim and the WFRC model contain destination choice models which can be adjusted to affect the distribution of trip length. However, as the figure shows, the two models have similar TLFDs, so no adjustment was necessary. The most significant discrepancies are with transit trips, again likely due to this configuration of ActivitySim being calibrated to San Francisco, making transit more attractive. Note that though these distributions match well enough for the purposes of this research, further calibration would be required to create a production-ready ActivitySim implementation.

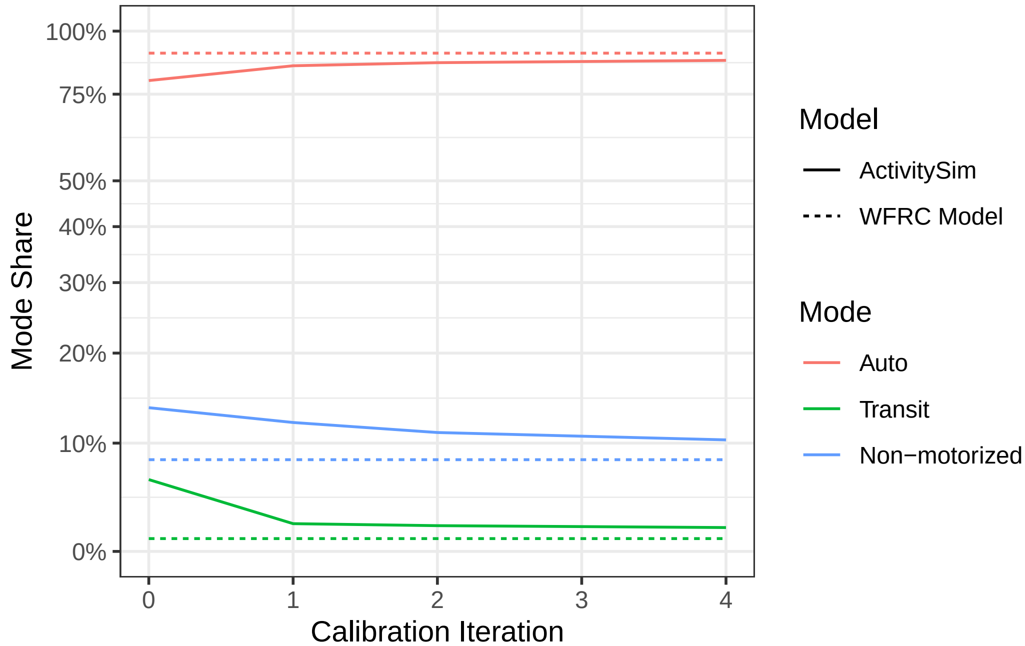


Figure 2.6: calibration plot

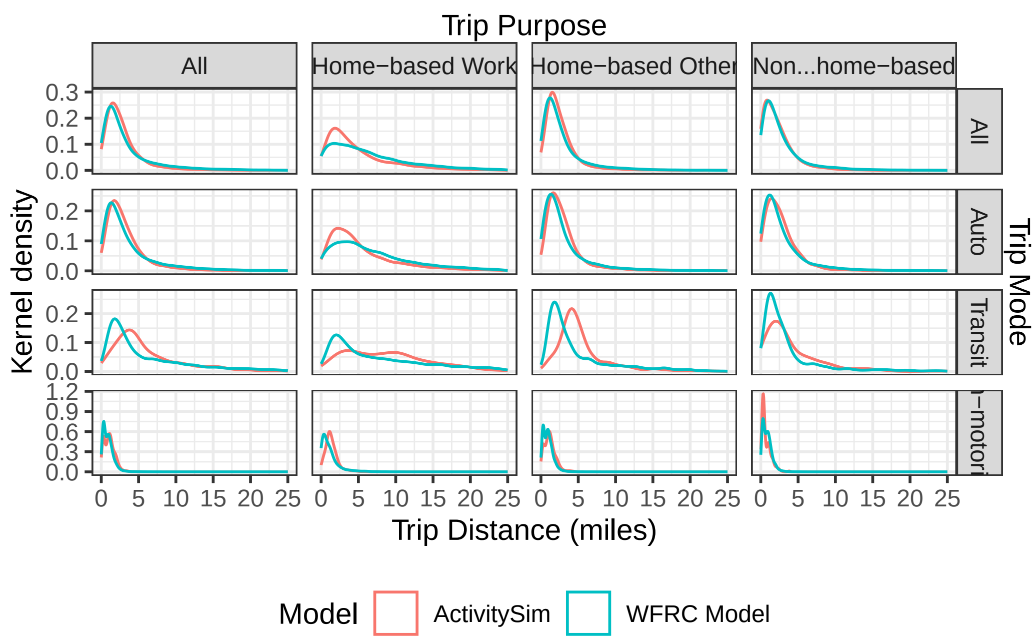


Figure 2.7: tlfd comparison

The WFRC model has basic support for predicting telecommuting and work-from-home trips. This includes a lookup table of telecommute percentages based on job type and year. ActivitySim also has this functionality, and can additionally use individual- and household-level variables in its predictions. It is worth noting that both the WFRC model and ActivitySim make a distinction between "telecommuting", where an individual commutes to work some days and does not others, and "work-from-home" (or "home-based jobs" in the WFRC model), where an individual's workplace is always at their home.

