SAN DIEGO ASSOCIATION OF GOVERNMENTS

ABM3 MODEL DEVELOPMENT REPORT

Report | February 2, 2024


PREPARED FOR

San Diego Association of Governments

PREPARED BY

RSG

Data Perspectives Consulting, LLC

55 Railroad Row
White River Junction, VT 05001
802.295.4999
www.rsginc.com

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

The San Diego Association of Governments serves as the forum for regional decision-making for the San Diego region. SANDAG is governed by a Board of Directors composed of mayors, councilmembers, and county supervisors from each of the region's 19 local governments. SANDAG also serves as the Metropolitan Planning Organization (MPO) for San Diego County, whose role it is to prioritize spending on transportation projects to improve efficiency, promote safety, increase equity, and address other transportation planning objectives. The regional travel demand model is a key tool in SANDAG's toolbox used to analyze transportation and land-use projects and investments, quantify their impacts, and assess their performance relative to these objectives.

In 2009, SANDAG began development of an activity-based travel demand model, in the family of travel models referred to as CT-RAMP (Coordinated Travel Regional Activity-Based Travel Modeling Platform). The model was completed in 2013 and used for the 2015 RTP. The model was updated several times since the initial development - calibrated to new survey data, enhanced for additional sensitivities, expanded to consider emerging transportation technologies, etc. The latest version of the SANDAG ABM is referred to as ABM2+. The objective of this project is to develop Activity-Based Model 3 (ABM3) for the 2025 Regional Plan (2025 RP). The ABM3 development for the 2025 RP includes model estimation using recent surveys, ABM software update to ActivitySim, model calibration and validation, sensitivity tests, policy analysis enhancements, streamlining processes, risk evaluation, and general ABM support. The ABM3 model development report describes the models estimated as part of ABM3 model development and the data used to estimate them, the implementation of a more streamlined set of parking cost calculations, updates to the mode choice model for emerging mobility, and a description of the implementation of special market models for airport ground access and overnight visitors in ActivitySim.

2.0 MODEL ESTIMATION DATA

Household Survey Data

Household Travel Survey (HTS) data from 2016 and 2022 is the primary data component used in model development. This source of data contains the socio-demographic information of the household and the persons living in the household as well as detailed travel information for all members. This rich source of data is used in all three phases of model development: HTS data is used to estimate model coefficients, calibration targets for the majority of submodels in ActivitySim are derived from the weighted distribution of the HTS data, and statistics like VMT computed from the HTS data are a vital component of model validation.

RSG's rMove application is the source of the HTS data for this development project. rMove is a survey that can be performed on a smartphone app or through a web-based interface. Participants in the survey answer a host of questions about their household, themselves, and their typical travel patterns. Answers are typically broken down into three categories: household, person, and trip. Typical household questions include the income of the household, the number of vehicles owned, the location of the household, household type, etc.\ Person questions include age, gender, work status, student status, work/school location, work from home and telecommuting, etc.

If the person is participating in the survey through the smartphone version of the rMove app, the app will record the person's location throughout the day. The app will then ask questions around the trips that person took including checking the accuracy of the start and end times and locations , what the purpose of the trip was, the trip mode, and whether there were other people who took that trip with them. People participating on the web-based version of rMove input this information manually instead of having the app help track location and times.

Once raw rMove app data is captured, the first step is to go through an initial cleaning phase. This includes the removal of partially completed surveys, table formatting, assigning valid ID numbers to each household/person/trip, manually inspecting and correcting trip locations and times, and geocoding the locations to the zone system being used. The result of this processing is three primary tables: households, persons, and trips.

After processing raw rMove app data into household, persons, and trip tables, a weighting process is performed and can be summarized in the following four primary steps:

1. **Initial Expansion**: Calculating an "initial weight" based on the probability of selection in the survey sample design. This step essentially "reverses" the sample plan, providing higher initial weights to areas where less sampling occurred.
2. **Reweighting to account for non-response bias**: Performing an entropy maximization-based list balancing routine to match several key household and person dimensions to

ensure the weighted data accurately represent the entire survey region (and reduce sampling biases). This routine is performed using the open-source application, PopulationSim[1]. To do this step, missing values for income, gender, and race/ethnicity were imputed for those who did not provide that information.

3. **Creating day-level weights to account for multi-day survey data:** Adjusting the day-level and trip-level data to account for the fact that smartphone respondents provided multi-day travel diaries, while online respondents provided a single-day travel diary (this is the "multi-day adjustment"). These relatively simple adjustments ensure that travel analyses accurately reflect the entire survey region for a "typical" weekday (Mon-Thu) and do not over-represent smartphone respondents with multiple travel days.
4. **Adjusting for non-response bias in day-pattern and trip rates**: Adjusting the trip-level weights by data collection method (smartphone vs. online vs. call center) to account for under-reporting biases that RSG has detected in this survey and prior travel surveys. These adjustments help make the day and trip-level data more consistent and increase the accuracy of trip rates across survey participation methods.

The next phase of HTS survey data processing is to compute variables that relate to the activity-based model structure of ActivitySim. This step is performed via the Survey Processing Application (SPA). The primary purpose of the SPA tool is to group trips into tours and works by looping through the trip records for each person. Tours start and end at home (or work for atwork subtours). For each trip record that is involved with that tour, the SPA tool will transform trips into linked trips if necessary. This is often only necessary for transit trips where access and egress trips are combined with the transit trip to create a linked trip. For example, if a person drives their car to a transit center, hops on the train, and then walks to work, those three legs will be combined into one park-and-ride linked trip. The SPA tool will also check if the trip is reported to be with other household members. If so, there is logic to see if a matching trip exists for the household member that was reported to be traveling with them. In this way, joint tours and escorting can be determined by looking at which household members are present for each leg of the tour.

The SPA tool is also responsible for the general calculation of variables used in the ActivitySim framework. This includes the determination of worker status, student status, and person types for each person, the work from home and telecommute status of each person, whether the person takes mandatory, non-mandatory, or no travel during that day, calculating the tour mode

[1] https://activitysim.github.io/populationsim/

based on a hierarchy of trip modes, creating fully joint tours and determining their participants, deciding on a tour purpose based on the trip purposes in the tour, and creating atwork subtours.

Output from the SPA tool consists of a format that mimics the output of ActivitySim and includes household, person, tour, trip, and joint tour tables. This data can then be summarized in the same manner as the ActivitySim output to produce direct comparisons of model output with observed distributions. Additionally, this data is used to perform model estimation where the parameters are fit to the observed records in the HTS data. For additional details on exactly how this data is used for model estimation and calibration, please see the following sections.

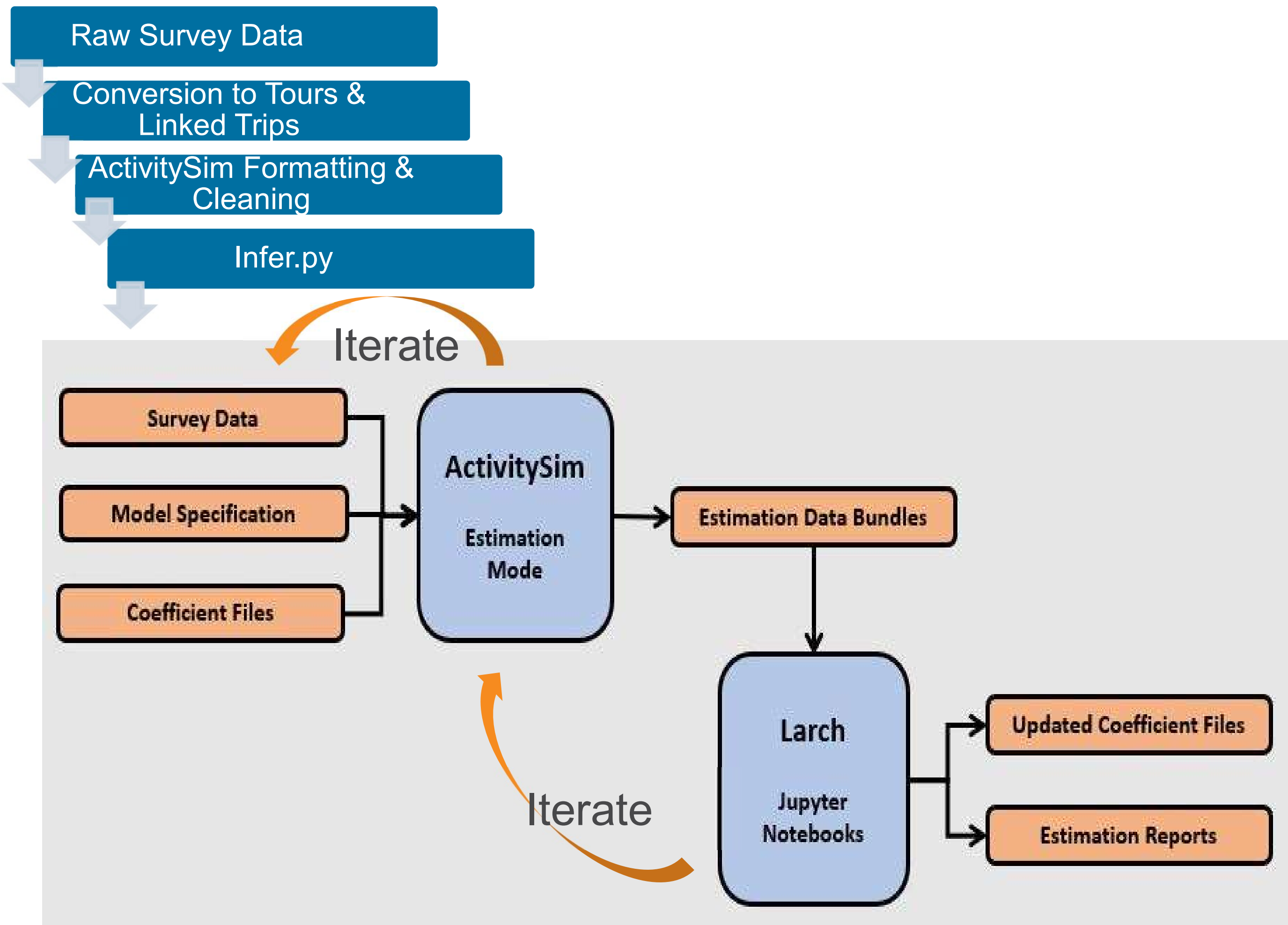
Overview of Data Preparation and Model Estimation

Model estimation in the context of activity-based development is the process by which household travel survey data is fit with a given model structure to determine the coefficients of the model. These model coefficients are then used when calculating utilities to make predictions in the model application. ActivitySim allows for multinomial or nested logit model structures and each submodel in the ActivitySim framework is estimated independently.

The general process of estimation is as follows: the SPA tool processes raw survey data into tours and linked trips as described in the previous section, then the data is further cleaned to be ActivitySim compliant and processed to have all the inputs necessary to run in ActivitySim's estimation mode. Running in ActivitySim's estimation mode produces Estimation Data Bundles (EDBs) that are read by the software package Larch[2] which fits the model to the data to produce estimated model coefficients. This procedure is shown schematically in Figure 1.

[2] See https://larch.newman.me/v5.7.0/intro.html for additional documentation.

FIGURE 1: MODEL ESTIMATION WORKFLOW



There are a couple of steps in the estimation workflow that are iterative. ActivitySim is very particular about the data that it accepts in estimation. If data is missing or represents travel not captured by the ActivitySim framework, then ActivitySim will crash when trying to create the EDBs. Thus, an extensive data preparation stage is required to make sure all the data is ActivitySim compliant. Finding every single edge case for survey data that may throw off the estimation process is often an iterative procedure where ActivitySim will crash, the user needs to figure out what in the data is non-compliant, and the input survey data is modified to avoid the crash. RSG has gone through this process in a number of different regions and has a good handle on the types of things that need to be addressed in the survey data in order for ActivitySim to work. These are described in the following data preparation section.

The other iterative part of the estimation workflow is between the model estimation and the ActivitySim config files. EDBs are created according to the specification given to ActivitySim when ActivitySim is run in estimation mode. There are often cases that exist when going through the estimation process where the user will come up with a new variable or expression to test. This often necessitates re-generating the EDBs with updated config files.

The following sections in this chapter will describe the common data preparation steps and variables needed to be able to run ActivitySim in estimation mode, actually running ActivitySim in estimation mode and the creation of Estimation Data Bundles, and the estimation procedure in Larch. Additional sections focus on how to deal with multiple survey days and multiple survey regions in the estimation process.

Data Preparation

Data preparation in this stage is centered around getting data to be completely ActivitySim compliant. In order to be compliant, every variable that is provided into ActivitySim must be a variable that could be generated by ActivitySim itself. This section will describe many of the common cases where data must be further cleaned from the output of the Survey Processing Application (SPA) described in the previous section.

If the survey data was completely accurate, and ActivitySim fully captured the travel that all individuals make, then we could just run ActivitySim in estimation mode with the formats from the SPA tool. Unfortunately, survey data is never perfect and ActivitySim does not model all possible travel scenarios. There are specific tour patterns that ActivitySim expects, and the survey data can be incomplete or inconsistent leading to ActivitySim crashes. This section contains the cleaning of the household travel survey data to get it into valid ActivitySim data and the assumptions made along the way. Considering the myriad of issues that may accompany survey data, it should be expected that other survey data sets will have issues that are not present in this document.

Households

Processing of household level data in ActivitySim is specified in the *annotate_households.csv* config file, so checking this file for region specific calculations is recommended. The household table typically requires minimal processing compared to the others, however a few modifications are still required:

- **Income:** Household income is often reported in a categorical variable in the survey and through the SPA tool. ActivitySim produces and expects actual dollar amounts for income. Income values are randomly generated by sampling from a uniform distribution of the household's income category. Households that were missing income values are assigned a randomly selected income drawn from the distribution of income values in the survey.

- **Household Size:** The household size variable needs to match the actual number of persons in the person file for that household. If persons are removed in subsequent processing steps, this field needs to be updated.

Persons

Review of the *annotate_persons.csv* config file should be performed prior to processing person level data to understand how student and employment categories are determined and how person type is calculated. When using a synthetic population generated from census targets, people are employed when the employment status code (*ESR*) is equal to 1. Part-time and full-time work is then determined by the usual hours worked per week (*WKHP*) and the number of weeks worked (*WKW*) in the past year. Student status is determined by the person's age and grade level attending (*SCHG*) code.

Certain person and tour level models only run for people that fall within a certain employment, student, or person type category. Thus, when fixing person type issues, special care needs to be taken to ensure the correct variables are being changed in the survey processor so that the persons are annotated correctly in ActivitySim.

- **Age:** In the Household Travel Survey (HTS), there are people who decline to answer the age question. Since student status, employment, and person type each have an "age < some number" filter, these missing ages need to be fixed. If a person has a valid school TAZ and a student category less than high school, the assumption used was to set their age 12 to represent non-driving students. Otherwise, they were treated as adults and their age is set to 30.
- **School Zone:** Not all students have a valid school zone. Either the student does not provide that information in the survey or the school zone they reported has no actual enrollment in the landuse file. If a school zone was not reported, and a school trip was made, that school trip location is used as their school zone. If a school zone had no enrollment in the land use, ActivitySim would crash in estimation mode because the size term used to calculate the utility is zero. These people have their school zones replaced with the closest zone (by TAZ/MAZ centroid distance) with the appropriate level of enrollment. Typically, surveys can only get ~70% of students to have a valid school zone even after the corrections.
- **Workplace Zone:** Like school, many people who report as workers do not report a workplace zone. If these people made a work tour, the first work tour destination is selected as their workplace. If a region contains zones without employment, additional logic would be needed to ensure a valid workplace zone like what was done for school. This can be done by industry or occupation segmentation if deemed necessary.

If there was no way to infer the school or workplace zones, their values were set to -1 which is ignored in ActivitySim destination choice estimation. Ignoring these values is only acceptable if that person does not make a work or school tour or else the downstream models will crash with an invalid destination.

- **Person Type:** If a person is not labeled as a worker or a student, but has a work or school tour, mandatory tour frequency and scheduling models will crash because work and school location choice does not get run for those people and no tour destination is set. Everyone who makes a work or school tour needs to be labeled as a worker or student. To fix this, the person types are changed for people who make a work or school tour.

  Persons who do not report being workers but take work trips have their *pemploy* status set to 2 (part-time worker). This flags them as employed in the annotate persons step. ActivitySim does not allow full-time workers to go to school. If a person takes a work tour and a school tour, the number of working hours and the work hours per week are decreased to fall below the part time threshold so that both school and work tours can be performed.

  There are also often people who report as non-student workers but perform a school tour and not a work tour. These people are changed from workers to students. Similarly, people who report as non-worker students but only make a work tour and no school tour are changed from students to workers.

Tours

In addition to determining correct school and workplace zones, tour level processing constitutes the bulk of the survey processing code. ActivitySim requires that all tours happen within a 24-hour day and that the number of tours match acceptable tour patterns.

- **Tour Purpose:** ActivitySim configuration files specify the types of tours allowed. A map between SPA output tour purposes and ActivitySim purposes must be created. Most of the categories map one-to-one with ActvitySim purposes. ActivitySim does not have a university tour type, so university tour types were changed to school. Another common tour purpose mapping is to code work-related travel as other maintenance. Loop tours are dropped from the dataset since ActivitySim does not model them.
- **Tour Category:** ActivitySim expects tours to be categorized into mandatory, non-mandatory, joint, and at-work subtours. Only tours that contain multiple members from the household for the entire tour are considered joint tours in ActivitySim.
- **Tour Type:** Atwork subtours contain an additional tour type variable into the ActivitySim categories *business*, *eat*, and *maint*. Work tours are mapped to *business*, eat out tours are mapped to *eat*, and all other atwork subtour purposes are mapped to *maint*.

- **Tour Mode:** Tour modes need to match the specified tour modes in the ActivitySim configuration files.
- **Tour Locations:** If tour destinations are not within the model region or are not reported, those tours do not have a valid start or end zone. These tours are removed.
- **Tour Times:** The *tour_departure_and_duration_alternatives.csv* config file contains the allowed tour start and end times. Time bins from a full 24-hour day are broken down into 48 half-hour time bins starting at 3:00 A.M. Tours were removed if it did not fit this specification, such as tours that take place overnight or do not complete during that 24-hour period.
- **Tour Frequencies:** There are configuration files for each of the tour categories that specify the allowed sets of tour frequencies. ActivitySim's default setting is that a person can only take two mandatory tours with the following possible combinations: one work, two work, one school, two school, or one work and one school. A much larger set of possible alternatives exists for the non-mandatory purposes (see the config file non_mandatory_tour_frequency_alternatives.csv for a list of all allowed possibilities).

  Code was developed to count the number of tours for each person and each tour category and summarize them up in such a way that matches the ActivitySim tour frequency alternatives files. Tours are then removed if a certain tour exists outside the allowed tour frequencies. For example, if a person were to take 3 eat out tours, but the specification only allows for up to two, then the third eat out tour is removed. Tours are numbered starting at the beginning of the day and the first tours are the ones selected. A slight bias may have been introduced from this sampling method, but previous comparisons between the estimation results and survey data that did not remove these tours showed no significant difference.

If a tour was removed for any of the above reasons, all subtours for that parent tour were also removed.

Joint Tours

ActivitySim lists joint tours only once in the tours file and instead lists the participants on each joint tour in the joint tour participants file. Additionally, all joint tours in ActivitySim are fully joint tours, meaning that all persons on the tour follow the same trip patterns and modes. Besides these formatting requirements and performing the transferrable checks on all tours including locations and start/end times, there are a few additional joint tour checks that need to be made.

- **Joint Tour Type**: ActivitySim tour types follow the same purposes as individual non-mandatory tours except there is no escorting purpose. Joint escort tours are changed to other maintenance instead.

- **Joint Tour Frequency:** Similar to non-joint tours, joint tours have restrictions on frequency. The procedure for removing joint tours that do not fall in the frequency alternatives is the same as for non-joint tours.
- **Joint Tour Participants:** Each joint tour needs to have an adult or child on the tour to determine joint tour composition. This is achieved by ensuring only fully joint tours from the SPA output are treated as joint tours. Each fully joint tours has participants listed in the joint tour participants file.

Trips

Trips follow many of the same processing rules as tours with some exceptions around stop frequency.

- **Trip Destination:** If trip destinations are not within the model region or are not reported, those trips do not have a valid start or end zone. These trips are removed.
- **Trip Purpose:** ActivitySim configuration files specify the types of trips allowed. A map between SPA output trip purposes and ActivitySim purposes must be created. Most of the categories map one-to-one with ActvitySim purposes, just like tours.
- **Trip Departure Time:** The departure time of the trip is given by bin number where the bins start at 3:00 AM and are in increments of half-hour. The time bin structure for trips is the same as for tours.
- **Stop Frequency:** Technically a tour-level variable, stop frequency is the number of stops on the tour in the outbound and inbound directions. The default number of total stops is three for inbound and outbound directions separately. Additional trips beyond the allowed number of stops are removed. Removed trips are those that take place between the maximum number of stops and the half-tour destination.
- **Trip Mode:** Trip modes need to match the modes specified in the ActivitySim configuration files.

Trips that take place on tours that were removed in the previous tour cleaning steps need to also be removed since trips must belong to a corresponding tour. Generally, the trip cleaning process removes about 30% of the available trip records in the raw survey data.

Infer.py

Infer.py is a python script that gets run on the cleaned survey data and adds ActivitySim specific variables to the data files before the data gets run in ActivitySim’s estimation mode. This step is necessary to ensure the tour and trip records created internally in ActivitySim match the corresponding tours and trips as well as ensuring the proper observed survey values are used for all downstream models.

These tour and trip IDs are how ActivitySim matches the observed survey tours with the corresponding tour/trip generated by ActivitySim. Thus, they are an essential part of the data preparation process. Part of the infer.py module includes checking these tour frequencies and assigning ActivitySim tour ids. If there are tour or trip combinations that are not allowed, they will not be assigned an ID and the estimation process will fail.

The other main function of the infer.py script is to append model alternatives to the input survey data. For example, the output of ActivitySim's tour scheduling model is an alternative number that corresponds to a tour start, end, and duration. Rather than expecting the user to assign the alternative number to each survey record, the infer.py script will determine the alternative number based on the tour time-of-day information and the ActivitySim configuration files for that region. This process occurs for many models, including non-mandatory tour frequency, tour scheduling, school escorting, joint tour frequency, atwork subtour frequency, joint tour composition, etc.

The infer.py module takes the human-readable survey data values and derives the necessary codes that ActivitySim uses under-the-hood. The output of the infer.py script is the same household, person, tour, and trip files, but with appended columns containing the required information for ActivitySim's estimation mode, particularly the tour and trip IDs and the chosen alternative numbers for specific models.

Creating Estimation Data Bundles

Estimation Data Bundles (EDBs) are created by running ActivitySim in estimation mode and are used as input into Larch for estimating model coefficients. The process of running ActivitySim in estimation mode is much like running ActivitySim in non-estimation mode with the exception of needing to provide additional estimation config files. An example run command would look like the following:

activitysim run -c configs_estimation -c configs -d data -o output

where the run directory would contain a *configs_estimation* folder with the estimation and run setting yaml files, a normal *config* folder with the typical model configs used in non-estimation mode, a *data* folder containing the override tables from the infer.py module, land use, and skims, and an *output* folder where the estimation data bundles would be written.

The *estimation.yaml* file *is located in the configs_estimation* directory This file contains a list for all the models for which estimation data bundles are written. At the time of this writing, multi-threading is not supported for ActivitySim's estimation mode. Thus, run times for estimation are often longer than running ActivitySim in production mode. Run time for full data samples take roughly 8 hours, but this can vary greatly depending on the zone system used and the size of the survey data.

The cleaning process can be time consuming considering a few fringe cases can cause estimation mode to crash. This means that while using a small sample is useful to test configurations and setup, problems are often not found until running the entire sample.

Debugging typically involves running the entire sample until an error arises. The household(s) that crash can then be traced in a smaller sample to uncover the error. Typically, the error is a result of unclean survey data. The full survey sample is run again after fixing the issue and tested with the traced household with the test sample.

However, the full survey sample must be run from the start since the input data is changing. This means that ActivitySim's restart functionality cannot be fully taken advantage of for most problems. Without multi-threading and hours long run times to get to later models, this process can be slow. Having the necessary cleaning outlined in this design should reduce the time it takes to get estimation data bundles written out for different regions.

Estimation Data Bundles are the output created by ActivitySim's estimation mode and contain everything needed to estimate a model in Larch. They consist of the following files:

- **Model Settings:** The model settings are specified in a yaml file and detail model specific settings. For example, the tour mode choice nesting structure would be listed in the *tour_mode_choice.yaml* model setting file.
- **Model Specification:** The model specification csv file specifies the utility structure. Each row in this file contains a term in the utility specification. This file is the main file that will be modified when trying new utility equations in model estimation to obtain the best model with the most explanatory power.
- **Model Coefficients:** The model coefficients csv file lists the name of the coefficients used in the model specification as well as their starting value and whether that value should be fixed in the model estimation procedure. Fixed coefficients often correspond to a reference alternative or an availability condition.
- **Model Choosers:** The model choosers csv file contains a row for each record in the survey that is going to make a model choice. For example, the chooser file would contain a row for each household for the auto ownership estimation data bundle and the chooser file for tour mode choice would list a tour in each row. The columns of the model chooser data correspond to the rows in the model specification, i.e. each utility expression is a column in the chooser data.
- **Model Alternatives:** The model alternatives csv file is only used for the "interaction simulate" models in ActivitySim that have many different alternatives. These models include the destination choice models, the scheduling models, and school escorting. This file would have the alternative listed as the columns and the rows would be a list of the utility expressions evaluated for the alternative for each chooser. So, if there were 10

tours, 5 utility expression terms, and 100 destination alternatives, the total number of rows would be 10*5 = 50 and the number of columns would be 100. As such, the size of this table can grow to be quite large (up to 10 GBs) for some models.

Estimation Data Bundles are written out for each individual model for ActivitySim as well as for each model segmentation. For example, non-mandatory tour scheduling models are often broken down by purpose. Each tour purpose would have its own EDB and contain just the observed survey records for that purpose so they can be estimated independently of each other.

Included in the model choosers table of the EDB are two columns: *override_choice* and *model_choice*. The *model_choice* column contains what the model would have chosen with the provided specification file and the *override_choice* column contains the observed value from the survey data. Thus, comparisons of how well a model estimation fits the survey data can be performed by comparing these two columns and can help validate the estimated model.

Model Estimation in Larch

The actual estimation of model coefficients is performed outside of the main ActivitySim application through the Larch python software package which is built on top of the SciPy python package. The user interacts with Larch through a Jupyter estimation notebook. The estimation notebooks load the EDB into Larch, allow the user to review the loaded data and specification, run the model estimation, and observe the estimated model fit and coefficient values and associated statistics. Many example estimation notebooks exist on the ActivitySim repository[3].

The final report out of the estimation notebooks is written to an easy-to-digest excel spreadsheet that lists the coefficient names, estimated values, standard error, t-statistic, and significance. These estimated coefficient values are then transferred back into the ActivitySim configuration files for use in model production.

The process of getting a valid model estimation specification is often more art than science. The user must ensure that the specification is functional, including proper reference coefficients, bug-free expressions, and utility equations that are not over-specified. Users change the utility equation by editing the model specification csv file in the estimation data bundle that is loaded into Larch.

As mentioned in the introduction section, getting the best estimated model is often an iterative procedure where the user will modify the specification, run ActivitySim to generate the EDB, and see how the estimation responds. The goal is to produce model estimations that predict the

[3] https://github.com/ActivitySim/activitysim/tree/main/activitysim/examples/example_estimation/notebooks

observed distributions as well as possible while producing coefficient values that can be explained, are statistically significant, and are temporally stable for forecasting.

Handling Multiple Survey Days

ActivitySim models a single 24-hour day, but many surveys now will record a single household for multiple days. This means that the survey data processing and estimation needs to be very careful in how to handle households that record data for more than one day.

Households that are surveyed for multiple days must be copied to produce unique household records when input into ActivitySim's estimation mode. This procedure is typically performed by duplicating rows in the survey household table by the number of days surveyed and assigning a new household id column. The same goes for persons. The new household and person ids are then propagated forward into the person, tour, and trip files. The result of this process is that each survey day will show up in the ActivitySim files as a unique household that performs one day's worth of travel. Flags are added to the data to keep track of whether a household or person record was duplicated in this process. Tours and trips are day-level attributes and are assigned to the appropriate household and person.

While the procedure of copying household and person attributes to create self-contained travel days solves the problem of running the survey data through ActivitySim's estimation mode, further care needs to be taken in the model estimation. Because records are copied, those copied records need to be removed from the estimation data set before loading the data into Larch. Improper estimation results would arise for instance when estimating mobility models, because household records that were surveyed using the smartphone app might appear up to seven times (the week long duration of the survey), whereas households filing out a single day of the online survey appear only once. The user would use the flags created in the data duplication procedure to remove the duplicated records as appropriate.

There can be additional nuance to this procedure for some models that are predicting day-level attributes but at a household or person level. For example, the CDAP model is coordinated across all household members, but if children are only surveyed on a single day, only that day can be used to determine CDAP, even if the household has records for adults on all days. Similarly, the school escorting and joint tour models should only be estimated with household-days where the whole household participated in the survey.

3.0 ABM3 MODEL SYSTEM DESCRIPTION

The ABM3 model system is based on the ActivitySim platform. The current version of ActivitySim follows the SANDAG resident model structure closely since they are both based on the Coordinated Travel Regional Activity-based Modeling Platform (CT-RAMP). Figure 2 shows the current ActivitySim model design for residents. In order to understand the flow chart, some definitions are required. These are described in more detail below.

- *Tour*: A sequence of trips that start and end at an anchor location. In ActivitySim, anchors are home or work.
- *Primary destination*: The “main” activity of the tour; this activity determines the tour purpose. It also divides the tour into two "legs"; the sequence of trips from the anchor location to the primary destination is the outbound leg, and the sequence of trips from the primary destination back to the anchor location is the inbound or return leg.
- *Mandatory activity:* Work or school
- *Non-mandatory activity*: Any out of home activity that is not work or school, including maintenance activities such as shopping as well as discretionary activities such as out-of-home recreation and eating out.
- *Fully joint tour*: A tour in which two or more household members travel together to all out-of-home activity locations and return home together. In other words, no household member is picked-up or dropped-off en route.
- *Intermediate stop*: An out-of-home activity location on the tour other than the anchor location or the primary destination. Intermediate stops are made on the way from the anchor location to the primary destination (outbound) or on the way from the primary destination back to the anchor location (inbound).
- *Tour mode*: The “main mode” or “preferred mode” of the tour. This is an abstract concept used categorize the tour with respect to accessibility and constrain the availability of modes for trips on the tour to ensure some consistency of modes used for each trip.

The first model in the sequence is disaggregate accessibilities. This is a recent addition to ActivitySim in which the tour destination choice model is run for a prototypical sample population covering key market segments and destination choice logsums from the model are written out for each tour in the population. These destination choice logsums are then merged with the actual synthetic population and used as accessibility variables in downstream models such as auto ownership, coordinated daily activity patterns, and tour frequency. are mandatory location choice; this model is run for all workers and students regardless of whether they attend work or school on the simulated day. Next a set of long-term and mobility models are run. The first model in the sequence predicts whether an autonomous vehicle is owned by the household. This model conditions the next model, which predicts the number of autos owned. If an autonomous vehicle is owned, multiple cars are less likely. Next, the mandatory (work and school) location choice models are run. The work location choice models includes a model to predict whether the worker has a usual out-of-home work location or exclusively works from home. If the worker chooses to work from home, they will not generate a work tour. An external worker identification model determines whether each worker with an out-of-home workplace location works within the region or external to the region. If they work external to the region, the external station is identified. Any primary destination of any work tours generated by the worker will be the external station chosen by this model. A work location choice model predicts the internal work location of each internal worker, and a school location choice model predicts the school location of each student.

Next, a set of models predicts whether workers and students have subsidized transit fares and if so, the percent of transit fare that is subsidized, and whether each person in the household owns a transit pass. A vehicle type choice model then runs, which predicts the body type, fuel type, and age of each vehicle owned by the household; this model was extended to predict whether each vehicle is autonomous, conditioned by the autonomous vehicle ownership model. Next, we predict whether each household has access to a vehicle transponder which can be used for managed lane use. We assume that all vehicles built after a certain year (configurable by the user) are equipped with transponders. Next we predict whether each worker has subsidized parking available at work. Finally we predict the telecommute frequency of each worker, which affects downstream models including the daily activity pattern model, the non-mandatory tour frequency model, and stop frequency models.

Next the daily and tour level models are run. The first daily model is the daily activity pattern model is run, which predicts the general activity pattern type for every household member. Then Mandatory tours are generated for workers and students, the tours are scheduled (their location is already predicted by the work/school location choice model), a vehicle availability model is run that predicts which household vehicle would be used for the tour, and the tour mode is chosen. After mandatory tours are generated, a school pickup/dropoff model forms half-tours where children are dropped off and/or picked up at school. The model assigns chaperones to drive or ride with children, groups children together into “bundles” for ride-sharing, and assigns the chaperone task to either a generated work tour or generates a new tour for the purpose of ridesharing. Fully joint tours – tours where two or more household members travel together for the entire tour - are generated at a household level, their composition is predicted (adults, children or both), the participants are determined, the vehicle availability model is run, and a tour mode is chosen. The primary destination of fully joint tours is predicted, the tours are scheduled, the vehicle availability model is run, and a tour mode is chosen. Next, non-mandatory tours are generated, their primary destination is chosen, they are scheduled, the vehicle availability model is run, and a tour mode is chosen for each. At-work subtours are tours that start and end at the workplace. These are generated, scheduled (with constraints that the start and end times must nest within the start and end time of the parent work tour), a primary destination is selected, the vehicle availability model is run, and a tour mode is chosen.

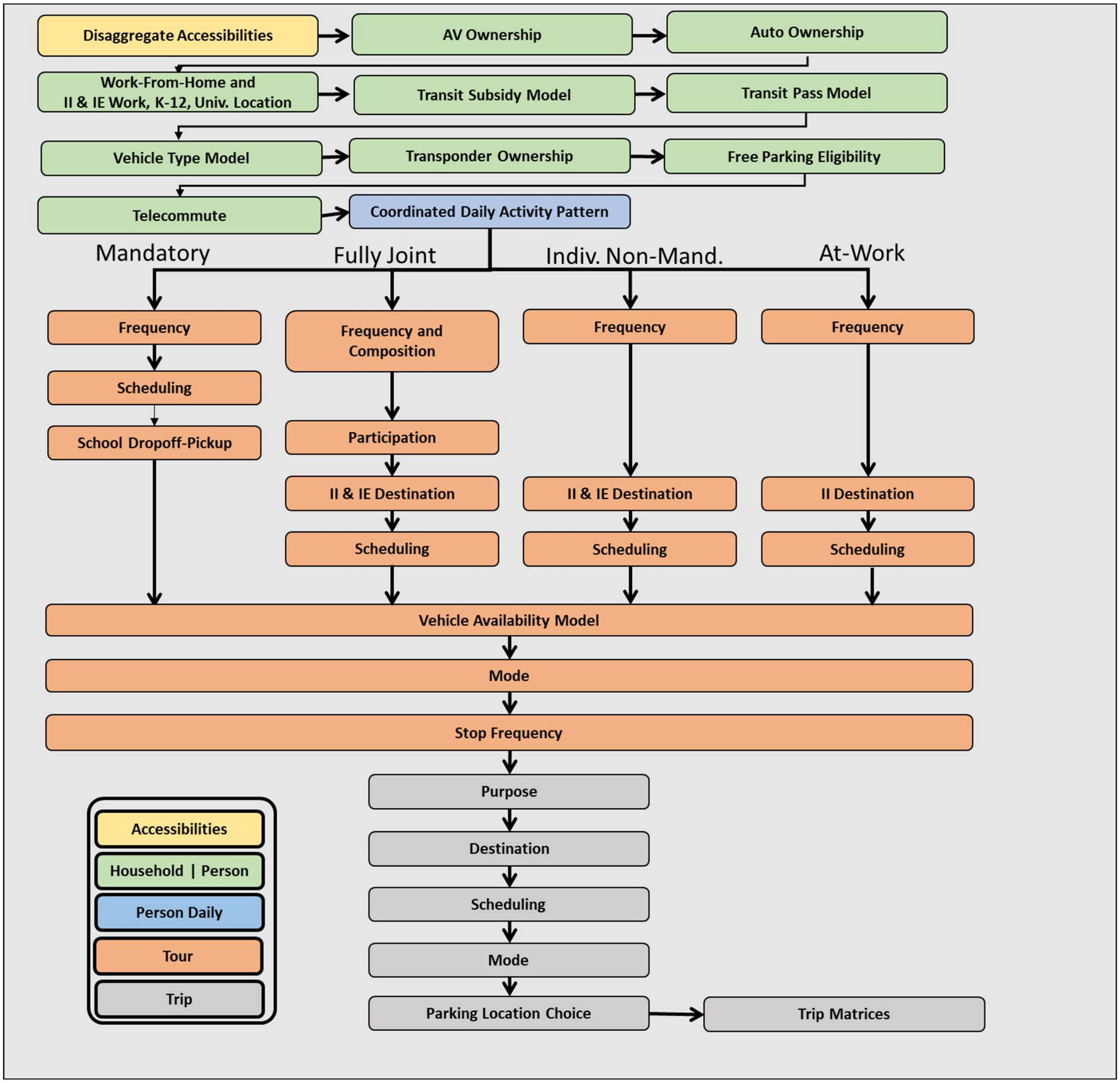
At this point, all tours are generated, scheduled, have a primary destination, and a selected tour mode. The next set of models fills in details about the tours - number of intermediate stops, location of each stop, the departure time of each stop, and the mode of each trip on the tour. Finally, the parking location of each auto trip to the central business district (CBD) is determined.

After the model is run, the output files listed above are created. The trip lists are then summarized into origin-destination matrices by time period and vehicle class or transit mode and assigned to the transport network. Skims are created based on congested times, and the model system is iterated multiple times until either some convergence threshold is attained, or a predetermined number of iterations is reached.

ActivitySim is used to represent all internal travel made by residents of the SANDAG region (modeled area). The decision-makers in the model system include both persons and households. These decision-makers are created (synthesized) for each simulation year and land-use scenario, based on Census data and forecasted distributions of households and persons by key socio-economic categories. The decision-makers are used in the subsequent discrete-choice models in a microsimulation framework where a single alternative is selected from a list of available alternatives according to a probability distribution. The probability distribution is generated from a logit model which considers the attributes of the decision-maker and the attributes of the various alternatives. The application paradigm is referred to as Monte Carlo simulation, since a random number draw is used to select an alternative from the probability distribution. The decision-making unit is an important element of model estimation and implementation and is explicitly identified for each model specified in the following sections.

A key advantage of using the micro-simulation approach is that there are essentially no computational constraints on the number of explanatory variables that can be included in a model specification. However, even with this flexibility, the model system will include some segmentation of decision-makers. Segmentation is a useful tool to both structure models (for example, each person type segment could have their own model for certain choices) and to characterize person roles within a household. Segments can be created for persons as well as households.

FIGURE 2: ACTIVITYSIM MODEL COMPONENTS



4.0 DISAGGREGATE ACCESSIBILITIES

In ABM3, a new disaggregate accessibilities component was added to the model system. These new disaggregate accessibilities consistent with the actual ActivitySim destination and mode choice models, and are used by ActivitySim model components requiring destination choice logsums. They can also be post-processed to create benefit cost assessments (social welfare summaries) based on logsums.

The accessibilities are a replacement for the aggregate origin-based (e.g. destination choice) accessibilities created by CT-RAMP and which are used in previous versions of the SANDAG ABM. We describe the previous method below and then discuss the new disaggregate accessibility calculations.

4.1 OVERVIEW OF ACCESSIBILITIES

Accessibilities used in activity-based travel models can be broadly categorized as follows:

- *Origin-destination accessibilities*. These accessibilities are calculated between an origin and a destination Transportation Analysis Zone (TAZ) or Micro-Analysis Zone (MAZ). These accessibilities can take the form of travel time, generalized cost, or a mode choice logsum. They are used in models where both the tour or trip origin and destination are known. For example, the mode choice logsum for work location choice can be used in auto ownership, because disaggregate accessibilities is run before auto ownership. Mode choice logsums are also used in destination choice models; often this requires sampling a subset of zones as potential destinations since calculating a mode choice logsum for every tour or trip to every potential destination zone is typically computationally prohibitive. The mode choice probability equation is shown in Equation 1, where $P_{a*,i}$ is the probability of choosing mode *i* for decision-maker *a**, and *V* is the utility for mode *i*. Equation 2 shows the mode choice logsum ML_{a*} for decision-maker *a*,* calculated as the natural log of the denominator of the mode choice model.
- *Origin-based accessibilities*. These accessibilities are for origin TAZ or MAZs. They represent the accessibility of the origin zone to all potential destinations. These accessibilities can be defined as the total employment or total retail employment accessible within a certain range of travel time for a certain mode; for example, total retail employment within 30 minutes of peak transit time. The logsum of the destination choice model is a more sophisticated treatment of origin-based accessibility due to its ability to capture accessibility for a certain tour or trip purpose, considering all modes of transportation and variables associated with mode choice such as cost and socio-economic variables. Equation 3 shows the destination choice logsum equation. S, is the

size term of zone *j*, γ is the coefficient on mode choice logsum, and ML_{a*} is the mode choice logsum term for decision-maker a* (*i* and *j* zones are implied.

EQUATION 1: MODE CHOICE EQUATION

$$P_{a*,i} = \frac{e^{V_{a*,i}}}{\sum_{a\in A} e^{V_{a,i}}}$$

EQUATION 2: MODE CHOICE LOGSUM

$$ML_{a*} = \ln\left(\sum_{a\in A} e^{V_{a,i}}\right)$$

EQUATION 3: DESTINATION CHOICE EQUATION

$$A_i = ln\left[\sum_{j=1}^{I} S_j \; x \; \exp(\gamma \; ML_{a*})\right]$$

The previous version of the SANDAG ABM use destination-choice based logsums calculated using simplified mode and destination choice models that do not utilize the same model specifications as the actual mode and destination choice models used in the disaggregate activity-based model. The previous model calculates 50 different destination choice logsums in a pre-processor; each logsum is a combination of purpose, auto sufficiency and mode. Because these logsums are calculated with simplified models (not the actual destination or mode choice models used in CT-RAMP), the utility equations are maintained in separate set of Utility Expression Files and there are separate Java classes to create them. The utility expressions are simplified versions of the real models; they do not include all of the alternatives in the real models, or the actual utilities used in the actual models. The alternative-specific constants in the models are approximated. Changes (new modes such as ride-hailing, policies such as VMT taxes, or calibration of constant terms to new data) made to the actual tour or trip mode choice models do not affect the origin-based accessibilities. The new disaggregate logsums address these shortcomings.