Both models treat "work-from-home"/"home-based jobs" similarly. The WFRC model's land use data contains employment by type in each TAZ, and it considers a "home-based job" as a separate job type, so these are not counted toward employment totals in trip generation and subsequent steps. Figure 2.8 shows the distribution of home-based job percentage by TAZ in the WFRC land use data. ActivitySim has a "work from home" submodel which assigns workers work-from-home status based on personal variables such as income, sex, and education (coefficients on these variables were left unchanged from the existing configuration). There is also a "target work-from-home percent" value that adjusts the model to reach the specified work-from-home proportion of all workers. Individuals with work-from-home status are then prohibited from making a mandatory tour. This target work-from-home percentage is set at 2.3%, based on a weighted average from the WFRC model data. There are no other adjustments to the ActivitySim work-from-home submodel.

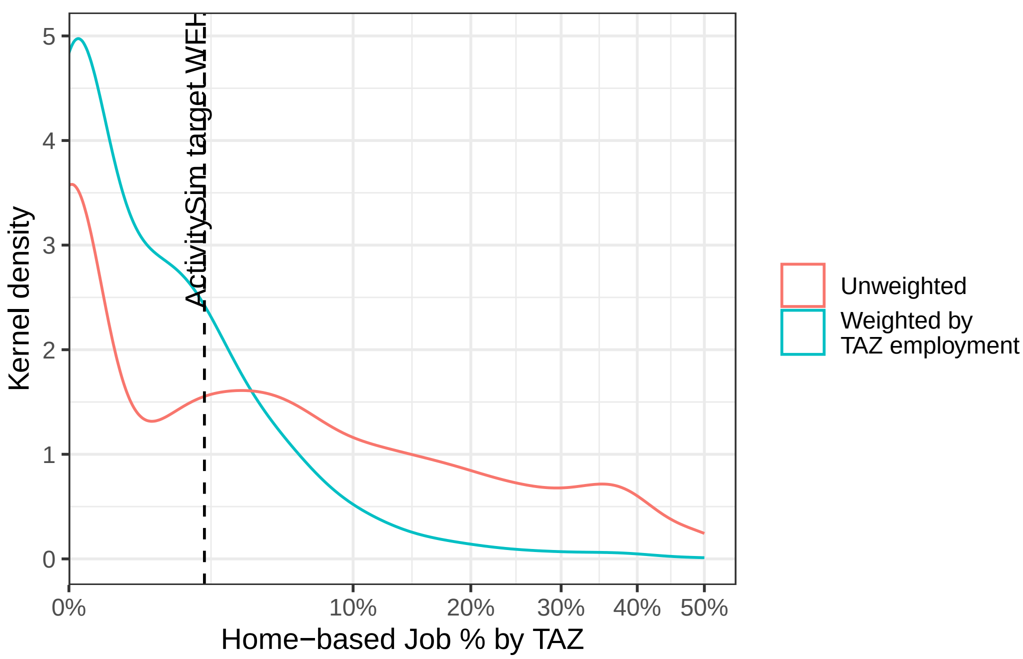


Figure 2.8

The two models differ in their approach to telecommuting, however. The WFRC model has a

Table 2.4: wfrc telecommute data

Job Type	WFRC Telecommute Proportion	ActivitySim Telecommute Utility Coefficients		
		1 day	2–3 days	4 days
Retail	2.70%	0.312	0.125	0.078
Food	1.87%	-0.368	-0.148	-0.092
Manufacturing	2.02%	0.038	0.015	0.010
Office	6.66%	1.782	0.712	0.445
Gov't/Education	1.67%	-0.560	-0.224	-0.140
Health	2.86%	0.158	0.063	0.039
Agriculture	6.93%	2.262	0.904	0.566
Mining	0.53%	-2.030	-0.810	-0.511
Construction	3.28%	0.816	0.326	0.204
Other	5.37%	1.535	0.614	0.384

lookup table of telecommuting shares based on job type (see Table 2.4), including predictions for future years. ActivitySim has a "telecommute frequency" submodel which assigns workers a telecommute status indicating the number of days they work remotely per week. Based on this status, ActivitySim adjusts the likelihood of making a mandatory tour. Telecommute status depends on personal variables similar to those in the work-from-home submodel by default. In order to calibrate to the WFRC data, however, we added additional job type variables to ActivitySim to match those given in Table 2.4. Because these are choice coefficients rather than target percentages, the values needed to be calibrated to match the WFRC targets. The calibration allowed ActivitySim to match these targets exactly, and the coefficients are given in Table 2.4.

2.4 Example Scenarios

There are three proposed scenarios to implement and run in each model for comparison.
This is not a comprehensive list covering all potential scenario possibilities, but is taken to represent many of the types of scenarios an agency would analyze.

The first scenario involves a change in land use near the former state prison site in Draper, Utah. Current plans for this site involve new development known as "The Point", which will add high-density housing and commercial development to the area.
This research scenario will be based on this development, but will include only the land use changes. The actual development plans also include expansion of transit, but this will not be a part of this scenario.

The second scenario centers around an augmentation of transit service along the Wasatch

Front. The FrontRunner, a CRT line connecting Provo to Ogden, is slated for expansion. The expansion includes additional stations and increased travel speeds due to vehicle electrification. This scenario models these changes in accordance with the planned expansion of the service.

The third scenario addresses the growing trend of remote work. Given technological advancements and the notable surge in remote work during the COVID-19 pandemic, this scenario models a substantial increase in remote work based on projections from WFRC.

Each of these scenarios is based on the baseline 2019 scenario in the respective model, and ignores any additional expected growth or development beyond the specific changes of each scenario. For example, the increased WFH scenario uses WFH projections from 2050, but land use and socioeconomic data from 2019. These scenarios are therefore not realistic, but they serve as illustrative examples of the types of planning and development scenarios agencies may wish to analyze.

All three of these scenarios are coded in both the WFRC model and ActivitySim. The results (Chapters 3–5) describe the process of coding each scenario and analyzing them, as well as the analyses themselves.

3 Land Use

In our first scenario, we changes the land use at The Point.

The change is based on WFRC 2050 SE data forecast, in turn based on development plan of The Point.

This adds X jobs and X households, etc. (show map/table)

Table 3.1 shows the change in lu data, and Figure 3.1 shows these TAZs.

3.1 Creating the scenario

In the WFRC model the change is trivial to implement. We just copy/pasted the se data from 2050 to 2019 for the affected zones.

ActivitySim requires two things for this scenario: updated SE data and a new pop. The updated SE data is the same as in the WFRC scenario, so is also trivial to implement. The pop, however, is not. New controls.

New population for only the affected area, then joined to existing. There were no existing households so we didn't need to remove any (we're not double-counting).

? Note in realistic scenario you could use just the land use forecast as controls directly, rahter than the whole census shenanigains but we didn't do that.