4.2 POPULATION SEGMENTS AND TOURS FOR DISAGGREGATE ACCESSIBILITIES

In order to create a set of disaggregate accessibilities using the ActivitySim destination and mode choice models, a 'fake' or 'prototypical' synthetic population must be created covering all market segments of interest, and a set of tours must be defined for each household and person covering all tour purposes of interest. This section describes the household, person, and tour segmentation that is used to create the accessibilities.

Table 1 shows all household variables in the synthetic population household file. Each variable name is shown as a row in the table, along with a brief description of the variable. There are three variable "types" indicated in the table, as follows:

- Constant: The variable will not change across households in the prototype population.
- Sequential: Variables such as IDs whose values range from 1 to n
- Controlled: Variables such as household income and auto ownership, whose values will be controlled by the user

Household income and household vehicles are systematically varied in the synthetic population, with three levels of income and three levels of auto ownership (0 autos, autos<workers, autos>=workers) so there are a total of 3x3 or 9 total household segments in the synthetic population.

Table 2 shows the person variables in the synthetic population. Each household has two persons; a full-time working female age 35 and a non-working male age 55. The full-time worker has a Mandatory activity pattern (1 work tour) and the non-working adult has a Non-mandatory activity pattern (2 non-mandatory tours). The default values for tour start and end times and auto operating costs are used in the mode choice logsum calculations. (This would correspond to the non-household vehicle option in the context of the vehicle allocation model.) The household composition and person characteristics were selected to ensure that both mandatory and non-mandatory tours would be generated. Household member ages were selected to ensure that mobility was generally unrestricted but the exact ages and other characteristics are arbitrary.

We generate the following tours for destination choice logsums:

- Person 1: Work tour
- Person 2: Other Maintenance tour (due to strong correlations of logsums between different types of maintenance activities with Other Maintenance purpose.
- Person 2: Other Discretionary tour (due to strong correlations of logsums between different types of discretionary activities with Other Discretionary purpose.

These setting are controlled by the disaggregate_accessibilities.yaml file. Because there are over 22,000 MGRAs (MAZs) in the SANDAG model, it is computationally prohibitive to populate each MAZ with households covering all market segments. Therefore we use a k-means cluster-based sampling method to sample MAZs to populate. We select 5000 zones using weighted k-means clustering where each MAZ is weighted by total population (Figure 3). We run the sample through ActivitySim's destination and mode choice models and save the destination choice logsums for each tour in the output choice table (Figure 4 and Figure 5). We then merge those logsums with the actual full synthetic population by finding the closest MAZ to the synthetic household with a prototypical household, and we merge all three logsums by auto sufficiency from the prototype household with the same range of household income as the synthetic household. These logsums by auto sufficiency are used in the auto ownership model; subsequent models use one of the logsums depending on the chosen number of autos.

Most model components that used to use the aggregate logsums were re-estimated; when they were the disaggregate logsums were used in place of the aggregate logsums. These components are described below. Model components that use the old logsums which were not re-estimated, such as the joint tour frequency and participation model, were modified such that the old logsums were replaced with the closest relevant new logsum. In such cases, the coefficient on the logsum was modified to reflect the difference in scale between the old logsum and the new logsum.

FIGURE 3: ZONES SELECTED BY WEIGHTED K-MEANS CLUSTERING

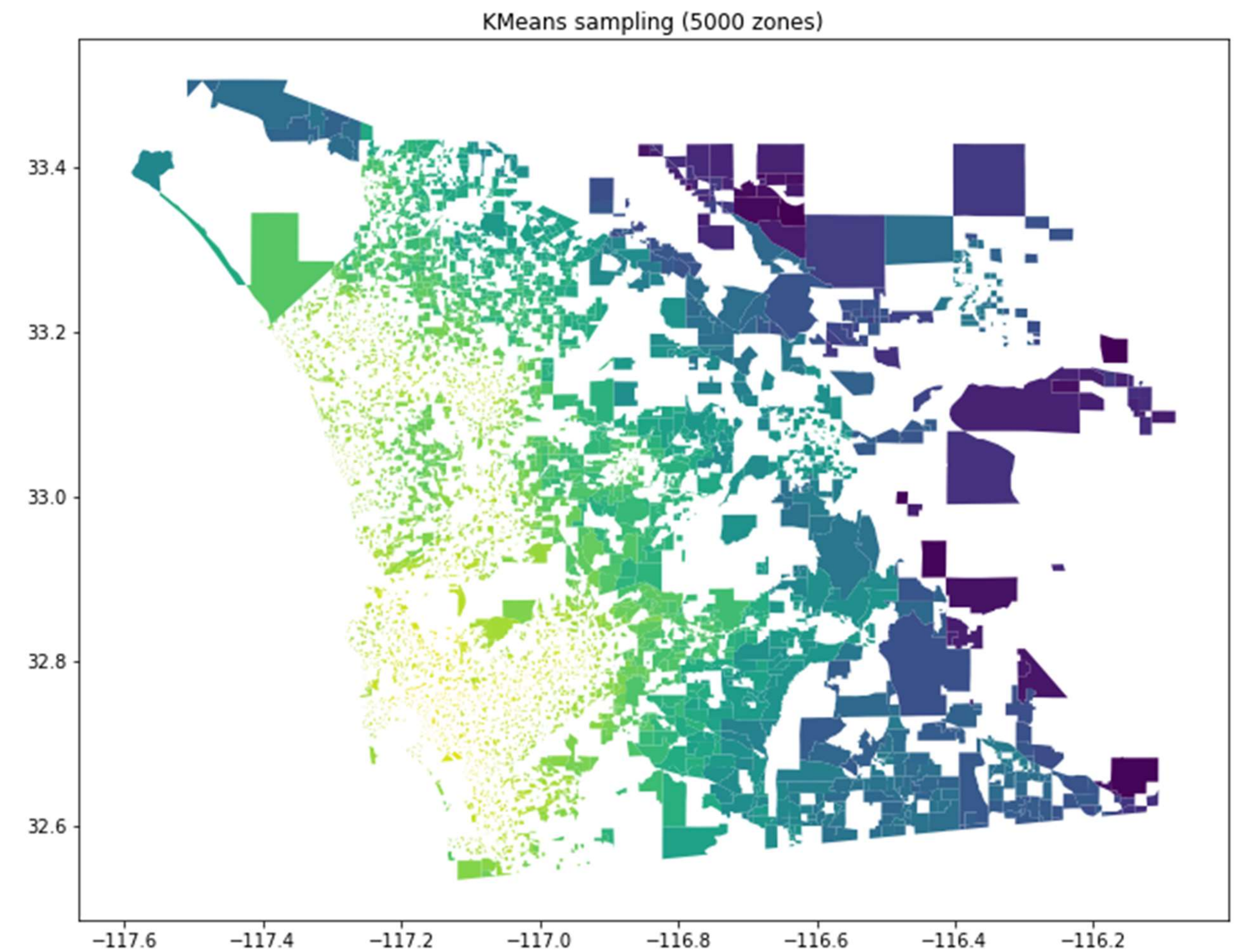


FIGURE 4: AVERAGE ACCESSIBILITIES BY AUTO OWNERSHIP AND TOUR PURPOSE

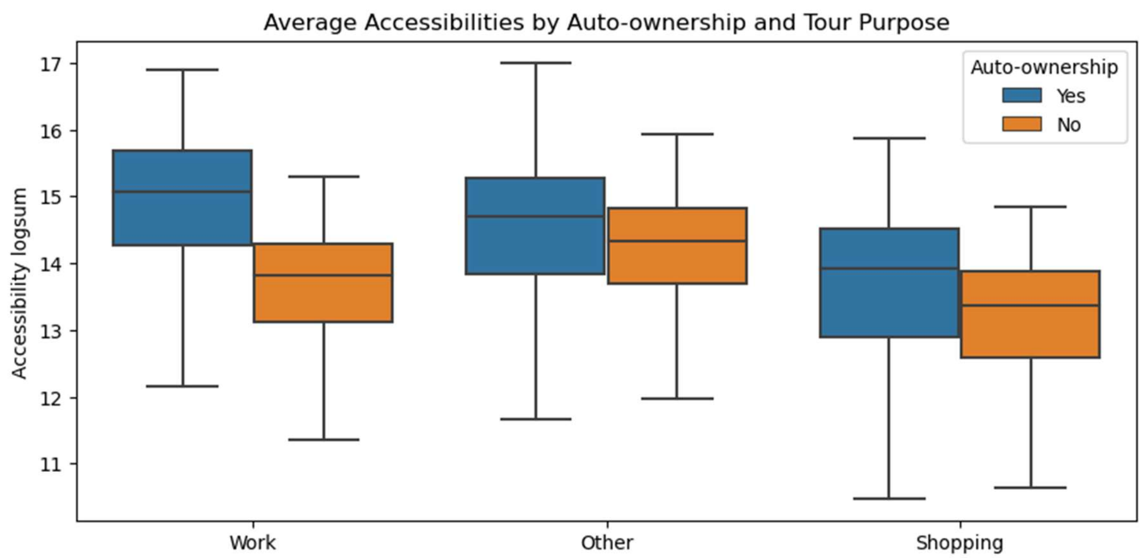


FIGURE 5: AVERAGE ACCESSIBILITIES BY HOUSEHOLD INCOME AND TOUR PURPOSE

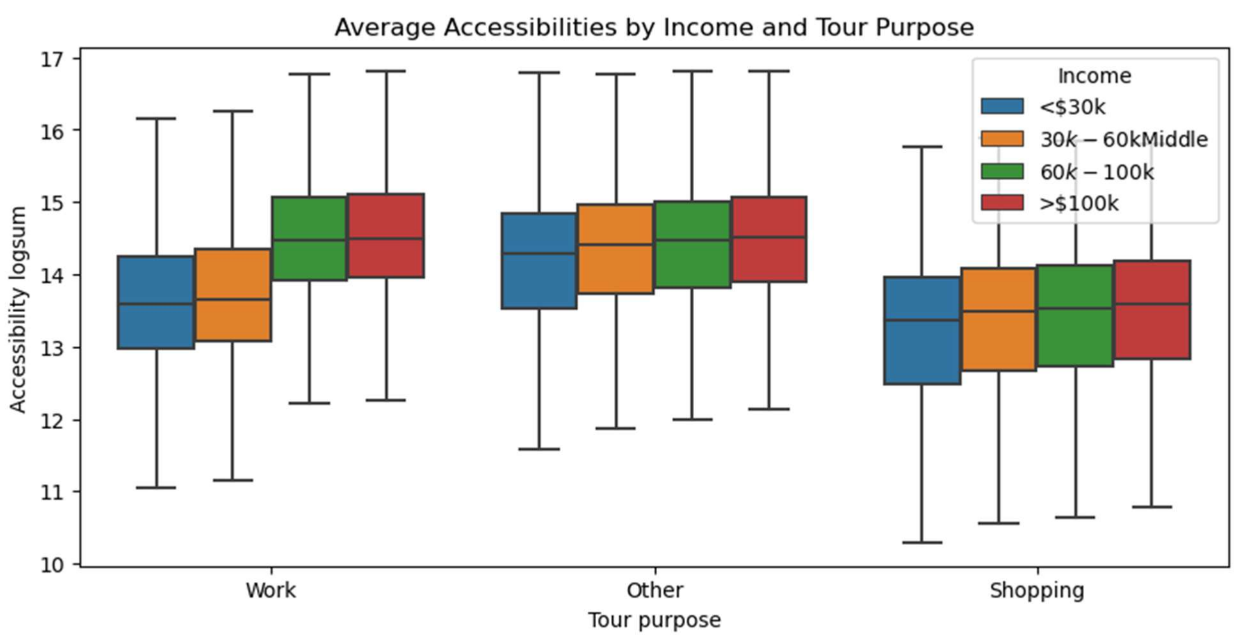


TABLE 1: HOUSEHOLD VARIABLES

Variable Name	Description	Type	Value(s)
hhid	Unique Household ID	Sequential	1 to n where n is total number of households in file
household_serial_no	Household serial number	Sequential	1 to n where n is total number of households in file
taz	TAZ of household	Controlled	The set of TAZs to populate
maz	MAZ of household	Controlled	The set of MAZs to populate (if applicable)
hinccat1	Household income category:	Controlled	1 (<$30k), 2 ($30k-$60k), 3 ($60-100k), 4 ($100-$150k), corresponding to hinc values below
hinc	Household income	Controlled	$14k, $45k, $67k, $120k (using 10th, 25th, 50th, and 75th percentiles)
hworkers	Number of workers in household	Constant	1
veh	Number of vehicles in household	Controlled	0, 1, 2
persons	Number of persons in household	Constant	2
hht	Household/family type:	Constant	1 = Family household: married-couple (2035 mode)

bldgsz	Building size - Number of Units in Structure & Quality:	Constant	2 = One-family house detached (2035 mode)

TABLE 2: PERSON VARIABLES

Column Name	Description	Unique value for Person 1	Unique value for Person 2	Other
hhid	Household ID			Set according to hhid in household file
perid	Person ID			Sequential 1 through n where n is total number of persons in file
pnum	Person Number	1	2	
age	Age of person	35	55	
sex	Gender of person	2 (female)	1 (male)	
military	Military status			4 (No)
pemploy	Employment status	1 (employed full-time)	3 (unemployed)	
pstudent	Student status			3 (not attending school)
ptype	Person type	1 (full-time worker)	4 (non-working adult)	

educ	Educational Attainment:			13 (Bachelors)
grade	Grade school attending			0 (not attending)
weeks	Weeks worked	1 (50 to 52 weeks)	0	
hours	Hours worked	35	0	
timeFactorWork	Work travel time factor			1 (mean)
timeFactorNonWork	Nonwork travel time factor			1 (mean)
DAP	Daily Activity Pattern	M (Mandatory)	N (Non-Mandatory)	

5.0 ESTIMATED MODELS

In ABM3 model development, the project team focused estimation activities on key models based on the following conditions:

- Travel behavior that demonstrated a significant behavioral shift from pre-COVID to current conditions. For example, the share and type of workers who work from home, telecommute frequency, and models describing the frequency of travel. While many aspects of travel experienced significant changes from pre-COVID conditions, we focus mostly on models upstream in the system whose effects cannot be accounted for simply by calibration of alternative-specific constants.
- New model components for which there were no existing parameters to calibrate. For example, the external worker identification model, the external worker destination choice model, and so on.
- Model components whose specification in CT-RAMP included variables that are no longer relevant or whose specification might significantly improve by replacing older variables with new previously unavailable variables such as industry and/or disaggregate accessibilities.

In this section we describe the estimation results of each re-estimated model component.

5.1 AUTO OWNERSHIP MODEL

The auto ownership model predicts the number of autos owned by the household. The choices are 0 autos, 1 auto, 2 autos, 3 autos, and 4+ autos. The model was estimated with 8,980 combined household records from 2016 and 2022 and is informed by variables that describe household attributes, the composition of the household, the accessibility of the household, and land-use effects.

The final coefficients are shown in Table 3 and t-statistics are shown in Table 4. The initial log likelihood is -16,090.00 and the final log-likelihood is -8,120.90. The rho-squared is 0.4953.

The initial CT-RAMP model included a set of variables that related household attributes to number of drivers in the household. This formulation was tested in estimation and resulted in many insignificant parameters. Therefore, the variables that were expressed as ratios to number of drivers were represented explicitly in the specification.

Below we summarize the estimated coefficients:

- The number of drivers in the household (identified as persons age 16 or higher) is positively correlated with autos owned. The model specification sets the base auto ownership category as the category equal to the number of drivers in the household, so the coefficients can be interpreted as the effect of drivers in the household on the likelihood of the household owning fewer cars than drivers or more cars than drivers. Although the coefficient for one-driver households owning 2 cars is insignificant, it was retained in the specification because the coefficient was reasonable in magnitude and correctly signed.

- The number of workers per household also has a positive effect on number of autos owned, except that 3+ worker households are less likely to own 1 car than zero cars. This could be due to correlations between households with multiple workers and other non-observed attributes such as lack of accumulated wealth that makes auto ownership difficult. However, there are also positive offsetting coefficients on owning 3 or more cars.
- The presence of children age less than 5 years old has a positive effect on owning 2 cars, indicating the mobility impacts of young children in the household. The presence of persons age 18-24 is positively correlated with owning 0 cars; this may be due to younger households who have not accumulated enough wealth to purchase an automobile. The presence of persons age 25-34 is negatively correlated with owning more than 2 cars, and the presence of persons age 80+ is negatively correlated with owning more than 1 car. Other age effects tested in the model were insignificant.
- Residential type other than single-family detached housing is negatively correlated with auto ownership. This may be due to limited auto parking at apartments and condos.
- Household income (in $2022) is positively correlated with auto ownership.
- Intersection density is positively correlated with owning 0 cars, reflecting the effect of more walkable environments on auto ownership. Intersection density is measured as the count of intersections with three or more legs within ½ mile of the microzone, divided by 520 (number of acres in a circle with a ½ mile radius).
- Disaggregate accessibilities, defined as tour destination choice logsums for non-mandatory tour purposes, are correlated with the number of autos owned. In other words, as the non-mandatory accessibility increases for a specific auto ownership category, the likelihood of owning cars in that category increases. The non-mandatory tour logsum is measured as the average of the tour destination choice logsum for shopping and for other discretionary tour purposes.
- In addition to the disaggregate accessibilities, aggregate accessibilities were also tested. Of these, the non-motorized accessibility was found to be significant and negatively correlated with autos owned. Non-motorized accessibility is measured as the distance-

weighted number of retail jobs within 3.0 miles of total round-trip distance (3.0 miles), as shown below.

$$NonMotorizedAccessibility_o = \sum_{od} RetailJobs_d * e^{-1.0*Distance_{od}}$$

- Population density (population in all microzones with ½ mile radius of the household microzone, divided by number of acres) and employment density (defined similarly) has an additional negative effect on the number of autos owned.
- Group quarters households are likely to own fewer cars. Note that there are few group quarters households in the observed dataset.
- A set of alternative-specific constants were estimated specifically for 2016 data. These constants are positive with respect to number of autos owned, reflecting generally lower auto ownership levels for 2022. This may be the result of survey bias so we suggest checking auto ownership model results against Census data for pre-pandemic model runs.

TABLE 3: AUTO OWNERSHIP MODEL COEFFICIENTS

VARIABLE DESCRIPTION	COEFFICIENT				
	0 autos	1 auto	2 autos	3 autos	4+ autos
1 driver	-3.40		-0.353	-1.65	-2.22
2 drivers	-5.64	-2.08		-1.11	-1.67
3 drivers	-5.66	-2.22	-0.428		-0.159
4+ drivers	-6.09	-3.25	-1.26	-0.736	
1 worker	-0.627			0.415	0.415
2 workers	-0.688		0.776	0.964	0.964
3+ workers	1.25			1.19	1.19
Has children 0-4			0.589		
Has persons age 18-24	0.303				
Has persons age 25-34				-0.205	-0.205
Has persons age 80+			-0.719	-0.973	-0.973
Not single-family detached	0.583		-0.654	-1.43	-1.86
Income < $15k	2.03		-1.26	-1.83	-2.74
Income $15k-$30k	1.12		-0.824	-1.24	-2.05
Income $30k-$60k			-0.285	-0.511	-0.903
Income $100k+	-0.902				
Intersection density	0.789				
Non-mandatory logsum	0.423	0.517	0.496	0.496	0.496
Non-motorized accessibility	0.305		-0.0515	-0.0686	-0.134
Population density	0.0160		-0.0123	-0.0157	-0.0157
Employment density	0.0142			-0.0109	-0.0168
Group quarters household	2.47				
2016 households	-0.818		0.256	0.433	0.529

TABLE 4: AUTO OWNERSHIP MODEL T-STATISTICS

VARIABLE DESCRIPTION	COEFFICIENT				
	0 autos	1 auto	2 autos	3 autos	4+ autos
1 driver	-5.13		-1.05	-4.37	-5.17
2 drivers	-7.85	-6.21		-6.98	-7.66
3 drivers	-7.05	-5.59	-2.21		-0.65
4+ drivers	-6.87	-6.70	-4.18	-2.73	
1 worker	-6.19		NA	4.21	4.21
2 workers	-2.51		10.13	7.60	7.60
3+ workers	2.32		NA	4.75	4.75
Has children 0-4			6.14		
Has persons age 18-24	1.57				
Has persons age 25-34				-2.08	-2.08
Has persons age 80+			-5.33	-4.90	-4.90
Not single-family detached	4.24		-9.34	-12.71	-9.99
Income < $15k	15.57		-8.91	-6.54	-5.10
Income $15k-$30k	8.82		-7.99	-7.41	-7.12
Income $30k-$60k			-3.34	-4.38	-5.67
Income $100k+	-3.03				
Intersection density	1.86				
Non-mandatory logsum	6.59	14.98	15.64	15.64	15.64
Non-motorized accessibility	4.95		-1.86	-1.92	-2.87
Population density	3.21		-3.02	-2.29	-2.29
Employment density	4.46			-1.87	-1.51
Group quarters household	4.98				
2016 households	-7.57		3.09	3.72	3.37

5.2 WORK FROM HOME MODEL

The work from home model predicts the propensity of an individual to work from home. The choices are work from home or not. The model was estimated using 9,020 individuals from the 2016 and 2022 survey data. 975 persons work from home. Duplicate person days were omitted in this model.