Table 3.1: the point overview

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

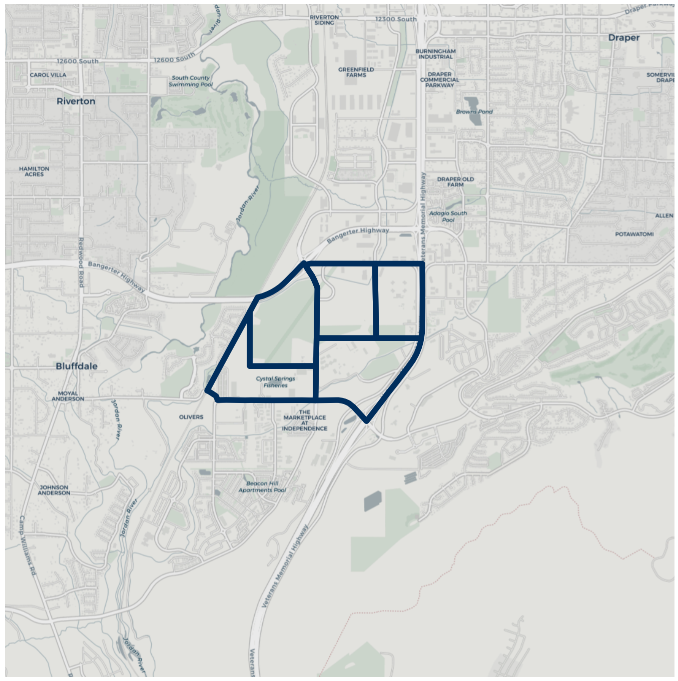


Figure 3.1: the point zones

3.2 Analyses

We looked at person-miles for both scenarios. Figure 3.2 shows the increase in person-miles for the ABM. Shaded based on “in-zone”. (Not many transit trips since we didn’t add transit). Figure 3.3 shows for TBM.

Can’t really shade since nhb relocation happens before mode choice. Best you can do is difference the NHB matrices (Figure 3.4 (2-panel)) (If there are nhb trips in/out of zones): clearly some of these are not nhb from residents since they occur in these zones. **Big problem:** They’re *all* in the new zones

With abm, you can follow people and trace their exact paths (Figure 3.5).

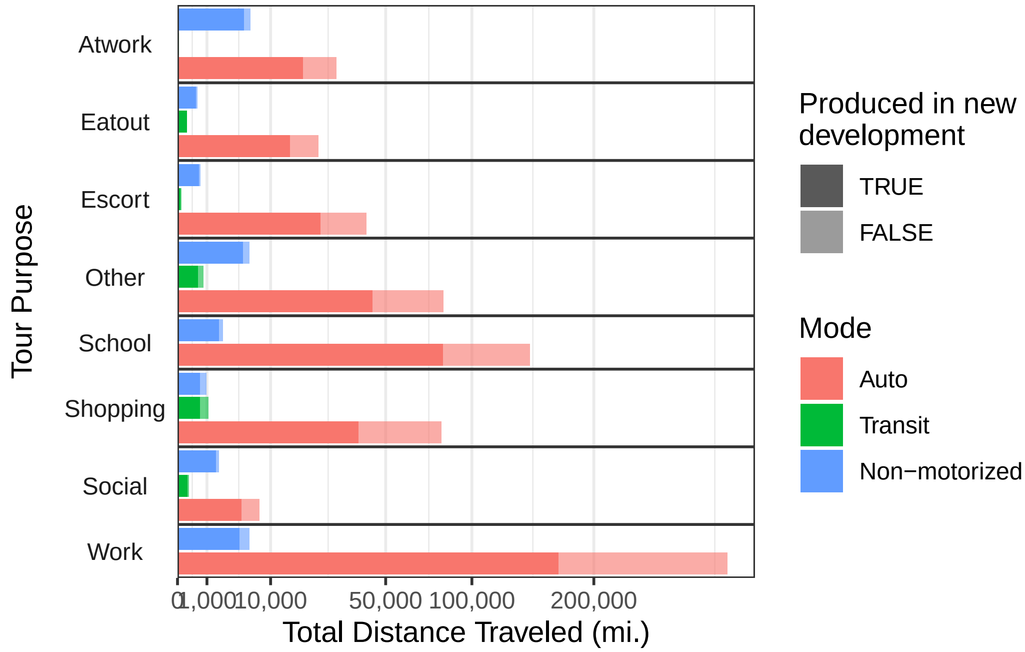


Figure 3.2: personmiles abm

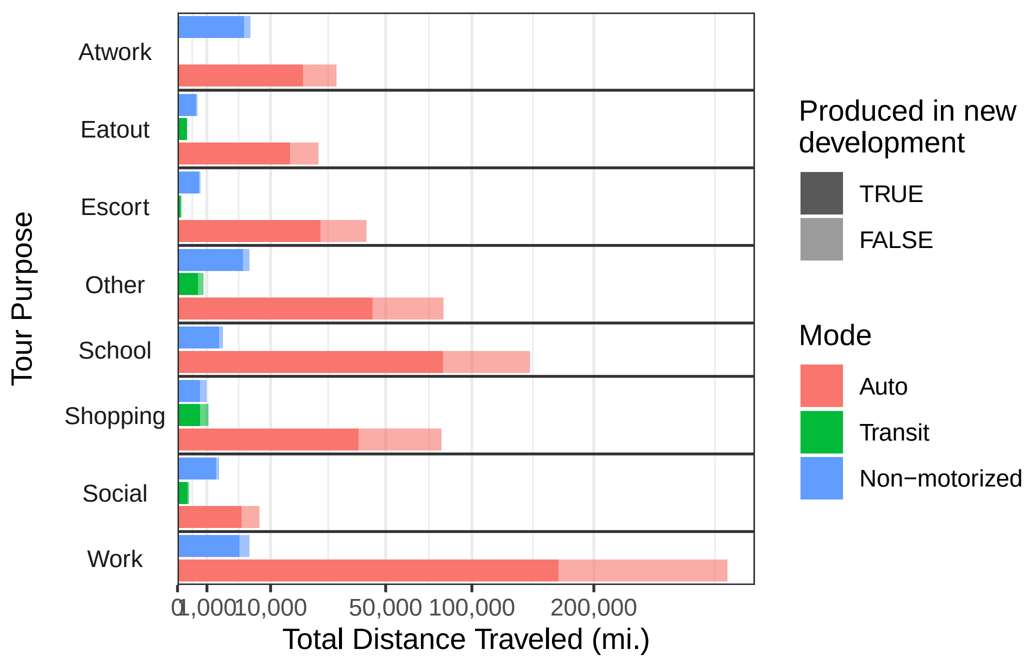
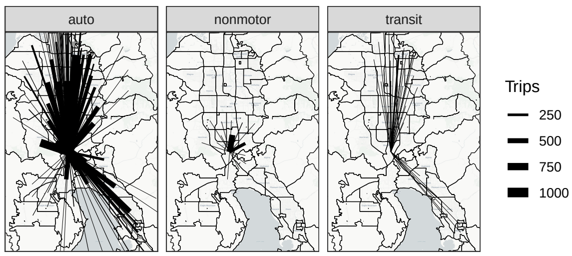
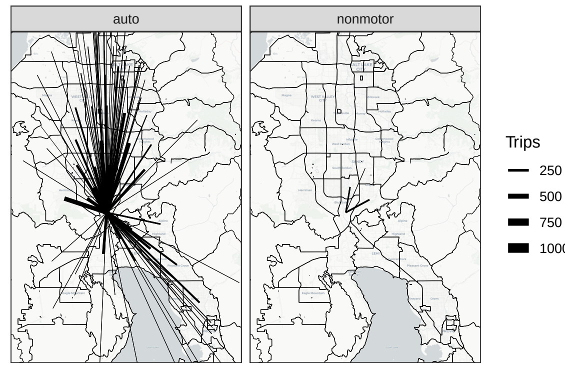