It is estimated using variables that describe the individual and household, as well as one land use variable.

The final coefficients and t-statistics are provided in Table 5. The initial log likelihood is -3663 and the final log-likelihood is -2546. The rho-squared is 0.3047.

The estimated coefficients used in the final model are summarized below:

- Access to workplaces: This variable is the destination choice logsum from the workplace location choice model, and is the only explanatory variable in the work away from home choice. The structure represents a nested model with work location choice nested under work from home choice. The original estimated coefficient was 0.05 and insignificant. Testing indicated that the coefficient was approximately 0.15 for 2016 data and insignificant for 2022 data. Therefore it was fixed at 0.1.
- Age coefficients: The age coefficients were broken into the following age groups: Age less than 35, age 35-44, age 45-54, age 55-64, age 65-79 and age 80 plus. Age less than 35 is used in this model as the reference group for the other binary indicators. Estimation results indicate that working from home is positively correlated with worker age.
- Income coefficients: The annual household income coefficients were broken into the following binary groupings: Income less than 15,000, income between 15,000- 24,999, income between 25,000 and 49,999, income between 50,000 and 99,999, income between 100,000 and 149,999, income between 150,000 and 250,000 and income greater than 250,000. Income between 50,000 and 100,000 was used as the reference group for this set of coefficients. Estimation results indicate that workers in low-income households and households with income ranging from 150k to 250k are more likely to work from home.
- Industry coefficients: The industry type was used in model estimation for the following categories: “Accommodations”, “Agriculture”, “Business Services”, “Construction”, “Education”, “Entertainment”, “Food Services”, “Government”, “Healthcare”, “Manufacturing”, “Management Services”, “Military” and all other as and additional “Other” category. The reference for this grouping of coefficients is the “Other” industry types. Estimation results indicate that workers in Business Services, Entertainment, and Management Services are more likely to work from home than workers in other industries. Workers in Accommodations, Construction, Education, Food Services, Government, Health, Manufacturing, and the Military are less likely to work from home than workers in other industries.
- Part time worker: The part time worker status was included in model estimation. Estimation results indicate that part-time workers are more likely to work from home.
- CBD indicator: If the worker’s household is in the CBD, they are more likely to work from home.

- Pre-COVID indicator: Workers in 2016 were less likely to work from home than workers in 2022.

The following variables were tested and found to be insignificant:

- Household composition including presence of non working adult in household, presence of preschool aged child or children, and presence of school aged child or children.
- Auto sufficiency (0 autos, autos less than adults, autos greater than adults)
- Non-motorized access to retail employment

TABLE 5: WORK FROM HOME MODEL COEFFICIENTS AND T-STATISTICS

VARIABLE	COEFFICIENT	T-STATISTIC
Accessibility to workplaces (work away from home alternative)	0.100	NA
Part-time worker	-1.45	-16.35
Worker Age		
Age 35-44	0.354	3.24
Age 45-54	0.468	4.24
Age 55-64	0.495	4.35
Age 65-79	1.24	9.42
Age 80 plus	0.771	1.53
Household Income		
Income < $15k	0.614	3.37
Income $150k-$250k	0.419	3.29
Employment Type		
Accommodations	-1.64	-3.51
Business Services	0.721	6.90
Construction	-0.738	-3.95
Education	-0.509	-3.34
Entertainment	1.03	6.43
Food Services	-1.20	-2.34
Government	-0.754	-3.22
Health	-0.442	-3.18
Manufacturing	-0.567	-2.58
Management Services	0.387	2.91
Military	-2.26	-3.82
Home is in CBD	0.499	2.98
Alternative-specific constant (works from home)	-0.387	-3.48
Pre-COVID constant	-1.06	-13.06

5.3 EXTERNAL WORKER IDENTIFICATION MODEL

The external worker identification model predicts which workers have a usual out-of-home workplace that is out of the region. If these workers generate a work tour, the primary destination of the tour would be an external station, predicted by the External Worker Location Choice Model.

Data used to estimate this model includes 8,074 unique workers in the combined 2016 and 2022 SANDAG household travel survey data. Note that duplicate person days were not used for model estimation; only one record per worker was used. Of these workers, 147 workers chose to work in a location outside the region. These workers are tabulated by the closest external station number to the worker in Table 6. Figure 6 shows the share of external workers by distance to the closest external station.

TABLE 6: WORKERS BY CLOSEST EXTERNAL STATION AND EXTERNAL CHOICE

STATION NUMBER	EXTERNAL	INTERNAL	TOTAL
24323	0	2	2
24324	1	3	4
24325	2	30	32
24326	63	1137	1200
24327	28	499	527
24328	43	5143	5186
24329	9	1082	1091
24330	1	31	32
All	147	7927	8074

Estimation results after dropping insignificant variables are shown in Table 7. The results are intuitive, as follows.

- Distance to the nearest external station is negative, indicating decreasing probability to work externally the further away from an external station a worker lives.
- The logged size of the external station is positive and approximately 0.8, indicating that all else being equal, the probability of being an external worker is nearly proportional to the amount of traffic at the nearest external station.

- A positive coefficient was estimated if the worker lives within 2.5 miles of an external station, consistent with the observed share of external workers by distance in Figure 6.
- Working externally to the region is positively correlated with household income.
- The alternative-specific constant for 2016 data is positive, indicating that workers in 2016 were generally more likely to work external to the region than workers in 2022.
- There is a strongly negative constant for working external to the region, reflecting the relatively low overall share of workers who work external to the region.

The initial log likelihood of the model is -5,596 and the final log likelihood is -669. The model rho-squared is 0.8803.

FIGURE 6: SHARE OF EXTERNAL WORKERS BY DISTANCE TO CLOSEST EXTERNAL STATION

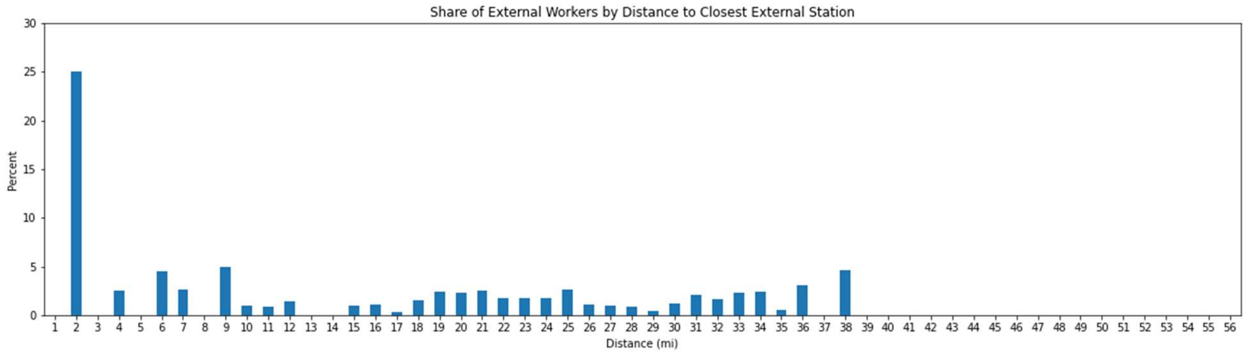


TABLE 7: EXTERNAL WORKER IDENTIFICATION ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Distance to the nearest external station	-0.0364	0.0117	-3.10
Logged size (Base-year or future year traffic count) of nearest external station	0.789	0.0788	10.02
Worker home is within 2.5 of an external station	2.30	1.27	1.81
Household income is less than $15k	-1.33	0.729	-1.83
Household income is greater than or equal to $15k and less than $25k	-0.933	0.432	-2.16
Household income is greater than or equal to $25k and less than $50k	-0.573	0.205	-2.80
Household income is greater than or equal to $150k and less than $250k	0.832	0.381	2.19
Household income is greater than or equal to $250k	1.23	0.455	2.70
Alternative-specific constant for 2016 survey	0.711	0.249	2.86
Alternative-specific constant	-10.7	0.775	-13.83

5.4 EXTERNAL WORKPLACE LOCATION CHOICE MODEL

The external workplace location choice model predicts which external station is the primary destination for the work tours generated by external workers. The 147 external workers who are identified as external workers were used to estimate the external workplace location choice model. The alternatives in the model are each external station.

Model coefficients are shown in Table 8. There are only three parameters in the model; the distance to the external station (negatively signed), the mode choice logsum to the external station (positively signed but relatively small), and the logged size of the external station (constrained to 1.0). The initial log likelihood is -224 and the final log likelihood is -220. The rho-squared is 0.0210.

TABLE 8: EXTERNAL WORKPLACE LOCATION CHOICE ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Distance to the external station (capped at 10 miles)	-0.519	0.258	-2.01
Mode choice logsum to the external station	0.0641	0.0489	1.31
Logged size (Base-year or future year traffic count) of external station	1.0	NA	NA

5.5 TRANSIT SUBSIDY MODEL

The transit pass subsidy model is a binary outcome choice model that predicts the likelihood of an individual receiving a subsidy for a transit pass. The transit subsidy can be provided through employment or through school. The model was estimated using 12,171 individuals who are either employed (full time or part time) or are students (of any education level). Of these individuals, 819 persons have a transit pass subsidy. The model was estimated using several sociodemographic variables and land use variables both at the home and workplace locations.

The final coefficients and t-statistics are provided in Table 9. The initial log likelihood for the estimation was -8,436.29 and the final log likelihood at convergence is -2,451.07. The rho-squared is 0.7095.

The estimated coefficients used in the final model and the results are summarized below:

Income: The income brackets that were significant in the model included household income less than $15,000, and household income between $100,000 and $200,000. Both of these had a negative correlation with transit pass subsidy.

Industry: Significant industries included business services, construction, education, government, healthcare, management services, military (all having a positive loading on the utility) and food services and retail (both having a negative loading on utility).

Employment status: Part time employment is correlated with a negative likelihood of having a transit pass with respect to the full time equivalent.

Student status: Being a student is negatively correlated with the likelihood of having a transit subsidy, however being a university student is positively correlated.

Parking cost at work: Higher parking costs at work are positively correlated with the likelihood of having a transit subsidy.

Transit access: Higher transit access both from households to employment and from employment to households is positively correlated with the likelihood of having a transit subsidy.

TABLE 9: TRANSIT SUBSIDY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Household income less than 15,000	-0.866	0.306	-2.83
Household income between 100,000 and 199,999	-0.340	0.134	-2.54
Industry is business services	1.07	0.124	8.63
Industry is construction	0.904	0.167	5.43
Industry is education	0.635	0.146	4.35
Industry is food services	-2.11	0.722	-2.92
Industry is government	1.46	0.205	7.13
Industry is healthcare	0.895	0.138	6.51
Industry is management services	0.800	0.143	5.59
Industry is military	1.72	0.200	8.60
Industry is retail	-1.59	0.514	-3.10
Is student	-0.870	0.191	-4.55
Cost of parking at work	0.0630	0.00348	18.08
Is parttime	-0.685	0.150	-4.55
Employment transit access to households	0.0431	0.00910	4.73
Household transit access to employment	0.0434	0.0113	3.84
Alternative Specific Constant	-4.19	0.183	-22.94

5.6 TRANSIT PASS OWNERSHIP MODEL

The transit pass ownership model predicts the likelihood of an individual to own a transit pass or not. This binary choice of ownership is to own or not own a transit pass. The model was estimated using 17,495 person records from 2016 and 2022. There were 1,920 individuals who owned transit passes. The ownership of a transit pass was estimated using socioeconomic variables of the person and household, transit attributes and land use variables.

The final coefficients and t-statistics are provided in Table 10. The initial log likelihood is –12,125.46 and the final log-likelihood is –4,418.10. The rho-squared is 0.6356.

The estimated coefficients used in the final model and results are summarized below:

Age: Age was broken into cohort groups, and only significant groups were retained in the model.

University status: Being a university student has a positive correlation with the likelihood of pass ownership.

Income: Income was broken into groups and significant incomes were retained in the model. The comparison group in the model was $25,000-49,999. The results indicated that household incomes less than the comparison group and greater than the comparison group were both significant, indicating that there may be two different factors contributing to pass ownership from different groups. There is likely an influence by the university students for the lower income groups, and employed persons who are more likely to receive a subsidy through work for the higher income groups.

Vehicle ownership at the household level: an indicator for zero car households, and one for car deficient households were used. Car deficient households are those households in which there are more workers than there are vehicles in the household.

Parking cost at work: A higher cost of parking in dollars at the place of work has a positive correlation with the likelihood of owning a transit pass.

Total accessibility by transit: Higher total accessibility by transit is positively correlated with the likelihood to own a transit pass.

Trip mode choice cost: The monthly transit pass is assumed to be $100/month. The trip cost for the transit pass is than computed as the transit pass cost divided by 44 (22 working days per month and 2 transit trips per day). This yields an expected value of the transit pass to be $2.27

per trip. For people under the age of 18 or age 65+, this is divided by two to reflect their reduced fare as indicated on the SDMTS fares description. The transit cost per trip is then scaled based on income and multiplied by the trip mode choice cost coefficient.

Transit subsidy provided: The provision of a subsidy through work or school is positively correlated with the likelihood of transit pass ownership.

TABLE 10: TRANSIT PASS OWNERSHIP RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Age 19 to 34	0.520	0.0869	5.98
Age 35 to 44	0.346	0.0981	3.53
Age 55 to 64	0.384	0.0925	4.15
Age 65 to 79	0.315	0.0944	3.34
Household income less than 15,000	0.492	0.104	4.70
Household income between 15,000 and 24,999	0.336	0.104	3.24
Household income between 50,000 and 99,999	0.233	0.0767	3.04
Household income between 100,000 and 199,999	1.17	0.0938	12.44
Household income greater than or equal to 200,000	1.17	0.143	8.15
University status	1.22	0.0891	13.73
Parking cost at work	0.0484	0.00569	8.51
Persons ages 0 to 4 in household	-0.551	0.0957	-5.76
Persons ages 5 to 15 in household	-0.578	0.0644	-8.98
Has transit subsidy	1.49	0.0964	15.44
Total transit accessibility	0.0692	0.00993	6.97
Trip mode choice cost coefficient	-1.25	NA	NA
Auto deficient household	1.21	0.0963	12.53
Household with zero autos	3.32	0.0920	36.11
Alternative Specific Constant	-4.10	0.137	-29.97

5.7 TOLL TRANSPONDER OWNERSHIP MODEL

The toll transponder ownership model predicts whether a household owns a toll transponder or not. *Fastrak* transponders are required in order to use I-15 managed lanes as a single-occupant vehicle, as there is no cash option. In the model we assume that if the household owns a transponder, it can be used in any household vehicle. If a transponder is owned, the members of that household can choose to drive alone on I-15 managed lanes. This is implemented by skimming the highway network separately for drive-alone vehicles by transponder ownership; drive-alone network skims for transponder owning households can include I-15 managed lanes, while drive-alone network skims for non-transponder owning households exclude I-15 managed lanes. Note that we also skim by value-of-time bin, so even if a household owns a transponder and could *potentially* use I-15 managed lanes, the skim *may or may not* include I-15 managed lanes if the time savings does not outweigh the cost of using the facility for the traveler's value of time. We assume that any household that does not own any vehicles also does not own a transponder.

The model was estimated using 8,980 households across both 2016 and 2022 surveys. The model replaces a previous model that was estimated from aggregate data on the number of households that owned a transponder by Census tract. That model was limited to using zonal attributes such as percent of households by auto ownership and zone-based accessibilities. This model uses household level variables such as actual cars owned, number of workers, and household income. Accessibility variables include distance to the nearest managed lane facility from the zonal centroid in miles, and the total distance traveled on managed lanes for work travel across all household members for the AM peak period in miles. The accessibility variables ensure that as SANDAG tests expansions to the managed lane network, transponder ownership will increase.

Two additional variables will be added to the model in application. One is that any household that owns an autonomous vehicle would also own a transponder. This assumes that autonomous vehicles are automatically configured to communicate with transponder sensors on managed lanes and handle billing for their use. The other is that any household that owns a vehicle of a certain age owns a transponder. That assumes that the same technology would be available in human-driven vehicles in future years. The minimum year where we assume this technology would be standard can be configured by the user.

Model coefficients are shown in Table 11. The initial log likelihood is -5724 and the final log likelihood is -3,164. The rho-squared is 0.4472. Following are some observations about the results:

- Zero worker households are much less likely to own a toll transponder.
- The likelihood of owning a toll transponder increases with income.
- The closer a household is to a toll road, the more likely they are to own a toll transponder.
- Toll transponder increases with two or more cars.
- Toll transponder ownership was slightly lower in 2016.

TABLE 11: TOLL TRANSPONDER OWNERSHIP ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Zero worker household	-0.291	0.0831	-3.50
Two or more autos owned	0.252	0.0714	3.53
Very low income (<$15,000)	-1.64	0.232	-7.08
Low income ($15k-$50k)	-1.07	0.139	-7.67
Medium income ($50k -	-0.312	0.124	-2.51
Very high income (	0.514	0.143	3.61
Distance to nearest managed lane facility	0.0332	0.00647	5.13
Total distance traveled on managed lanes for work across all workers	0.137	0.0143	9.64
Alternative-specific constant	-1.62	0.130	-12.45
Constant for 2016 households	-0.0239	0.0871	-0.27

5.8 TELECOMMUTE FREQUENCY MODEL

The telecommute frequency model predicts the frequency that each worker telecommutes on average in terms of days per week. The choices are no telecommuting (or less than one day per week), 1 day per week, 2-3 days per week, or 4+ days per week. The model was estimated with 7,309 unique workers from the combined 2016 and 2022 survey data. The model is informed by variables that describe household and person attributes (including industry of the worker), the composition of the household, and distance to work.

The final coefficients are shown in Table 12 and t-statistics are shown in Table 13. The initial log likelihood is -10,132 and the final log-likelihood is -5,283. The rho-squared is 0.4786.

Below we summarize the estimated coefficients:

- The presence of children less than 6 years old is negatively correlated with telecommuting 2 or more days per week.
- Workers who are the only adult in the household are more likely to telecommute 1 day per week.
- Part-time workers are less likely to telecommute.
- Workers in households earning between $60k and $100k are more likely to telecommute 1 day per week.
- Auto ownership is negatively correlated with telecommute frequency.
- Commute distance is positively correlated with telecommute frequency.
- The industry of the worker is significantly correlated with telecommute frequency. Workers in accommodations, education, food services, health care, military, and retail

industries are less likely to telecommute. Workers in business services, government and management services are more likely to telecommute.

- A set of alternative-specific constants were estimated specifically for 2016 data. These constants are negative for telecommute choices, indicating much lower levels of telecommuting in pre-COVID data.

TABLE 12: TELECOMMUTE FREQUENCY MODEL COEFFICIENTS

VARIABLE	NO TELECOMMUTE/LESS THAN 1 DAY/WEEK	1 DAY/ WEEK	2-3 DAYS/WEEK	4 DAYS/WEEK
Has children 0 to 5 years old	0	0	-0.328	-0.328
One adult in hh	0	0.231	0	0
Part-time worker	0	-0.165	-0.165	-0.165
Income 60-100k	0	0.355	0	0
1 Auto	0	-0.204	0	0
3+ Autos	0	-0.237	-0.267	-0.509
Distance to work	0	0.00658	0.0113	0.0113
Accommodation industry	0	-1.17	-1.17	-1.17
Business services industry	0	0.805	0.923	1.12
Education industry	0	-0.0388	-0.0388	-0.0388
Food services industry	0	-1.61	-1.61	-1.61
Government industry	0	0.45	0	0
Healthcare industry	0	-0.253	-0.253	-0.253
Management services industry	0	0	0.572	0.97
Miliary industry	0	0	-0.513	-0.513
Retail industry	0	-0.832	-0.758	-2.23
Model year is 2016	0	-0.125	-1.14	-3.71
Alternative specific constant	0	-1.82	-1.77	-1.37

TABLE 13: TELECOMMUTE FREQUENCY MODEL T-STATISTICS

VARIABLE	NO TELECOMMUTE/LESS THAN 1 DAY/WEEK	1 DAY/ WEEK	2-3 DAYS/WEEK	4 DAYS/WEEK
Has children 0 to 5 years old		NA	-2.33	-2.33
One adult in hh		2.25	NA	NA
Part-time worker		-1.97	-1.97	-1.97
Income 60-100k		4.70	NA	NA
1 Auto		-1.98	NA	NA
3+ Autos		-2.52	-2.18	-3.14
Distance to work		2.50	4.05	4.05
Accomodation industry		-4.50	-4.50	-4.50
Business services industry		8.99	7.57	7.32
Education industry		-0.43	-0.43	-0.43
Food services industry		-4.66	-4.66	-4.66
Government industry		2.12	NA	NA
Healthcare industry		-2.74	-2.74	-2.74
Management services industry		NA	3.71	4.56
Miliary industry		NA	-1.81	-1.81
Retail industry		-3.81	-2.75	-3.79
Model year is 2016		-1.38	-11.27	-19.88
Alternative specific constant		-17.16	-17.86	-14.33

5.9 COORDINATED DAILY ACTIVITY PATTERN MODEL

The coordinated daily activity pattern model (CDAP) predicts whether each person has a mandatory pattern (M), non-mandatory pattern (N), or stays at home (H). If a person has an M DAP then that person must take at least one work or school tour and is only available for workers or students. If a worker is not a student and works from home, they are not allowed to have an M DAP. People with an M pattern are also allowed to take non-mandatory tours. An N DAP means that person takes at least one joint or individual non-mandatory tour. This alternative is available to all people. A home pattern means that they do not leave the house for travel on that day.