Figure 3.3: personmiles tbm



(a) hb



(b) nhb

Figure 3.4: lu desire tbm

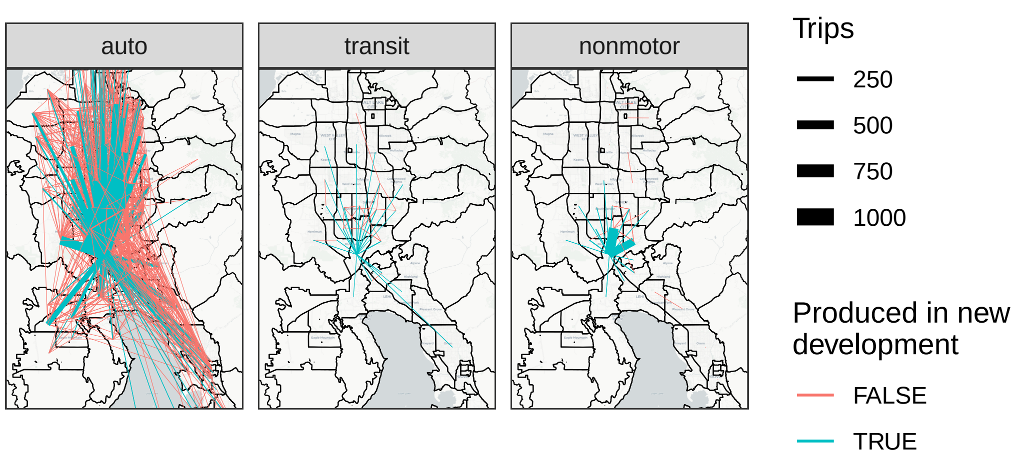


Figure 3.5: desire abm

4 Transit

The second sceario updates the transit network (frontrunner).

Data comes from the WFRC 2050 build plan. Includes new stations, new headways, and faster speeds.

Show map of changes.

4.1 Creating scenario

WFRC model needs three adjustments: edit the headways, point the model to the new frontrunner network (which includes speeds), and add park-and-ride option at new station locations in highway? network.

Asim just needs new skims, gotten from the new WFRC scenario.

4.2 Analyses

5 WFH

Third scenario models WFH.

Uses WFRC projections from 2050.

Show plot of some kind showing increase from 2019 to 2050.

5.1 Create scenario

WFRC needs only to move 2050 data to 2019 for TC% and HBJ (in SE file ***ask Kamryn if we did this part***).

Asim is the same as calibration for the baseline scenario, but with different target %s. List these. (table of %s and coefficients for 2019/2050 side-by-side?)

Table 5.1: wfrc telecommute data

5.2 Analyses

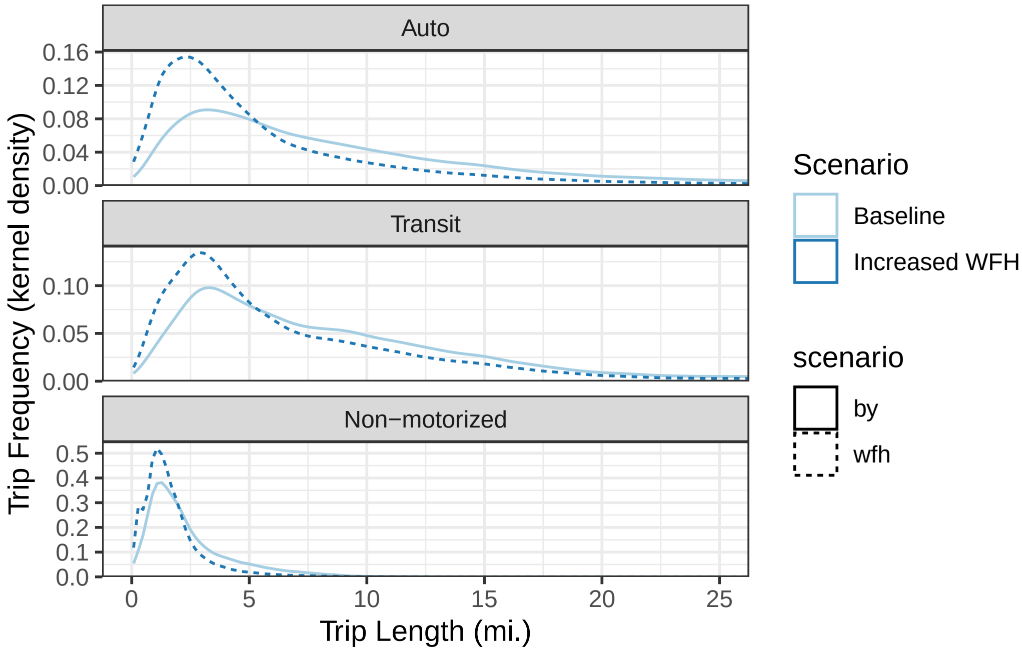


Figure 5.1: asim

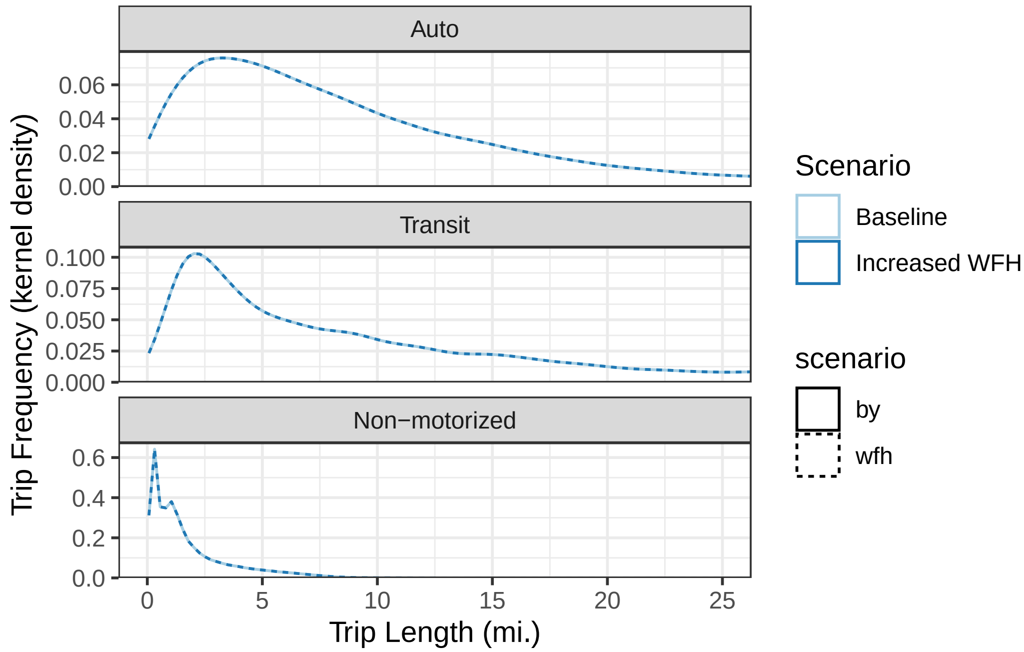


Figure 5.2: wfrc

6 Summary

In summary, this book has no content whatsoever.

6.1 Computational resources

This research also includes an analysis of the computational resources required to run each model.

The WFRC model….

ActivitySim….

?

References

Association of Metropolitan Planning Organizations. 2023. "ActivitySim." *ActivitySim*. https://activitysim.github.io/.

Bentley Systems. n.d. "CUBE Modeling Software." Accessed October 3, 2023. https://www.bentley.com/software/cube/.

Bills, Tierra S., and Joan L. Walker. 2017. "Looking Beyond the Mean for Equity Analysis: Examining Distributional Impacts of Transportation Improvements." *Transport Policy* 54 (February): 61–69. https://doi.org/10.1016/j.tranpol.2016.08.003.

Bowman, John L. 1998. "The Day Activity Schedule Approach to Travel Demand Analysis." Ph.D. Dissertation, Cambridge, MA: Massachusetts Institute of Technology. https://dspace.mit.edu/bitstream/handle/1721.1/16731/44704881-MIT.pdf?sequence=2.

Castiglione, Joe, Rachel Hiatt, Tilly Chang, and Billy Charlton. 2006. "Application of Travel Demand Microsimulation Model for Equity Analysis." *Transportation Research Record: Journal of the Transportation Research Board* 1977 (1): 35–42. https://doi.org/10.1177/0361198106197700105.

Chicago Area Transportation Study. 1959. *Final Report In Three Parts*. Vol. 1–3. Printed by authority of the State of Illinois. http://archive.org/details/chicagoareatrans01chic.