The CDAP model as implemented for SANDAG includes an additional joint indicator to determine which people participate in joint tours. Only the individual M, N, and H patterns were re-estimated in this work and the joint utility will be addressed in calibration.

It is important to only include data where the whole household has completed the survey for that day. In practice, this means that only data with non-zero person-day weights are included. In the survey processing procedures, many incomplete households are included because some members might have reported travel during that day, which we would like to keep, but it may not have been reported for all people. For example, child travel is only reported on one day for the 2022 data. Since CDAP is a coordinated decision by all household members, all members of the household must have completed the survey.

Interaction between household members is what the “coordinated” part of the CDAP model represents. There is an added utility component that is added for members of the household having the same M, N, or H pattern. Since the CDAP model must account for interactions between members across the household, the output is really a household level-decision where the alternatives are the specification for each person’s daily activity pattern.

To demonstrate, let’s look at a couple of household sizes. First of all, a single person household has just the individual component and contains three alternatives: M, N, and H. For a two-person household, each individual can have any of the three alternatives which totals to six possible choices: MM, MN, MH, NM, NN, NH, HM, HN, HH. (In implementation, the joint component extends the alternatives with non-home components to produce an additional 4 alternatives: MMJ, MNJ, NMJ, NNJ.) In this two-person household example, interaction coefficients would be applied to all cases where both household members are performing the same activity. If both people in the household are full-time workers, then the added utility for the MM alternative would include the coef_M_11 where the two 1’s indicate the person types of the first and second person. Interaction coefficients only apply to alternatives where members of the household are performing the same DAP and represent the added utility of people performing activities together. Interaction coefficients are therefore generally positive due to this added utility of coordinating your activities across other members of the household.

The number of alternatives grows combinatorically with additional household members up to a household size of five. For households with six or more people, the first five members of the household (ranked by worker status, student status, and age) are assigned a CDAP through the described process. The remaining household members have their CDAP assigned based on a probability lookup table segmented by person type. Probabilities were not adjusted as part of the estimation process.

Estimation results are broken out into the individual component shown in Table 14 and interaction coefficients shown in Table 17. Some observations in the individual component include:

- People who work from home make more N travel.
- As telecommuting increases, the amount of M travel decreases.

- Younger people tend to have fewer at home days compared to older people. People over the age of 80 are especially more likely to stay at home.
- An increase in auto sufficiency generally showed a decrease in the H pattern, but this was not significant for all person types.
- Low income full-time and part-time workers are more likely to stay at home.
- Pre-covid 2016 constants were generally negative for the M and N patterns. This might be correlated with the work from home increase post-covid.
- Accessibility terms were tried in many different forms: aggregate, disaggregate, workplace and school location mode choice logsums. The decided upon accessibility that would converge in estimation and produce reasonable results was the inclusion of non-mandatory accessibility binned into 4 levels. Non-mandatory accessibility in this context is the average between the shopping and other discretionary accessibilities produced by the disaggregate accessibility model. This was then binned into [<12, 13.5, 15, >15] which roughly corresponds to quantiles in the data. With the lowest bin acting as the base coefficient, what we see is that with increased accessibility, the CDAP pattern is more likely to be N, as expected. Note that the 2016 coefficients on accessibility dampen the effect for pre-coved, but the net effect of base coefficients plus 2016 coefficients is still positive.

TABLE 14: ESTIMATED CDAP INDIVIDUAL COEFFICIENT VALUES

VARIABLE DESCRIPTION	M	N	H
Full-time worker ASC	1.56	1.46	
Part-time worker ASC	1.64	1.61	
University student ASC	0.350	0.350	
Non-working adult ASC		1.18	
Retired ASC		1.15	
Driving-age child who is in school ASC	-0.03	-0.35	
Pre-driving-age child who is in school ASC	0.805	-0.07	
Preschool child ASC	-0.42	-0.78	
Full-time worker interaction with age less than 35	-0.12	-0.16	
Part-time worker interaction with age less than 35	0.232		
Age over 80		-0.41	
Zero auto ownership base	-0.38	-0.59	
Zero auto ownership university student			
Zero auto ownership non-working adult			
Auto Deficient part-time worker	0.377		
Auto Deficient university student			
Auto Deficient non-working adult		-0.84	
Auto Deficient retired		-0.58	
Auto Deficient driving age child		-0.95	
Auto Deficient pre-driving age child			
Auto Deficient preschool child	-0.62		

VARIABLE DESCRIPTION	M	N	H
Full-time worker interaction with income less than $30k			0.157
Part-time worker interaction with income less than $30k			0.234
University student interaction with income between $60k and $100k			
Preschool child interaction with income between $60k and $100k	0.272		
Base works from home		-0.36	
Works from home 2016		-0.41	
Part time worker who works from home		-1.23	
Part time worker who works from home 2016		1.43	
Telecommutes 1 day per week		-0.72	-0.78
Telecommutes 2-3 days per week		0.996	1.49
Telecommutes 4 days per week		0.350	0.350
non-mandatory accessibility > 15		0.449	
13.5 < non-mandatory accessibility < 15		0.238	
12 < non-mandatory accessibility < 13.5		0.210	
Full-time worker 2016 ASC	1.22	0.692	
Part-time worker 2016 ASC	0.298	0.352	
University student 2016 ASC	0.350	0.350	
Non-working adult 2016 ASC		0.233	
Retired 2016 ASC		0.310	
Driving-age child who is in school 2016 ASC	0.976	-0.25	
Pre-driving-age child who is in school 2016 ASC	0.111	0.104	
Preschool child 2016 ASC	0.0691	0.859	
Telecommutes 1 day per week 2016		0.530	-0.58
Telecommutes 2-3 days per week 2016		-0.60	-0.41
Telecommutes 4 days per week 2016		0.350	0.350
non-mandatory accessibility > 15, 2016		-0.31	
13.5 < non-mandatory accessibility < 15, 2016		-0.17	
12 < non-mandatory accessibility < 13.5, 2016		-0.15	

The above table contains only the estimated coefficient values. T-statistics for each coefficient are displayed below in Table 15. In general, only the coefficients with significance were kept with a couple exceptions. All alternative specific constants were kept, regardless of significance, for the model symmetry. This only applies to two coefficients: the N pattern for the university student and the preschool child 2016 N pattern.

TABLE 15: ESTIMATED CDAP COEFFICIENT T-STATS

VARIABLE DESCRIPTION	M	N	H
Full-time worker ASC	21.09	16.67	
Part-time worker ASC	9.41	9.12	
University student ASC	4.62	4.62	
Non-working adult ASC		10.05	
Retired ASC		15.98	
Driving-age child who is in school ASC	-0.09	-0.94	
Pre-driving-age child who is in school ASC	6.99	-0.50	
Preschool child ASC	-2.30	-4.09	
Full-time worker interaction with age less than 35	-2.02	-2.72	

VARIABLE DESCRIPTION	M	N	H
Part-time worker interaction with age less than 35	2.25		
Age over 80		-4.95	
Zero auto ownership base	-4.40	-8.74	
Zero auto ownership university student			
Zero auto ownership non-working adult			
Auto Deficient part-time worker	3.26		
Auto Deficient university student			
Auto Deficient non-working adult		-9.92	
Auto Deficient retired		-7.34	
Auto Deficient driving age child		-2.01	
Auto Deficient pre-driving age child			
Auto Deficient preschool child	-3.26		
Full-time worker interaction with income less than $30k			1.84
Part-time worker interaction with income less than $30k			1.78
University student interaction with income between $60k and $100k			
Preschool child interaction with income between $60k and $100k	2.52		
Base works from home		-3.76	
Works from home 2016		-3.07	
Part time worker who works from home		-4.22	
Part time worker who works from home 2016		4.17	
Telecommutes 1 day per week		-6.23	-3.84
Telecommutes 2-3 days per week		7.08	8.66
Telecommutes 4 days per week		4.62	4.62
non-mandatory accessibility > 15		5.75	
13.5 < non-mandatory accessibility < 15		3.61	
12 < non-mandatory accessibility < 13.5		3.40	
Full-time worker 2016 ASC	14.76	6.90	
Part-time worker 2016 ASC	1.58	1.81	
University student 2016 ASC	4.62	4.62	
Non-working adult 2016 ASC		1.93	
Retired 2016 ASC		4.00	
Driving-age child who is in school 2016 ASC	2.75	-0.62	
Pre-driving-age child who is in school 2016 ASC	1.03	0.78	
Preschool child 2016 ASC	0.38	4.48	
Telecommutes 1 day per week 2016		4.23	-7.34
Telecommutes 2-3 days per week 2016		-3.67	-1.89
Telecommutes 4 days per week 2016		4.62	4.62
non-mandatory accessibility > 15, 2016		-3.53	
13.5 < non-mandatory accessibility < 15, 2016		-2.20	
12 < non-mandatory accessibility < 13.5, 2016		-2.08	

The significant interaction coefficients are shown in Table 16. The structure of these coefficients is displayed such that the numbers represent the different person types. Person types are for reference are: 1) Full-time worker 2) part-time worker 3) university student 4) non-worker 5) retired 6) driving-age student 7) school-age student 8) preschool student. So, for example, coef_H_18 would be the interaction between a full-time worker and a preschool age child. Coefficients that are labeled with "_xxxxx" represent the interaction term for all members of the household performing the same activity. Many of the three-way interaction coefficients were combined for symmetry and to get better statistics for these groups, as can be seen in the table.

There were many interaction coefficients which were removed because they were not significant, particularly for three-way interactions. As mentioned previously, most interaction coefficients are positive and represent the added utility of household members. Negative interaction coefficients show up in the 3+ person interactions and are just balancing out the large positive two-way interaction coefficients that also apply to that household pattern leading to a net positive interaction effect as expected.

TABLE 16: ESTIMATED CDAP INTERACTION COEFFICIENTS

INTERACTION COEFFICIENT	VALUE	T-STAT
coef_H_11	0.908	8.28
coef_H_12	0.588	2.27
coef_H_13	0.529	2.98
coef_H_14	0.612	5.07
coef_H_15	0.786	6.65
coef_H_17	0.169	1.87
coef_H_18	0.370	4.04
coef_H_22	1.00	2.81
coef_H_25	0.584	2.20
coef_H_33	1.35	7.49
coef_H_34	0.482	1.94
coef_H_44	0.557	3.35
coef_H_45	0.501	3.31
coef_H_55	0.272	2.45
coef_H_66	3.14	5.95
coef_H_77	1.97	18.20
coef_H_78	1.62	7.31
coef_H_88	1.95	14.84
coef_M_11	0.206	4.70
coef_M_13	0.138	2.07
coef_M_17	0.101	2.96
coef_M_18	0.309	4.47
coef_M_27	0.246	3.71
coef_M_33	0.255	1.80
coef_M_77	1.18	13.64
coef_M_78	0.539	2.19
coef_M_88	1.31	7.97
coef_N_15	-0.306	-3.81

INTERACTION COEFFICIENT	VALUE	T-STAT
coef_N_17	0.149	3.04
coef_N_18	0.225	3.62
coef_N_26	0.979	3.54
coef_N_28	0.301	2.62
coef_N_35	-0.802	-3.15
coef_N_55	-0.15	-2.08
coef_N_66	2.24	2.80
coef_N_67	2.00	3.61
coef_N_77	1.62	15.04
coef_N_78	1.12	4.74
coef_N_88	1.22	8.78
coef_H_777_778_788_888	-1.17	-6.54
coef_M_111	-0.323	-1.94
coef_M_122	0.530	2.97
coef_M_127_128	0.285	2.66
coef_M_777_778_788_888	-0.567	-4.22
coef_M_xxxxx	0.327	4.35
coef_N_777_778_788_888	-0.732	-4.23
coef_N_xxxxx	0.350	4.62

5.10 NON-MANDATORY TOUR FREQUENCY MODEL

The non-mandatory tour frequency model predicts the number of non-mandatory tours taken by each person. There are six non-mandatory tour purposes: escorting, shopping, social, eat out, other discretionary (gym, religious services, etc.), and other maintenance (medical appointments, auto repair, etc.). Non-mandatory tour frequency configuration files are segmented by person type, and each person type is estimated separately. Each alternative in the non-mandatory tour frequency model consists of the number of tours of each of the six purposes. Additionally, a variable is added to the alternatives to determine the total number of individual non-mandatory tours in that alternative.

All person types have the same base model specification. The specification begins with purpose specific coefficients multiplied by the number of tours with that purpose for that alternative. For example, if an alterative has two shopping tours and one escort tour, and the estimated coefficient is -2 for shopping and -1 for escort, the total utility from the purpose specific constants would be -2 * 2 (shopping) + -1 * 1 (escort) = -5 utiles. In addition to purpose specific constants, there are also coefficients on the total number of non-mandatory tours in each alternative. A common structure would be to have a constant for 1, 2, and 3plus total tours. In the above example, the alternative has 3 total non-mandatory tours, so the 3plus coefficient would be applied to the utility.

Beyond the purpose specific and total number of individual non-mandatory tours, there are additional variables that were tried for each person type. Variables attempted for all person types are income, age, and the presence of previous mandatory, joint, and school escorting trips. For workers, work from home and telecommute frequency variables were tried.

Coefficients on shopping or other discretionary accessibilities calculated from the disaggregate accessibility model multiplied by the number of tours of that purpose were included in the estimation for all person types. If the resulting accessibility coefficient was positive, it was included in the final model specification.

Pre-covid variables were also added to all person types. These pre-covid coefficients were applied based on the total number of individual non-mandatory tours and only to the 2016 survey data.

All person types had coefficients based on the total amount of time available to that person. A person would be unavailable for times when previously determined mandatory, joint, or school escorting tours were scheduled. The available time was restricted between 7am and 10pm, representing the typical time of day for non-mandatory travel. Time variables were split into two categories, one which denoted the total number of consecutive hours available to a person between, and the second was the log of the total number of hours available. Both time categories were them multiplied by the number of tours of each purpose. In general, the logged number of total available hours provided more explanatory power across person types and tour purposes than the consecutive hours term.

The results of non-mandatory tour frequency estimation can be seen below in Table 17. This model demonstrates several trends that can be seen across all person types. First, there are large negative coefficients that are multiplied by the number of tours by purpose. These negative coefficients are counteracting the positive coefficients on the total number of individual non-mandatory tours. Negative coefficients on the purpose specific constants indicate that a person is much less likely to make multiple of the same type of non-mandatory tour on the same day.

There is a strong positive correlation between the length of the available time for non-mandatory tours between the hours of 7am and 10pm. Two time-availability variables were tried: a logged term based on the total number of free hours, and a linear term on the largest consecutive number of hours. The total number of hours was logged to represent the decreasing utility one might expect to make a tour with a larger number of free hours while the linear number of consecutive hours accounts for non-mandatory tours that might be long in duration. For most person types and tour purposes, the logged term was often the more significant of the two. Some person types had significance in both variables for some purposes. For example, in the full-time worker table below, shopping and other discretionary purposes had positive and significant coefficients for the largest consecutive number of free hours available.

Other trends in the full-time worker estimation results that are seen across all person types include:

- The total number of non-mandatory tours in 2016 was larger than in 2022.
- Positive coefficients on the disaggregate accessibilities multiplied by the number of tours by purpose were kept regardless of their significance to provide sensitivity to accessibility. Not all purposes produced positive coefficients. Escorting for example did not, which is to be expected considering accessibility is often not a large consideration for escorting people to a particular location.
- The presence of mandatory and joint tours produced a negative correlation to the number of individual non-mandatory tours.
- The presence of previous school escorting tours makes making an escort tour less likely.
- The number of children in the household increases the non-mandatory tour rates, especially for the escort purpose.
- As the household size increases, the number of tours go down, likely due to other members of the household being able to accomplish them.
- Zero auto and auto deficient households make less tours compared to auto sufficient households.
- As income increases, the number of tours increases likely due to the increase in disposable income.

For brevity, individual non-mandatory was shortened to iNM in the following tables.

TABLE 17: FULL-TIME WORKER NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-9.98	0.525	-19.01
Number of iNM other discretionary tours	-7.46	0.518	-14.42
Number of iNM shopping tours	-8.73	0.636	-13.71
Number of iNM other maintenance tours	-11.5	0.594	-19.35
Number of iNM social tours	-9.14	0.746	-12.25
Number of iNM eatout tours	-8.00	0.561	-14.24
Total number of iNM tours is 1	3.66	0.410	8.93
Total number of iNM tours is 2	3.78	0.805	4.70
Total number of iNM tours is 3	3.95	1.20	3.28
Total number of iNM tours is 4p	4.04	1.64	2.46
Total number of iNM tours is 1 for 2016	0.545	0.0654	8.33
Total number of iNM tours is 2 for 2016	0.449	0.0849	5.28
Total number of iNM tours is 3p for 2016	0.357	0.128	2.79

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Total number of iNM tours is 1 and has a mandatory tour	-2.71	0.111	-24.42
Total number of iNM tours is 2p and has a mandatory tour	-1.86	0.155	-12.01
Total number of iNM tours is 1 and has a joint tour	-2.68	0.112	-23.98
Total number of iNM tours is 2p and has a joint tour	-2.58	0.138	-18.75
Number of iNM escort tours * number of school escort tours determined by the school escorting model	-0.566	0.186	-3.05
Number of iNM other maintenance tours & income < $30k	0.168	0.0574	2.93
Number of iNM other discretionary tours & income < $30k	-0.425	0.0832	-5.11
Number of iNM escort tours & $30k < income < $60k	-0.262	0.0674	-3.89
Number of iNM other discretionary tours & $30k < income < $60k	-0.218	0.0537	-4.06
Number of iNM other discretionary tours & $100k < income < $150k	0.286	0.0786	3.64
Number of iNM escort tours & income > $150k	0.529	0.0844	6.27
Number of iNM other discretionary tours & income > $150k	0.552	0.0678	8.15
Total number of iNM tours is 1 & zero auto household	-0.778	0.196	-3.97
Total number of iNM tours is 2 & zero auto household	-1.75	0.325	-5.38
Total number of iNM tours is 3p & zero auto household	-1.49	0.498	-3.00
Total number of iNM tours is 1 & auto deficient household	-0.133	0.0848	-1.57
Total number of iNM tours is 2p & auto deficient household	-0.166	0.112	-1.48
Number of iNM escort tours * number of children	0.784	0.0272	28.81
Number of iNM shopping tours * number of children	0.164	0.0564	2.92
Number of iNM other maintenance tours * number of children	0.230	0.0393	5.85
Number of iNM social tours * number of children	0.436	0.146	2.98
Number of iNM other discretionary tours * number of children	0.213	0.0528	4.04
Number of iNM escort tours * number of full-time workers	0.266	0.0482	5.51
Number of iNM other discretionary tours * number of full-time workers	-0.127	0.0498	-2.55
Number of iNM shopping tours * household size	-0.130	0.0402	-3.23
Number of iNM other maintenance tours * household size	-0.114	0.0283	-4.02
Number of iNM eatout tours * household size	-0.268	0.0341	-7.88
Number of iNM social tours * household size	-0.496	0.102	-4.86
Number of iNM other discretionary tours * household size	-0.187	0.0404	-4.63
Number of iNM escort tours & age 18 to 24	-0.785	0.342	-2.30
Number of iNM escort tours & age 25 to 35	-0.332	0.0748	-4.45
Number of iNM eatout tours & age 25 to 35	-0.147	0.0754	-1.94
Number of iNM escort tours & age 50 to 79	-0.166	0.0755	-2.20
Number of iNM eatout tours & age 50 to 79	-0.337	0.0778	-4.33
Number of iNM other discretionary tours & age 50 to 79	-0.117	0.0481	-2.43
Total number of iNM tours is 2 & age 18 to 24	0.518	0.205	2.52

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	2.31	0.120	19.25
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	1.99	0.250	7.98
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	3.36	0.224	15.01
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	1.86	0.112	16.61
Number of iNM social tours * log of total number of free hours between 7am to 10pm	1.61	0.164	9.78
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	1.94	0.0989	19.57
Number of iNM shopping tours * number of consecutive free hours between 7am to 10pm	0.0385	0.0209	1.85
Number of iNM other maintenance tours * number of consecutive free hours between 7am to 10pm	0.0541	0.0164	3.30
Number of iNM shopping tours * shopping accessibility	0.0249	0.0189	1.32
Number of iNM eatout tours * other discretionary accessibility	0.0469	0.0139	3.37
Number of iNM other discretionary tours * other discretionary accessibility	0.0293	0.00977	2.99
Number of iNM social tours * other discretionary accessibility	0.0797	0.0245	3.26
Total number of iNM tours is 2 & work from home	0.288	0.0762	3.78
Total number of iNM tours is 3p & work from home	0.433	0.124	3.50
Total number of iNM tours is 2 & telecommute 2, 3, or 4 days per week	0.0602	0.0923	0.65
Total number of iNM tours is 3p & telecommute 2, 3, or 4 days per week	0.140	0.159	0.88

Full-time worker results had some results not seen in other person types. Work from home and telecommute frequency saw a positive correlation to the number of individual non-mandatory tours. This can be explained by people having more freedom throughout their day to take these tours and the creation of tours that might otherwise have been stops to or from work. The 36 to 49 reference age group had the most tours relative to other age groups, but this was only significant for some purposes.

Much of the success in estimating household and person demographic variables for full-time workers comes from the good statistics for that person type. The ability to get significant coefficients out of the estimation decreases significantly with the number of people captured by the household travel survey for each person type.

Part time worker non-mandatory tour frequency estimation results are shown in Table 18. Besides the trends previously discussed that are common across all person types, part time workers did not show significance towards auto sufficiency levels and had only one significant term on telecommute frequency, showing a negative correlation of having one individual non-mandatory tour and telecommuting 4 days per week. Also similar to the full-time worker results, part-time workers had a positive and significant coefficient on shopping tours for the number of consecutive free hours. All purposes had positive and significant coefficients related to the log of the total number of free hours.

TABLE 18: PART-TIME WORKER NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-6.84	0.868	-7.89
Number of iNM other discretionary tours	-7.06	0.815	-8.66
Number of iNM shopping tours	-7.27	1.20	-6.06
Number of iNM other maintenance tours	-10.6	0.891	-11.94
Number of iNM social tours	-7.51	1.41	-5.31
Number of iNM eatout tours	-8.03	1.05	-7.65
Total number of iNM tours is 0 and no prior tours	4.35	0.396	11.00
Total number of iNM tours is 1	5.42	0.605	8.97
Total number of iNM tours is 2	5.89	0.997	5.92
Total number of iNM tours is 3p	5.94	1.50	3.96
Total number of iNM tours is 1 for 2016	0.174	0.151	1.16
Total number of iNM tours is 2 for 2016	0.188	0.187	1.00
Total number of iNM tours is 3 for 2016	0.702	0.305	2.30
Total number of iNM tours is 4p for 2016	1.29	0.556	2.32
Total number of iNM tours is 1 and has a joint tour	-0.291	0.138	-2.11
Number of iNM escort tours * number of school escort tours determined by the school escorting model	-0.612	0.238	-2.57
Number of iNM escort tours & income < $30k	-0.418	0.150	-2.79
Number of iNM other discretionary tours & income < $30k	-0.195	0.107	-1.82
Number of iNM escort tours & $30k < income < $60k	-0.379	0.120	-3.17
Total number of iNM tours is 1 & zero auto household	-1.14	0.367	-3.10
Total number of iNM tours is 2p & zero auto household	-1.37	0.474	-2.88
Total number of iNM tours is 1 & auto deficient household	-0.426	0.170	-2.50
Total number of iNM tours is 2p & auto deficient household	-0.761	0.231	-3.30
Number of iNM escort tours * number of children	0.649	0.0502	12.93
Number of iNM shopping tours * number of children	0.148	0.0538	2.75
Number of iNM other discretionary tours * number of children	0.0924	0.0500	1.85
Number of iNM eatout tours * number of full-time workers	-0.244	0.131	-1.87
Number of iNM escort tours & age 18 to 24	-1.28	0.510	-2.51
Number of iNM escort tours & age 25 to 35	-0.710	0.160	-4.43

Number of iNM shopping tours & age 25 to 35	-0.459	0.157	-2.92
Number of iNM other discretionary tours & age 25 to 35	-0.323	0.133	-2.43
Number of iNM escort tours & age 50 to 79	-0.805	0.137	-5.86
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	1.46	0.215	6.78
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	1.05	0.525	2.01
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	3.13	0.279	11.19
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	1.45	0.291	4.98
Number of iNM social tours * log of total number of free hours between 7am to 10pm	1.00	0.403	2.49
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	1.59	0.213	7.48
Number of iNM shopping tours * number of consecutive free hours between 7am to 10pm	0.0849	0.0362	2.35
Number of iNM escorting tours * shopping accessibility	0.0349	0.0394	0.89
Number of iNM shopping tours * shopping accessibility	0.0538	0.0372	1.44
Number of iNM eatout tours * other discretionary accessibility	0.0684	0.0328	2.09
Number of iNM other discretionary tours * other discretionary accessibility	0.0241	0.0220	1.10
Number of iNM social tours * other discretionary accessibility	0.0347	0.0529	0.66
Total number of iNM tours is 1 & telecommute 4 days per week	-1.02	0.311	-3.28

University student non-mandatory tour frequency estimation results are shown below in Table 19. Survey collection methodology excludes the presence of university students living in group quarter households. The university student estimation differs from the worker person types above in that fewer household coefficients are significant for this model. This is likely due to the household make-up of University students being different in general from the worker population.

A couple of household characteristics that did stick around as significant were an increase in escort tours based on the number of children in the household and an increase in the number of eatout tours based of full-time workers balanced by a decrease in the number of eatout tours based on household size. A couple of purpose specific age variables were significant for age buckets less than 35 years old, likely due to the usual age of university students being younger. Interestingly, social tours did not have a significant coefficient on the number of available hours in the day. This is likely due to the unique social characteristics of university students.

TABLE 19: UNIVERSITY STUDENT NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-8.50	1.29	-6.59
Number of iNM other discretionary tours	-4.81	0.692	-6.95
Number of iNM shopping tours	-12.7	1.94	-6.54

Number of iNM other maintenance tours	-11.1	1.59	-6.96
Number of iNM social tours	-4.46	0.437	-10.21
Number of iNM eatout tours	-6.16	1.36	-4.54
Total number of iNM tours is 0 and no prior tours	5.00	0.519	9.63
Total number of iNM tours is 1	5.71	0.650	8.78
Total number of iNM tours is 2	6.83	0.964	7.09
Total number of iNM tours is 3p	7.21	1.37	5.26
Total number of iNM tours is 2p for 2016	-0.679	0.168	-4.04
Total number of iNM tours is 1 & zero auto household	-1.49	0.498	-2.99
Total number of iNM tours is 2p & zero auto household	-2.14	0.641	-3.34
Total number of iNM tours is 1 & auto deficient household	-0.461	0.189	-2.44
Total number of iNM tours is 2 & auto deficient household	-0.693	0.254	-2.73
Total number of iNM tours is 3p & auto deficient household	-1.20	0.479	-2.51
Number of iNM escort tours * number of children	0.770	0.0641	12.01
Number of iNM eatout tours * number of full-time workers	0.423	0.177	2.39
Number of iNM eatout tours * household size	-0.390	0.117	-3.34
Number of iNM shopping tours & age 18 to 24	-0.404	0.203	-1.99
Number of iNM eatout tours & age 18 to 24	0.625	0.214	2.93
Number of iNM shopping tours & age 25 to 35	-0.289	0.172	-1.68
Number of iNM other discretionary tours & age 25 to 35	0.308	0.138	2.23
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	1.88	0.445	4.23
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	3.04	0.633	4.80
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	2.88	0.513	5.62
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	0.806	0.368	2.19
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	0.119	0.0232	5.11
Number of iNM shopping tours * shopping accessibility	0.162	0.0545	2.96
Number of iNM other maintenance tours * shopping accessibility	0.0713	0.0447	1.59
Number of iNM eatout tours * other discretionary accessibility	0.0889	0.0417	2.13
Number of iNM other discretionary tours * other discretionary accessibility	0.0448	0.0277	1.62

The non-worker person type non-mandatory tour frequency estimation results stands apart from the previous estimation results in a couple of interesting ways. First of all, the number of significant time availability coefficients is much less than the worker or student person types, which can be explained by the non-worker's lack of mandatory tours in their day. This person type also shows more significance around the number of previously decided joint tours. Additionally, non-workers show less sensitivity to auto sufficiency levels. Other than these specific instances, the non-worker estimation shown in Table 20 has similar trends to the previously discussed estimations.

TABLE 20: NON-WORKER NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-11.2	2.71	-4.13
Number of iNM other discretionary tours	-2.79	0.408	-6.84
Number of iNM shopping tours	-3.39	0.479	-7.08
Number of iNM other maintenance tours	-2.06	0.414	-4.97
Number of iNM social tours	-4.54	0.701	-6.48
Number of iNM eatout tours	-2.11	0.179	-11.77
Total number of iNM tours is 1	3.15	0.249	12.66
Total number of iNM tours is 2	3.20	0.334	9.58
Total number of iNM tours is 3p	3.16	0.463	6.81
Total number of iNM tours is 1 for 2016	0.382	0.196	1.94
Total number of iNM tours is 2p for 2016	0.189	0.206	0.92
Total number of iNM tours is 1 and has a joint tour	-3.60	0.177	-20.32
Total number of iNM tours is 2p and has a joint tour	-3.46	0.222	-15.63
Number of iNM escort tours * number of school escort tours determined by the school escorting model	-1.01	0.153	-6.59
Number of iNM escort tours & income < $30k	-0.352	0.118	-2.99
Number of iNM other maintenance tours & income < $30k	-0.404	0.105	-3.85
Number of iNM escort tours & $30k < income < $60k	-0.232	0.0996	-2.32
Number of iNM other maintenance tours & $30k < income < $60k	-0.281	0.0946	-2.96
Number of iNM other discretionary tours & $100k < income < $150k	0.325	0.175	1.86
Total number of iNM tours is 3p & zero auto household	-1.06	0.427	-2.49
Total number of iNM tours is 1 & auto deficient household	-0.216	0.176	-1.23
Total number of iNM tours is 2p & auto deficient household	-0.158	0.188	-0.84
Number of iNM escort tours * number of children	0.558	0.0662	8.43
Number of iNM social tours * number of children	0.217	0.0942	2.31
Number of iNM escort tours * number of full-time workers	0.265	0.0931	2.85
Number of iNM other discretionary tours * number of full-time workers	-0.180	0.0886	-2.04
Number of iNM escort tours * household size	0.111	0.0580	1.91
Number of iNM other maintenance tours * household size	-0.101	0.0395	-2.57
Number of iNM eatout tours * household size	-0.139	0.0503	-2.77
Number of iNM social tours & age 18 to 24	1.14	0.541	2.12
Number of iNM escort tours & age 25 to 35	-0.561	0.119	-4.71
Number of iNM shopping tours & age 25 to 35	-0.311	0.107	-2.90
Number of iNM other discretionary tours & age 25 to 35	-0.343	0.0962	-3.57
Number of iNM escort tours & age 50 to 79	-0.381	0.105	-3.64
Total number of iNM tours is 2 & age 18 to 24	-0.947	0.483	-1.96
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	3.23	0.976	3.30
Number of iNM shopping tours * number of consecutive free hours between 7am to 10pm	0.0681	0.0254	2.67

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM other maintenance tours * number of consecutive free hours between 7am to 10pm	0.0531	0.0253	2.09
Number of iNM other discretionary tours * number of consecutive free hours between 7am to 10pm	0.0621	0.0224	2.77
Number of iNM shopping tours * shopping accessibility	0.0578	0.0252	2.29
Number of iNM other discretionary tours * other discretionary accessibility	0.0265	0.0144	1.84
Number of iNM social tours * other discretionary accessibility	0.0453	0.0430	1.05

Retired person type non-mandatory tour frequency estimation results shown in Table 21 have a lot of the general trends already discussed. Where the retired person type sticks out is in the couple of coefficients on aged 80 and over people traveling less. There are a few more eatout related terms compared to other person types likely due to the increased role of eating out for retired people. Interestingly, most of the low-income purpose specific coefficients were significant and showed that low-income retired people travel less.

TABLE 21: RETIRED NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-13.3	4.62	-2.89
Number of iNM other discretionary tours	-5.03	1.65	-3.05
Number of iNM shopping tours	-9.14	2.39	-3.82
Number of iNM other maintenance tours	-9.95	2.49	-4.00
Number of iNM social tours	-4.19	0.445	-9.42
Number of iNM eatout tours	-4.45	0.773	-5.75
Total number of iNM tours is 1	6.00	0.548	10.95
Total number of iNM tours is 2	5.88	0.933	6.31
Total number of iNM tours is 3	5.83	1.35	4.32
Total number of iNM tours is 4p	5.93	1.83	3.25
Total number of iNM tours is 1 for 2016	0.170	0.154	1.10
Total number of iNM tours is 2 for 2016	0.324	0.167	1.94
Total number of iNM tours is 3p for 2016	0.296	0.201	1.48
Total number of iNM tours is 1 and has a joint tour	-5.80	0.333	-17.41
Total number of iNM tours is 2p and has a joint tour	-5.76	0.357	-16.16
Number of iNM escort tours & income < $30k	-0.684	0.154	-4.43
Number of iNM shopping tours & income < $30k	-0.368	0.0864	-4.26
Number of iNM other maintenance tours & income < $30k	-0.773	0.0969	-7.98
Number of iNM eatout tours & income < $30k	-0.662	0.131	-5.04
Number of iNM other discretionary tours & income < $30k	-0.533	0.0779	-6.84
Number of iNM other discretionary tours & $30k < income < $60k	-0.216	0.0698	-3.09
Number of iNM shopping tours & income > $150k	-0.251	0.129	-1.94
Total number of iNM tours is 1 & auto deficient household	-0.787	0.171	-4.61
Total number of iNM tours is 2 & auto deficient household	-1.05	0.197	-5.30
Total number of iNM tours is 3p & auto deficient household	-1.12	0.272	-4.10
Number of iNM escort tours * number of children	0.754	0.112	6.72

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escort tours * number of full-time workers	0.334	0.114	2.92
Number of iNM shopping tours * household size	-0.143	0.0522	-2.73
Number of iNM other maintenance tours * household size	-0.205	0.0534	-3.84
Number of iNM eatout tours * household size	-0.297	0.0854	-3.48
Number of iNM other discretionary tours * household size	-0.171	0.0460	-3.71
Number of iNM escort tours & retired only household	-0.329	0.142	-2.32
Number of iNM escort tours & age 80+	-0.591	0.240	-2.47
Number of iNM eatout tours & age 80+	-0.408	0.179	-2.28
Total number of iNM tours is 2 & age 80+	-0.254	0.135	-1.88
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	4.56	1.82	2.50
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	2.73	0.847	3.22
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	3.12	0.882	3.54
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	1.42	0.574	2.47
Number of iNM escort tours * number of consecutive free hours between 7am to 10pm	-0.145	0.0422	-3.44
Number of iNM eatout tours * number of consecutive free hours between 7am to 10pm	0.0767	0.0362	2.12
Number of iNM eatout tours * other discretionary accessibility	0.0879	0.0192	4.59

Driving age student non-mandatory tour frequency estimation results seen in Table 22 suffered from very low statistics. Only 278 person-days worth of data was available to estimate on compared to the 15,000 observations for full-time workers. For this reason, very few household or person demographics were able to be estimated. What did turn out to be significant was an increase in escort tours with an increase in the number of children in the household, an increase in eatout tours based on the number of full-time workers in the household, and an increase in shopping trips with an increase in household size. This last observation is interesting because for many of the other person types, household size shows a decrease in the number of tours. This could mean that shopping responsibilities fall to driving-age students in larger households. Time availability coefficients did still come out to be significant.

TABLE 22: DRIVING-AGE STUDENT NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-25.0	6.17	-4.05
Number of iNM other discretionary tours	-6.04	1.36	-4.45
Number of iNM shopping tours	-17.1	5.56	-3.07
Number of iNM other maintenance tours	-12.3	3.21	-3.82
Number of iNM social tours	-7.34	2.44	-3.02

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM eatout tours	-16.5	3.81	-4.33
Total number of iNM tours is 1	0.982	0.463	2.12
Total number of iNM tours is 1 for 2016	0.163	0.388	0.42
Total number of iNM tours is 2p for 2016	0.503	0.920	0.55
Number of iNM escort tours * number of children	0.457	0.222	2.06
Number of iNM eatout tours * number of full-time workers	1.59	0.614	2.60
Number of iNM shopping tours * household size	0.525	0.289	1.82
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	8.11	2.23	3.64
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	1.76	1.11	1.58
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	3.37	1.20	2.80
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	4.12	1.33	3.09
Number of iNM social tours * log of total number of free hours between 7am to 10pm	1.25	0.981	1.27
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	1.28	0.512	2.50
Number of iNM shopping tours * shopping accessibility	0.553	0.392	1.41

The school person type estimation suffered from many of the same problems as the driving age student previously discussed in that not many coefficients turned out to be significant as demonstrated in Table 23. While the sample size was larger at 2,522 person-days, almost all person and household level variables were removed due to insignificance. Like driving-age students, the explanatory power of the model is concentrated around the alternative specific constants and the time availability variables.

TABLE 23 : SCHOOL NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-9.69	1.04	-9.33
Number of iNM other discretionary tours	-7.57	0.861	-8.79
Number of iNM shopping tours	-8.98	1.22	-7.36
Number of iNM other maintenance tours	-13.2	2.21	-5.99
Number of iNM social tours	-14.6	2.84	-5.16
Number of iNM eatout tours	-6.73	1.10	-6.15
Total number of iNM tours is 0 and no prior tours	1.11	0.152	7.29
Total number of iNM tours is 1	1.75	0.357	4.89
Total number of iNM tours is 2p	1.37	0.768	1.78
Total number of iNM tours is 2p for 2016	0.583	0.271	2.15
Total number of iNM tours is 1 and has a joint tour	-0.912	0.182	-5.02
Number of iNM escort tours * number of full-time workers	-0.158	0.0986	-1.60

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escort tours * household size	0.155	0.0510	3.04
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	2.05	0.281	7.30
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	1.96	0.435	4.51
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	3.38	0.927	3.64
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	0.992	0.397	2.50
Number of iNM social tours * log of total number of free hours between 7am to 10pm	2.49	0.843	2.96
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	1.76	0.290	6.06
Number of iNM other maintenance tours * number of consecutive free hours between 7am to 10pm	-0.147	0.0662	-2.23
Number of iNM escort tours * shopping accessibility	0.105	0.0454	2.32
Number of iNM other maintenance tours * shopping accessibility	0.177	0.102	1.74
Number of iNM social tours * other discretionary accessibility	0.208	0.103	2.02

The final person type for non-mandatory tour frequency estimation is preschool aged children. The non-mandatory tour frequency estimation results look very similar to the school person type above. Not many significant demographic variables and the explanatory power is again concentrated in available time windows and the alternative specific constants. An interesting observation for preschool estimation was that social tours decreased with an increase in the number of full-time workers in the household. This is likely explained by lack of free time for parents to take their preschool child to social engagements if they are working. Unsurprisingly, there is an increase in the number of escort tours for preschoolers as the number of household children grows as they are likely dragged along to transporting other children to their tours.

TABLE 24: PRESCHOOL NON-MANDATORY TOUR FREQUENCY ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Number of iNM escorting tours	-11.1	1.65	-6.68
Number of iNM other discretionary tours	-9.63	1.89	-5.10
Number of iNM shopping tours	-9.04	1.84	-4.92
Number of iNM other maintenance tours	-19.4	5.43	-3.57
Number of iNM social tours	-6.60	1.55	-4.25
Number of iNM eatout tours	-17.1	6.25	-2.74
Total number of iNM tours is 0 and no prior tours	3.25	0.220	14.77
Total number of iNM tours is 1	2.46	0.333	7.39
Total number of iNM tours is 2p	2.68	0.562	4.76
Total number of iNM tours is 1 for 2016	0.689	0.236	2.92

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Total number of iNM tours is 2p for 2016	0.493	0.307	1.61
Number of iNM escort tours * number of children	0.345	0.0631	5.47
Number of iNM social tours * number of full-time workers	-0.799	0.360	-2.22
Number of iNM escort tours * log of total number of free hours between 7am to 10pm	3.14	0.592	5.30
Number of iNM shopping tours * log of total number of free hours between 7am to 10pm	1.66	0.600	2.76
Number of iNM other maintenance tours * log of total number of free hours between 7am to 10pm	4.73	1.92	2.46
Number of iNM eatout tours * log of total number of free hours between 7am to 10pm	4.92	2.27	2.17
Number of iNM other discretionary tours * log of total number of free hours between 7am to 10pm	2.06	0.626	3.29
Number of iNM shopping tours * shopping accessibility	0.146	0.0714	2.05
Number of iNM other maintenance tours * shopping accessibility	0.264	0.104	2.53
Number of iNM other discretionary tours * other discretionary accessibility	0.0633	0.0453	1.40
Number of iNM social tours * other discretionary accessibility	0.202	0.0949	2.13

As discussed in this section, many trends were found across all person types estimated for individual non-mandatory tour frequency. These include large negative alternative specific constants decreasing many tours of the same purpose counteracted by positive constants on the overall number of tours, time availability variables being very important, and an increase in the tour rate in 2016 compared to 2022.

The driving-age student, school, and preschool person types lacked rich household and person specific terms seen in the adult person types. This is likely due to their day patterns being driven more by the schedules of other members of the household which is difficult to tease out in the estimation process, especially with the statistics for child person types being much lower than that for the adults.

5.11 EXTERNAL NON-MANDATORY TOUR IDENTIFICATION MODEL

The external non-mandatory tour identification model predicts which non-mandatory tours have a primary destination activity that is out of the region, in which case the primary destination of the tour would be an external station, predicted by the External Non-Mandatory Tour Destination Choice Model.