Dong, Xiaojing, Moshe E. Ben-Akiva, John L. Bowman, and Joan L. Walker. 2006. "Moving from Trip-Based to Activity-Based Measures of Accessibility." *Transportation Research Part A: Policy and Practice* 40 (2): 163–80. https://doi.org/10.1016/j.tra.2005.05.002.

Ferdous, Nazneen, Lakshmi Vana, John L. Bowman, Ram M. Pendyala, Gregory Giaimo, Chandra R. Bhat, David Schmitt, Mark Bradley, and Rebekah Anderson. 2012. "Comparison of Four-Step Versus Tour-Based Models for Prediction of Travel Behavior Before and After Transportation System Changes." *Transportation Research Record* 2303 (1): 46–60. https://doi.org/10.3141/2303-06.

Jones, Peter M. 1979. "New Approaches to Understanding Travel Behaviour: The Human Activity Approach." In *Behavioural Travel Modelling*. Routledge.

Lemp, Jason D., Laura B. McWethy, and Kara M. Kockelman. 2007. "From Aggregate Methods to Microsimulation: Assessing Benefits of Microscopic Activity-Based Models of Travel Demand." *Transportation Research Record: Journal of the Transportation Research Board* 1994 (1): 80–88. https://doi.org/10.3141/1994-11.

McNally, Michael G. 2007. "The Four-Step Model." In *Handbook of Transport Modelling*, edited by David A. Hensher and Kenneth J. Button, 1:35–53. Emerald Group Publishing Limited. https://doi.org/10.1108/9780857245670-003.

Miller, Eric. 2023. "The Current State of Activity-Based Travel Demand Modelling and Some Possible Next Steps." *Transport Reviews* 0 (0): 1–6. https://doi.org/10.1080/01441647.

2023.2198458.

Moeckel, Rolf, Nico Kuehnel, Carlos Llorca, Ana Tsui Moreno, and Hema Rayaprolu. 2020. "Agent-Based Simulation to Improve Policy Sensitivity of Trip-Based Models." *Journal of Advanced Transportation* 2020 (February): e1902162. https://doi.org/10.1155/2020/1902162.

National Academies. 2012. *Travel Demand Forecasting: Parameters and Techniques*. Washington, D.C.: National Academies Press. https://doi.org/10.17226/14665.

Ortúzar, Juan De Dios, and Luis G. Willumsen. 2011. *Modelling Transport*. 1st ed. Wiley. https://doi.org/10.1002/9781119993308.

Park, Keunhyun, Sadegh Sabouri, Torrey Lyons, Guang Tian, and Reid Ewing. 2020. "Intrazonal or Interzonal? Improving Intrazonal Travel Forecast in a Four-Step Travel Demand Model." *Transportation* 47 (5): 2087–2108. https://doi.org/10.1007/s11116-019-10002-0.

Pinjari, Abdul Rawoof, and Chandra R. Bhat. 2011. "Activity-Based Travel Demand Analysis." In *A Handbook of Transport Economics*, edited by André de Palma, Robin Lindsey, Emile Quinet, and Roger Vickerman. Edward Elgar Publishing.

Rasouli, Soora, and Harry Timmermans. 2014. "Activity-Based Models of Travel Demand: Promises, Progress and Prospects." *International Journal of Urban Sciences* 18 (1): 31–60. https://doi.org/10.1080/12265934.2013.835118.

Vovsha, Peter, Mark Bradley, and John L. Bowman. 2005. "Activity-Based Travel Forecasting Models in the United States: Progress Since 1995 and Prospects for the Future." In *Progress In Activity-Based Analysis*, edited by Harry Timmermans, 389–415. https://trid.trb.org/view/759300.

Walker, Joan L. 2005. "Making Household Microsimulation of Travel and Activities Accessible to Planners." *Transportation Research Record: Journal of the Transportation Research Board* 1931 (1): 38–48. https://doi.org/10.1177/0361198105193100105.

Weiner, Edward. 1997. "Urban Transportation Planning in the United States. An Historical Overview." DOT-T-97-24. https://doi.org/10.21949/1526591.