Data used to estimate this model includes 24,924 non-mandatory tours in the combined 2016 and 2022 SANDAG household travel survey data. Of these tours, 282 tours chose a primary destination outside the region. Non-mandatory tours are tabulated by the closest external station number to the tour origin in Table 25, and by tour purpose in Table 26. Figure 7 shows the share of external tours by distance to the closest external station.

TABLE 25: NON-MANDATORY TOURS BY CLOSEST EXTERNAL STATION AND EXTERNAL CHOICE

STATION NUMBER	EXTERNAL	INTERNAL	TOTAL
24323	2	0	2
24324	5	11	16
24325	3	74	77
24326	96	3929	4025
24327	36	2038	2074
24328	107	15292	15399
24329	32	3188	3220
24330	1	110	111
All	282	24642	24924

TABLE 26: NON-MANDATORY TOURS BY PURPOSE AND EXTERNAL CHOICE

STATION NUMBER	EXTERNAL	INTERNAL	TOTAL
Escort	23	4084	4107
Shopping	34	4265	4299
Other Maintenance	140	6721	6861
Eating out	23	2451	2474
Social/Visiting	12	680	692
Other discretionary	50	6441	6491
All	282	24642	24924

Estimation results after dropping insignificant variables are shown in Table 27. The results are intuitive and follow closely the results for the external worker identification model, as follows.

- Distance to the nearest external station is negative, indicating decreasing probability to work externally the further away from an external station a worker lives.
- The logged size of the external station is positive and approximately 0.4. This indicates a non-linear relationship between the estimated traffic volume for non-mandatory travel at an external station and the probability of selection. It suggests that distance plays a

more significant role in the selection process for non-mandatory tours and could point to some inaccuracies in the estimation of station size that is input to the model.

- A positive coefficient was estimated if the worker lives within 2.5 miles of an external station, consistent with the observed share of external tours by distance in Figure 7.
- Working externally to the region is positively correlated with household income.
- The coefficient for 2016 data was insignificant and therefore dropped from the specification.
- There is a strongly negative constant for working external to the region for each tour purpose, reflecting the relatively low overall share of external non-mandatory tours.

The initial log likelihood of the model is -17276 and the final log likelihood is -1422. The model rho-squared is 0.9177.

FIGURE 7: SHARE OF EXTERNAL NON-MANDATORY TOURS BY DISTANCE TO CLOSEST EXTERNAL STATION

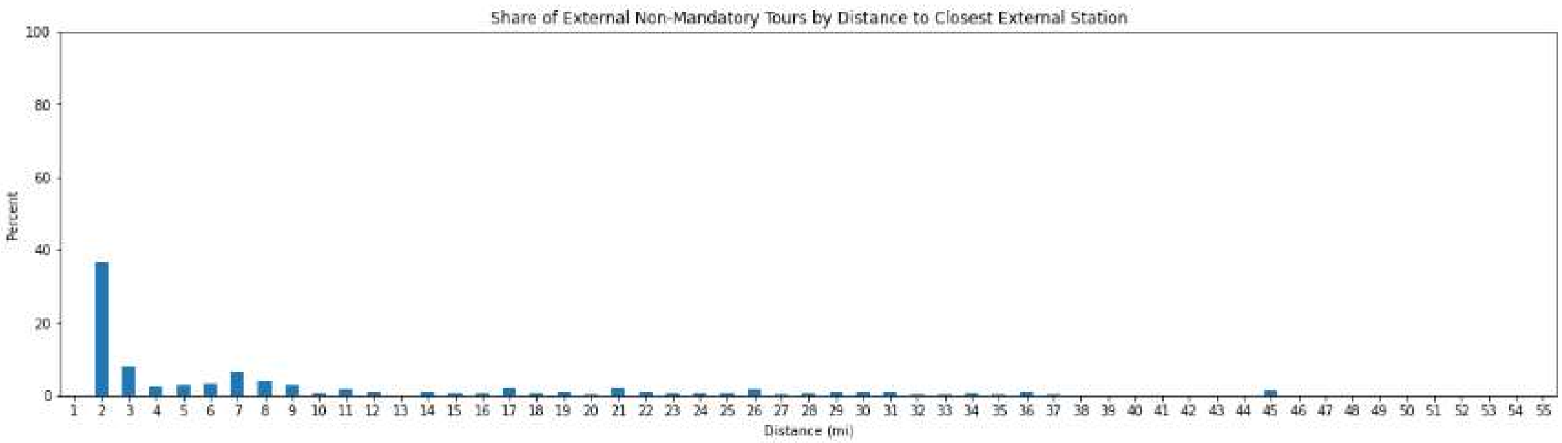


TABLE 27: EXTERNAL WORKER IDENTIFICATION ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Distance to the nearest external station	-0.0691	0.00847	-8.17
Logged size (Base-year or future year traffic count) of nearest external station	0.404	0.0856	4.72
Worker home is within 2.5 of an external station	2.65	0.433	6.12
Household income is less than $15k	-0.707	0.336	-2.10
Household income is greater than or equal to $15k and less than $50k	-0.253	0.140	-1.81
Household income is greater than or equal to $100k and less than $250k	0.118	0.186	0.64
Household income is greater than or equal to $250k	0.622	0.353	1.76
Alternative-specific constant for escort tours	-8.06	0.912	-8.84
Alternative-specific constant for shopping tours	-7.41	0.900	-8.24
Alternative-specific constant for other maintenance tours	-6.45	0.883	-7.31
Alternative-specific constant for other eating out tours	-7.28	0.902	-8.07
Alternative-specific constant for other social/visiting tours	-6.57	0.922	-7.12
Alternative-specific constant for other discretionary tours	-7.42	0.894	-8.30

5.12 EXTERNAL NON-MANDATORY TOUR DESTINATION CHOICE MODEL

The external non-mandatory tour destination choice model predicts which external station is the primary destination for external non-mandatory tours. There are 282 external non-mandatory tours used to estimate the model. The alternatives in the model are each external station.

Model coefficients are shown in Table 28. As expected, distance to the external station (capped at 20 miles is negatively signed. This coefficient is segmented by tour purpose. The mode choice logsum to the external station is positively signed but relatively small, and the logged size of the external station was constrained to 1.0. The initial log likelihood is -547 and the final log likelihood is -468. The rho-squared is 0.1452.

TABLE 28: EXTERNAL NON-MANDATORY TOUR DESTINATION CHOICE ESTIMATION RESULTS

VARIABLE DESCRIPTION	VALUE	STD ERR	T STAT
Distance to the external station (capped at 20 miles) – Escort tours	-0.356	0.148	-2.40
Distance to the external station (capped at 20 miles) – Shopping tours	-0.420	0.202	-2.07
Distance to the external station (capped at 20 miles) – Other Maintenance tours	-0.180	0.0470	-3.82
Distance to the external station (capped at 20 miles) – Eating out tours	-0.312	0.154	-2.03
Distance to the external station (capped at 20 miles) – Social/visiting tours	-0.273	0.0848	-3.22
Distance to the external station (capped at 20 miles) – Other discretionary tours	-0.345	0.133	-2.60
Mode choice logsum to the external station	0.270	0.0759	3.56
Logged size (Base-year or future year traffic count) of external station	1.0	NA	NA

5.13 INTERMEDIATE STOP FREQUENCY MODEL

The stop frequency model predicts the number of stops one makes on outbound and inbound segments of a tour. Tour types included in the stop frequency model are “work”, “school”, “university”, “escorting”, “shopping”, “other maintenance”, “eating out”, “social/ visiting”, “other discretionary”, and “at work”. The frequency of stops is broken into 16 different outcomes. No stops outbound or inbound, and each combination of zero to three stops outbound and inbound (for example zero out/one in, one out/ one in, etc.). For each of the models, the alternative of zero stops outbound or inbound is used as the base alternative. The utility for each alternative is composed of the Alternative Specific Constant, as well as the coefficient multiplied by the value for each of the significant variables in the model.

For each of the models, the same set of base variables were tested in sequence for significance. Additional tour specific variables were tested as well. The set of base variables included in the model development include the following variables:

- Personal and household level variables: age in cohorts (ages 5-18, 19-35, 45-55, 55-65, 65-79 and over 80), household income (less than $15,000, $15,000-$24,999, $25,000-$49,999, $50,000- $99,999, $100,000- $199,999, and greater than $200,000), employment status (worker, as well as disaggregated into part time or full time), student status, number of children in household as well as binary variables for presence of children in age brackets or disaggregated into preschool age, school age, driving, presence of part time worker, non-worker and retiree binary variables, household size broken into 2, 3 and 4 or larger.
- Land use variables: non-mandatory accessibility
- Tour variables: whether the tour is a joint tour (for non-mandatory tours), tour mode (SOV, HOV, school bus, taxi, transit, non-motorized, tour duration and tour distance, start time of tour broken into morning (3:00 am- 12:30 pm) or afternoon (12:30pm-5:30pm), number of non-work tours, work from home and telecommute.
- Year: data collected in the 2016 survey is identified with a specific variable to test the possible influence of Covid, temporal differences or survey differences between the two collection efforts

For escort, university, work and school tours, a model specific coefficient is utilized when appropriate to prevent stops for school escort tours. For at work tours, tour type variables of eat or maintenance as well as industry of employment was tested. For each of the variable described, when appropriate, a reference variable is chosen based on the distribution of the sample for the corresponding variable (for instance omitting one age cohort as a reference).

Each of the tour type models and estimation results will be discussed in more detail below.

Work

The work model is estimated using 10260 tours of these 5246 were 0 out 0 in tours. Due to the larger number of tours, disaggregation of the stops is possible. Tour stops were aggregated based on direction of stop, number of stops and variable significance. Due to a lack of significance of the highly disaggregated alternative specific constants at the higher frequency of stops, and the lack of data, the alternative specific constants were collapsed to create one constant termed 2 plus out, 2 plus in. As previously described, the base set of variables were sequentially tested for significance. The log-likelihood at null is -28,446.76 and is -16,073.92 at convergence. The rho squared is 0.4349. The final model coefficients and t-statistics are presented in Table 29. Significant variables included the following:

Variables that contribute in a positive way to the likelihood of at least one stop during work tours include ages 35-44, 45 to 54 and 55-64, distance to tour destination and tour duration, part time work status, non-mandatory accessibility, presence of school aged and preschool children in the household, telecommuting two to four days per week, and tour mode of HOV and SOV. Presence of non-workers in the household has a negative impact in the utility of an individual making at least one stop during the work tour.

TABLE 29: WORK STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Age 35 to 44 1 plus out 1 plus in work	0.286	0.0919	3.12
Age 45 to 54 1 plus out 1 plus in work	0.350	0.0972	3.60
Age 55 to 64 0 out 1 plus in work	0.174	0.0642	2.71
Age 55 to 64 1 plus out 0 in work	0.219	0.0949	2.31
Age 55 to 64 1 plus out 1 plus in work	0.314	0.118	2.67
Distance to tour destination 0 out 1 plus in work	0.00753	0.00238	3.17
Distance to tour destination 1 plus out 0 in work	0.0122	0.00314	3.88
Distance to tour destination 1 plus out 1 plus in work	0.0104	0.00327	3.17
Is part time worker 0 out 1 plus in work	0.286	0.0757	3.77
Is part time worker 1 plus out 1 plus in work	0.330	0.112	2.95
Non-mandatory accessibility 1 plus out 0 in work	0.0556	0.0197	2.83
Presence of non-worker in household 0 out 1 in work	-0.250	0.0899	-2.78
Presence of non-worker in household 0 out 2 in work	-0.390	0.150	-2.59
Presence of non-worker in household 0 out 3 in work	-0.703	0.214	-3.29
Presence of non-worker in household 1 plus out 1 plus in work	-0.464	0.122	-3.79
Preschool child in household 1 out 0 in work	0.492	0.106	4.64
Preschool child in household 1 plus out 1 plus in work	0.434	0.103	4.21
Presence of preschool child in household 2 plus out 0 in work	0.514	0.173	2.96
School aged child in household 1 out 0 in work	0.523	0.0904	5.78
School aged child in household 1 out 1 in work	0.282	0.116	2.44
School aged child in household 1 out 2 in work	0.511	0.167	3.07
School aged child in household 1 out 3 in work	0.515	0.194	2.65
School aged child in household 2 plus out 0 in work	0.715	0.147	4.86
School aged child in household 2 plus out 1 plus in work	0.395	0.144	2.75
Telecommute 2 to 4 days per week 1 plus out 1 plus in work	0.601	0.134	4.48

	VALUE	STD ERR	T STAT
Tour duration in hours 0 out 1 in work	0.0968	0.0113	8.53
Tour duration in hours 0 out 2 in work	0.190	0.0183	10.42
Tour duration in hours 0 out 3 in work	0.337	0.0248	13.59
Tour duration in hours 1 out 0 in work	0.0667	0.0142	4.70
Tour duration in hours 1 out 1 in work	0.165	0.0198	8.33
Tour duration in hours 1 out 2 in work	0.255	0.0312	8.17
Tour duration in hours 1 out 3 in work	0.325	0.0384	8.47
Tour duration in hours 2 out 1 in work	0.196	0.0383	5.12
Tour duration in hours 2 plus out 0 in work	0.0594	0.0238	2.50
Tour duration in hours 2 plus out 2 plus in work	0.261	0.0373	7.00
Tour duration in hours 3 out 1 in work	0.150	0.0696	2.16
Tour mode HOV 0 out 1 in work	1.75	0.132	13.33
Tour mode HOV 0 out 2 in work	2.43	0.214	11.31
Tour mode HOV 0 out 3 in work	3.08	0.292	10.58
Tour mode HOV 1 out 0 in work	1.88	0.153	12.30
Tour mode HOV 1 out 1 in work	2.90	0.187	15.57
Tour mode HOV 1 out 2 in work	3.01	0.221	13.65
Tour mode HOV 1 out 3 in work	3.45	0.246	14.00
Tour mode HOV 2 out 0 in work	2.36	0.204	11.56
Tour mode HOV 2 out 1 in work	2.90	0.250	11.60
Tour mode HOV 2 plus out 2 plus in work	3.70	0.250	14.83
Tour mode HOV 3 out 0 in work	2.37	0.287	8.25
Tour mode HOV 3 out 1 in work	3.14	0.406	7.75
Tour mode SOV 0 out 1 in	0.666	0.111	6.01
Tour mode SOV 0 out 2 in	0.849	0.198	4.29
Tour mode SOV 0 out 3 in	1.02	0.281	3.63
Tour mode SOV 1 plus out 0 in	0.321	0.130	2.46
Tour mode SOV 1 plus out 1 plus in	0.454	0.163	2.78
ASC 0 out 1 in	-2.84	0.154	-18.51
ASC 0 out 2 in	-5.07	0.268	-18.92
ASC 0 out 3 in	-7.36	0.386	-19.06
ASC 1 out 0 in	-4.00	0.352	-11.36
ASC 1 out 1 in	-5.34	0.263	-20.33
ASC 1 out 2 in	-7.29	0.378	-19.28
ASC 1 out 3 in	-8.64	0.466	-18.55

	VALUE	STD ERR	T STAT
ASC 2 out 0 in	-5.71	0.404	-14.12
ASC 2 out 1 in	-7.15	0.440	-16.28
ASC 3 out 0 in	-6.78	0.425	-15.95
ASC 3 out 1 in	-8.19	0.756	-10.84
ASC 2 plus out 2 plus in	-9.53	0.448	-21.24

School

The school stop frequency model is estimated using 2,178 tours, of which 1,119 tours has zero stops in either direction. Alternatives were specified for 0 stops out/1 plus stops in, 1 plus stops out/ 0 stops in, and 1 plus stops out/ 1 plus stops in for most coefficients. The alternative specific constant is estimated for each possible combination of stop alternatives from 0 to 3 stops out and 0 to 3 stops in. The final model log likelihood at null parameters is -6038.70, and is -3,345.39 at convergence. The rho squared is 0.4460. The resulting model and coefficients is presented in Table 30.

Each of the variables described in the base set of variables above were tested for significance. Significant variables were retained in the model. The indicator for 2016 survey data, non-mandatory accessibility, presence of non-workers in the household, and presence of school aged child/children in the household all make a positive contribution to the likelihood of an individual making at least one stop on a school tour. Conversely, number of non-work tours, and non-motorized tour modes contribute negatively to the likelihood of at least one stop on the tour.

TABLE 30: SCHOOL STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
2016 survey 0 out 1 plus in	0.713	0.147	4.84
2016 survey 1 plus out 0 in	0.770	0.284	2.71
2016 survey 1 plus out 1 plus in	1.00	0.283	3.54
Non-mandatory accessibility 0 out 1 plus in	0.0853	0.0351	2.43
Non-mandatory accessibility 1 plus out 1 plus in	0.131	0.0584	2.24
Presence of non-worker in household 1 plus out 1 plus in	0.368	0.169	2.17
Number of non-work tours 0 out 1 plus in	-0.350	0.112	-3.14
Number of non-work tours 1 plus out 1 plus in	-0.606	0.207	-2.93
School child in household 0 out 1 plus in	0.267	0.0998	2.67

	VALUE	STD ERR	T STAT
School child in household 1 plus out 0 in	0.799	0.184	4.33
School child in household 1 plus out 1 plus in	0.871	0.172	5.07
Tour mode is non-motorized 0 out 1 plus in	-1.13	0.222	-5.07
Tour mode is non-motorized 1 plus out 0 in	-2.23	0.720	-3.09
Tour mode is non-motorized 1 plus out 1 plus in	-3.11	1.01	-3.08
Tour modeis school bus 0 out 1 plus in	-1.72	0.471	-3.65
ASC 0 out 1 in	-2.27	0.499	-4.55
ASC 0 out 2 in	-3.15	0.503	-6.26
ASC 0 out 3 in	-3.82	0.508	-7.52
ASC 1 out 0 in	-3.23	0.301	-10.71
ASC 1 out 1 in	-4.84	0.863	-5.61
ASC 1 out 2 in	-5.65	0.871	-6.49
ASC 1 out 3 in	-6.06	0.878	-6.90
ASC 2 out 0 in	-4.93	0.356	-13.86
ASC 2 out 1 in	-7.31	0.927	-7.88
ASC 2 out 2 in	-7.44	0.937	-7.95
ASC 2 out 3 in	-8.29	1.03	-8.02
ASC 3 out 0 in	-5.98	0.456	-13.12
ASC 3 out 1 in	-8.00	0.992	-8.06
ASC 3 out 2 in	-8.69	1.11	-7.83
ASC 3 out 3 in	-8.69	1.11	-7.83

University

The University stop frequency model is estimated using 764 university trips, of which

309 were 0 stops out/ 0 stops in trips. Due to data limitations, the alternatives for most coefficients in this model were collapsed to one alternative of 1 plus stops. This encompasses stops in either outbound of inbound directions. The alternative specific constants were collapsed for 0 out 1-3 inbound stops to create one term named “0 out 1 plus in”, and were specified for each remaining alternative combination of 1 to 3 stops outbound and 0 to 3 stops inbound.

The model log likelihood at null parameters is -2,115.49, and the final model log likelihood at convergence is -1,574.35. The rho squared is 0.2558. The resulting model and coefficients are reported in Table 31.

As with previous models, each of the variables described above were tested. Significant variables were retained in the model. Presence of preschool child in the household, tour duration and the modes of HOV and transit each were positively correlated with the likelihood of at least one stop on university tours. Non-motorized tour mode choice has a negative contribution to the utility of a tour with at least one stop.

TABLE 31: UNIVERSITY STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Presence of preschool child in household 1 plus stops	0.735	0.328	2.24
Tour duration in hours 1 plus stops	0.184	0.0254	7.25
Tour mode is HOV 1 plus stops	0.621	0.214	2.91
Tour mode is non-motorized 1 plus stops	-0.715	0.323	-2.22
Tour mode is transit 1 plus stops	0.436	0.195	2.23
ASC 0 out 1 plus in	-2.94	0.185	-15.91
ASC 1 out 0 in	-2.96	0.213	-13.89
ASC 1 out 1 in	-2.99	0.215	-13.96
ASC 1 out 2 in	-3.96	0.272	-14.55
ASC 1 out 3 in	-4.11	0.285	-14.41
ASC 2 out 0 in	-3.65	0.249	-14.66
ASC 2 out 1 in	-4.57	0.335	-13.65
ASC 2 out 2 in	-5.11	0.414	-12.33
ASC 2 out 3 in	-5.11	0.414	-12.33
ASC 3 out 0 in	-3.84	0.262	-14.63
ASC 3 out 1 in	-4.42	0.317	-13.95
ASC 3 out 2 in	-4.98	0.392	-12.69
ASC 3 out 3 in	-4.86	0.374	-12.99

Escorting

The escorting model is estimated using 4,107 escorting trips, of which 2,707 trips were conducted with zero stops in either direction. Due to data limitations again with this model, the alternatives for most coefficients were collapsed into one coefficient alternative with the terminology of "1 plus stops". Each alternative is estimated in a disaggregate form for the alternative specific constants. The log likelihood at null parameters is -11,387.02, and is -5,262.89 at convergence. The rho squared for the final model is 0.5378. The resulting model coefficients can be seen in Table 32.

Several variables are positively correlated with the likelihood of making at least one stop on an escort tour. These include a person being within the ages of 5 to 18 years or 65 to 79 years, distance to tour destination, a person having student status, non-mandatory accessibility, and if the tour is started in the afternoon (12:30-5:30pm). Household income less than $15,000, household sizes of 3 and 4 or larger, and a non-motorized tour mode contribute negatively to the utility of one or more stops during an escort tour.

Additionally, availability conditions were added to the stop frequency model of escort and were retained to enforce the condition prohibiting stops on half-tours that include school escorting. This is due to the accounting of school escorting trips elsewhere in the model. Lastly, the coefficient "Escort half tour without escortee for school escort tour 1 plus stops" reflects the increased likelihood of an individual stopping on the half tour that does not include an escorting responsibility for school escort tours.

TABLE 32: ESCORT STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Age is between 5 and 18 1 plus stops	0.373	0.140	2.67
Age is between 65 and 79 1 plus stops	0.390	0.129	3.01
Distance to tour destination 1 plus stops	0.0354	0.00506	7.00
Household income less 15 1 plus stops	-0.386	0.158	-2.45
Household size is 3 1 plus stops	-0.606	0.107	-5.67
Household size is 4 or more 1 plus stops	-0.584	0.0973	-6.00
Student status 1 plus stops	0.216	0.0973	2.22
Half tour without escortee for school escort tour 1 plus stops	0.591	0.116	5.10
Non-mandatory accessibility 1 plus stops	0.0692	0.0237	2.92
Start time is afternoon (12:30pm-5:30pm) 1 plus stops	0.176	0.0701	2.52
Tour mode is non-motorized 1 plus stops	-1.24	0.187	-6.64

	VALUE	STD ERR	T STAT
ASC 0 out 1 in	-2.49	0.344	-7.25
ASC 0 out 2 in	-3.67	0.351	-10.46
ASC 0 out 3 in	-4.32	0.360	-11.99
ASC 1 out 0 in	-2.93	0.346	-8.47
ASC 1 out 1 in	-3.95	0.356	-11.10
ASC 1 out 2 in	-4.79	0.374	-12.79
ASC 1 out 3 in	-5.30	0.395	-13.42
ASC 2 out 0 in	-4.19	0.358	-11.69
ASC 2 out 1 in	-5.06	0.384	-13.17
ASC 2 out 2 in	-6.58	0.509	-12.92
ASC 2 out 3 in	-6.44	0.491	-13.12
ASC 3 out 0 in	-4.42	0.362	-12.20
ASC 3 out 1 in	-5.09	0.385	-13.21
ASC 3 out 2 in	-6.12	0.455	-13.46
ASC 3 out 3 in	-6.22	0.465	-13.38

Shopping

The shopping stop frequency model is estimated using 4,791 tours, of which 2,338 tours had zero stops in either direction. The alternative specific constants were each estimated separately; however all other coefficients were estimated with one alternative to zero stops as the "1 plus stops" alternative. The log likelihood at null parameters is -13,282.47 and is -8,325.44 at convergence. The rho squared is 0.3732.

The resulting model and significant coefficients can be seen in Table 33. Coefficients that have a positive impact on the utility of stops during shopping tours include distance to tour destination, 2016 survey data, non-mandatory accessibility, presence of school child or children in the household, tour duration and tour mode is HOV. Joint tours, tours conducted by workers, number of non-work tours, and non-motorized and transit tour modes are associated with a decreased likelihood to make a stop on either outbound or inbound portions of an escort tour.

TABLE 33: SHOPPING STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Distance to tour destination 1 plus stops	0.0912	0.00964	9.45
Tour is a joint tour 1 plus stops	-0.938	0.116	-8.08
2016 survey 1 plus stops	0.188	0.0677	2.78
Worker 1 plus stops	-0.298	0.0634	-4.70
Non-mandatory accessibility 1 plus stops	0.0358	0.0195	1.83
Number of non-work tours 1 plus stops	-0.280	0.0366	-7.63
School child in household 1 plus stops	0.276	0.0932	2.96
Tour duration in hours 1 plus stops	0.0919	0.0223	4.13
Tour mode is HOV 1 plus stops	0.858	0.0799	10.74
Tour mode is non-motorized 1 plus stops	-1.07	0.132	-8.14
Tour mode is transit 1 plus stops	-0.496	0.187	-2.66
ASC 0 out 1 in	-1.99	0.282	-7.06
ASC 0 out 2 in	-3.00	0.288	-10.44
ASC 0 out 3 in	-3.82	0.298	-12.81
ASC 1 out 0 in	-1.93	0.282	-6.84
ASC 1 out 1 in	-3.08	0.288	-10.69
ASC 1 out 2 in	-3.95	0.300	-13.16
ASC 1 out 3 in	-4.45	0.313	-14.20
ASC 2 out 0 in	-2.90	0.287	-10.11
ASC 2 out 1 in	-3.88	0.299	-12.98
ASC 2 out 2 in	-5.06	0.339	-14.93
ASC 2 out 3 in	-5.06	0.339	-14.93
ASC 3 out 0 in	-3.47	0.293	-11.86
ASC 3 out 1 in	-4.04	0.302	-13.37
ASC 3 out 2 in	-5.36	0.358	-14.99
ASC 3 out 3 in	-5.79	0.394	-14.72

Other Maintenance

The other maintenance stop frequency model is estimated using 7,291 tours, of which 3,086 tours were tours with 0 stops in each direction. Like in previous models, the alternative specific constants were estimated for each stop frequency combination, but the remaining coefficients were collapsed into one alternative of “1 plus stops” to the zero-stop frequency comparison. The log likelihood at null parameters is -20,214.94 and at convergence it is -14,294.38. The rho squared is 0.2929. The model coefficients can be seen in Table 34.

Number of children in the household, tour start time in the morning (3:00am to 12:30pm) and afternoon (12:30pm to 5:30pm), tour duration, and tour mode of HOV all have a positive correlation with the likelihood of at least one stop on a tour associated with other maintenance activities. Household size of 4 or more people, joint tour, number of non-work tours and a tour mode that is non-motorized all have a negative contribution to the utility of number of stops.

TABLE 34: OTHER MAINTENANCE STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Distance to tour destination 1 plus stops	0.0292	0.00298	9.81
Household size is 4 or more 1 plus stops	-0.381	0.112	-3.39
Tour is joint tour 1 plus stops	-1.14	0.115	-9.90
Number of children in household 1 plus stops	0.236	0.0508	4.64
Number of non-work tours 1 plus stops	-0.233	0.0315	-7.40
Start time is afternoon (12:30pm-5:30pm) 1 plus stops	0.374	0.100	3.73
Start time is morning (3:00am-12:30pm) 1 plus stops	0.418	0.0978	4.27
Tour duration in hours 1 plus stops	0.0305	0.00913	3.34
Tour mode is HOV 1 plus stops	1.08	0.0673	16.09
Tour mode is non-motorized 1 plus stops	-0.620	0.119	-5.21
ASC 0 out 1 in	-1.56	0.114	-13.67
ASC 0 out 2 in	-2.38	0.119	-20.07
ASC 0 out 3 in	-2.61	0.121	-21.61
ASC 1 out 0 in	-2.31	0.118	-19.52
ASC 1 out 1 in	-2.94	0.125	-23.55
ASC 1 out 2 in	-3.56	0.136	-26.17
ASC 1 out 3 in	-3.36	0.132	-25.52
ASC 2 out 0 in	-3.19	0.129	-24.81
ASC 2 out 1 in	-3.80	0.142	-26.74
ASC 2 out 2 in	-4.39	0.163	-26.89
ASC 2 out 3 in	-4.09	0.151	-27.03
ASC 3 out 0 in	-3.57	0.136	-26.21
ASC 3 out 1 in	-4.03	0.149	-27.00
ASC 3 out 2 in	-4.60	0.173	-26.54
ASC 3 out 3 in	-4.16	0.154	-27.03

Eating out

The eating out stop frequency model is estimated using 3,007 tours, of which 1,988 tours have zero stops in both directions. Variables described above were each tested for significance, and were retained if significant. Like many of the stop frequency models, due to data limitations, the alternatives for the eating out model were condensed into zero stops, and one plus stops. The log likelihood at null parameters is -8,337.17, and is -3,840.98 at convergence. The rho squared is 0.5393. The resulting model can be seen in Table 35.

Older age individuals (between 65 and 79, and 80 and older) have a higher likelihood of making at least one stop on an eating out tour, with those 80 and over being more likely than those between 65 and 79. Additionally, non-mandatory accessibility, tour duration and distance, and a tour mode of HOV also contribute positively to the utility of making at least one stop. Joint tours, number of non-work tours, and a non-motorized tour mode each have a negative correlation with stops on eat out tours.

TABLE 35: EATING OUT STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Age between 65 and 79 1 plus stops	0.257	0.103	2.48
Age 80 or older 1 plus stops	0.613	0.298	2.06
Joint tour 1 plus stops	-0.799	0.119	-6.70
Non-mandatory accessibility 1 plus stops	0.0385	0.0245	1.57
Number of nonwork tours 1 plus stops	-0.143	0.0499	-2.86
Tour duration in hours 1 plus stops	0.0712	0.0247	2.88
Distance to tour destination 1plus stops	0.0419	0.00706	5.93
Tour mode is HOV 1 plus stops	0.728	0.0935	7.79
Tour mode is non-motorized 1 plus stops	-0.644	0.154	-4.17
ASC 0 out 1 in	-2.72	0.363	-7.49
ASC 0 out 2 in	-4.04	0.375	-10.77
ASC 0 out 3 in	-4.84	0.395	-12.26
ASC 1 out 0 in	-2.87	0.364	-7.88
ASC 1 out 1 in	-3.92	0.373	-10.49
ASC 1 out 2 in	-5.23	0.411	-12.74
ASC 1 out 3 in	-5.12	0.405	-12.62
ASC 2 out 0 in	-4.07	0.376	-10.83
ASC 2 out 1 in	-5.31	0.415	-12.81

	VALUE	STD ERR	T STAT
ASC 2 out 2 in	-5.88	0.453	-12.98
ASC 2 out 3 in	-7.06	0.615	-11.48
ASC 3 out 0 in	-4.81	0.394	-12.22
ASC 3 out 1 in	-6.05	0.469	-12.91
ASC 3 out 2 in	-6.37	0.504	-12.65
ASC 3 out 3 in	-7.35	0.680	-10.81

Social/visiting

The social/visiting stop frequency is estimated using 795 tours. Of these tours, 481 tours were comprised of zero stops in both directions. The same modeling procedure of testing the base set of variables for significance in the model is followed. The log likelihood with null parameters is -2,204.21, and is -1,125.68 at convergence. The rho squared for the final model is 0.4893.

As seen in the model results presented in Table 36, tour distance and duration as well as tour mode of HOV are all positively associated with at least one stop during a social/visiting tour. Younger age (age 5 to 18 years), a joint tour, and the number of non-work tours all have a negative influence on the utility of at least one stop.

TABLE 36: SOCIAL/ VISITING STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Age is between 5 and 18 1 plus stops	-0.771	0.380	-2.03
Distance to tour destination 1 plus stops	0.0432	0.00981	4.40
Tour duration in hours 1 plus stops	0.0731	0.0362	2.02
Tour is a joint tour 1 plus stops	-0.777	0.265	-2.93
Non-mandatory accessibility 1 plus stops	0.0400	0.0422	0.95
Number of nonwork tours 1 plus stops	-0.209	0.0922	-2.27
Tour mode is HOV 1 plus stops	0.524	0.169	3.11
ASC 0 out 1 in	-2.67	0.647	-4.13
ASC 0 out 2 in	-3.95	0.671	-5.88
ASC 0 out 3 in	-5.09	0.740	-6.88
ASC 1 out 0 in	-2.47	0.645	-3.84

	VALUE	STD ERR	T STAT
ASC 1 out 1 in	-3.82	0.667	-5.72
ASC 1 out 2 in	-4.84	0.719	-6.74
ASC 1 out 3 in	-5.09	0.740	-6.88
ASC 2 out 0 in	-3.35	0.656	-5.11
ASC 2 out 1 in	-5.43	0.778	-6.98
ASC 2 out 2 in	-6.35	0.952	-6.67
ASC 2 out 3 in	-5.65	0.810	-6.98
ASC 3 out 0 in	-4.84	0.719	-6.74
ASC 3 out 1 in	-5.09	0.740	-6.88
ASC 3 out 2 in	-7.04	1.19	-5.94
ASC 3 out 3 in	-7.04	1.19	-5.94

Other Discretionary

The stop frequency mode predicting stops on other discretionary tours is estimated using 7,153 tours. Of these tours, 4,406 tours were tours that had zero stops in both directions. The log likelihood with null parameters for the model is -19,832.33, and it is -9,989.99 at convergence. The rho squared of the final model is 0.4963.

As shown in Table 37, older individuals, larger distances to tour destination, higher non-mandatory accessibilities, presence of preschool child or children in the household, tour start time in the morning (3:00am-12:30pm) or afternoon (12:30pm- 5:30pm), and a tour mode of HOV contribute in a positive manner in the utility of one plus stops for tours conducting other discretionary purposes. Younger ages (5 to 18 years and 19 to 35 years), presence of retiree in the household, number of non-work tours, and non-motorized tour modes all have a negative correlation with at least one stop on an “other discretionary” tour.

TABLE 37: OTHER DISCRETIONARY STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Age is between 19 and 35 1 plus stops	-0.285	0.0788	-3.61
Age is between 5 18 1 plus stops	-0.566	0.136	-4.15
Age is between 65 79 1 plus stops	0.170	0.0695	2.44
Distance to tour destination 1 plus stops	0.0375	0.00393	9.54

	VALUE	STD ERR	T STAT
Presence of retiree in household 1 plus stops	-0.217	0.0795	-2.73
Tour is a joint tour 1 plus stops	-0.899	0.105	-8.58
Non-mandatory accessibility 1 plus stops	0.0691	0.0168	4.12
Number of nonwork tours 1 plus stops	-0.264	0.0322	-8.21
Presence of preschool child in household 1 plus stops	0.269	0.109	2.48
Start time is afternoon (12:30pm-5:30pm) 1 plus stops	0.182	0.0767	2.37
Start time is morning (3:00am-12:30pm) 1 plus stops	0.298	0.0651	4.58
Tour mode is HOV 1 plus stops	0.906	0.0674	13.45
Tour mode is non-motorized 1 plus stops	-1.51	0.0891	-16.94
ASC 0 out 1 in	-2.49	0.248	-10.04
ASC 0 out 2 in	-3.63	0.253	-14.35
ASC 0 out 3 in	-4.21	0.259	-16.26
ASC 1 out 0 in	-2.75	0.248	-11.09
ASC 1 out 1 in	-3.52	0.252	-13.96
ASC 1 out 2 in	-4.61	0.266	-17.35
ASC 1 out 3 in	-4.85	0.271	-17.89
ASC 2 out 0 in	-4.02	0.257	-15.66
ASC 2 out 1 in	-4.75	0.269	-17.67
ASC 2 out 2 in	-5.90	0.314	-18.79
ASC 2 out 3 in	-6.02	0.322	-18.71
ASC 3 out 0 in	-4.43	0.262	-16.87
ASC 3 out 1 in	-5.22	0.282	-18.51
ASC 3 out 2 in	-5.75	0.306	-18.83
ASC 3 out 3 in	-6.06	0.325	-18.67

At work

The at work model is estimated using 2,003 tours, of which 1,567 of these tours had zero stops in both directions. Like other models, the limited data necessitated the collapsing of stops for tours into one alternative, specifically 1 plus stops. The base set of variables were again used in this model to test for significance. In addition, tour type for the at work tour is utilized and is “eat out” with a comparison of “maintenance,” and 11 industry types were included in the model. These include accommodation, business services, construction, education, entertainment, food services, government, health, manufacturing, management services, and military. The log likelihood at null parameters is -5,553.50, and is -1,534.54 at convergence. The rho squared of the final model is 0.7237. The resulting model is presented in Table 38.

Records from 2016 were more likely to have at least one stop during the at work tour. In similar fashion, number of children aged 5 to 15 years old, and tour duration also contributed in a positive manner to the utility of at least one stop on an at work tour. Non-motorized modes as well as at work tour purposes of eating out were correlated with a lower likelihood of stopping during the tour.

TABLE 38: AT WORK STOP FREQUENCY ESTIMATION RESULTS

	VALUE	STD ERR	T STAT
Survey data from 2016 (1 plus stops)	0.576	0.211	2.72
Number of children aged 5 to 15 in household (1 plus stops)	0.284	0.0909	3.13
Tour duration (1 plus stops)	0.834	0.0668	12.47
Mode is non-motorized (1 plus stops)	-1.14	0.195	-5.85
Tour type is eating out (1 plus stops)	-1.51	0.147	-10.31
ASC 0 out 1 in	-3.09	0.242	-12.80
ASC 0 out 2 in	-4.90	0.292	-16.77
ASC 0 out 3 in	-5.77	0.361	-16.00
ASC 1 out 0 in	-3.55	0.248	-14.31
ASC 1 out 1 in	-5.39	0.325	-16.58
ASC 1 out 2 in	-6.54	0.469	-13.96
ASC 1 out 3 in	-6.72	0.503	-13.37
ASC 2 out 0 in	-5.56	0.340	-16.36
ASC 2 out 1 in	-6.25	0.422	-14.82
ASC 2 out 2 in	-7.64	0.744	-10.27
ASC 2 out 3 in	-7.24	0.622	-11.64
ASC 3 out 0 in	-5.77	0.361	-16.00
ASC 3 out 1 in	-6.95	0.550	-12.62
ASC 3 out 2 in	-8.33	1.03	-8.12
ASC 3 out 3 in	-6.54	0.469	-13.96

6.0 MODE CHOICE MODEL UPDATES AND ENHANCEMENTS

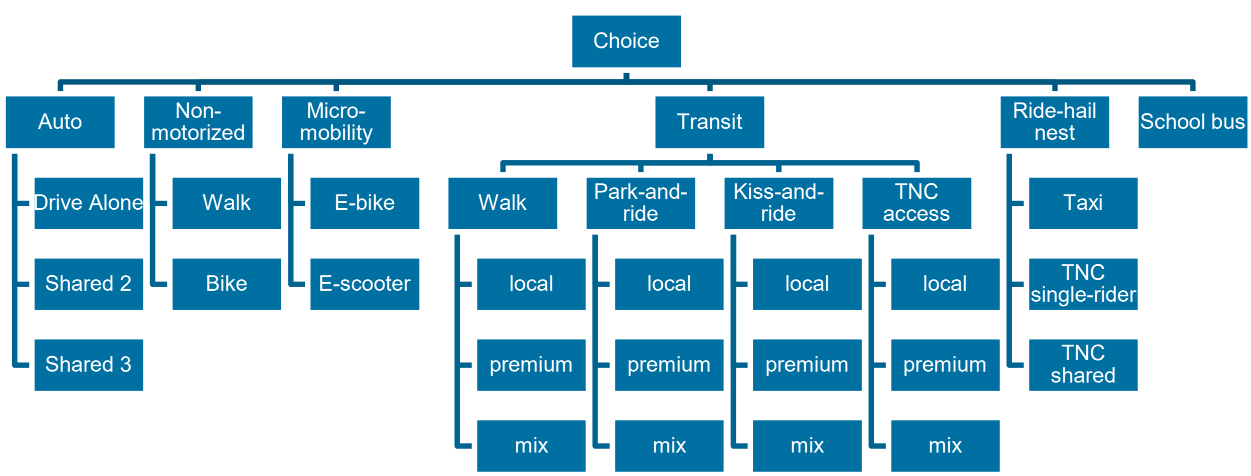
The mode choice model in ABM3 was revised and enhanced in several ways. New micro-mobility modes were added to the model, including e-scooter and e-bikes. The new model uses TAZs for transit skims (ABM2+ model used Transit Access Points) and overrides skimmed walk time to transit with the time from the MAZ to the nearest transit stop by type of transit service (local versus premium). This section describes the new models.

There are two mode choice models used to predict mode in ActivitySim:

- The tour mode level (upper-level choice),
- The trip mode level (lower-level choice conditional upon the upper-level choice).

The tour mode level can be thought of as a mode preference model, while the trip mode choice model can be thought of as a mode switching model. Tour mode choice is used to constrain stop location choice as well as trip mode choice. The modes for both models are the same, but the higher level of the nesting structure constrains lower-level decisions. Figure 8 shows the revised nesting structure for both tour and trip mode choice.

FIGURE 8: REVISED MODE CHOICE NESTING STRUCTURE



Tour modes are defined based on a set of rules pertaining to the combination of modes reported for each trip on the tour. The following set of rules are proposed, subject to modification and finalization based on observed combinations of trip modes on tours. Note that micromobility and ride-hail modes are constrained at that level, rather than the lower level trip modes.

1) If any mode is PNR-transit, the tour mode is PNR-transit
2) If any mode is KNR-transit, the tour mode is KNR-transit
3) If any mode is TNC-transit, the tour mode is TNC-transit
4) If any mode is walk-transit, the tour mode is walk-transit
5) If any mode is ride-hail (taxi, single pay TNC, shared TNC), the tour mode is ride-hail
6) If any mode is micromobility (e-scooter or e-bike), the tour mode is micromobility
7) If any mode is school bus, the tour mode is school bus.
8) If any mode is shared 3+, the tour mode is shared 3+
9) If any mode is shared 2+, the tour mode is shared 2+
10) If any mode is drive alone, the tour mode is drive alone
11) If any mode is bike, the tour mode is bike
12) If any mode is walk, the tour mode is walk

Note that although there are options in the nesting structure for transit sub-mode (local-only, premium-only, and mix), we do not calibrate the model for these modes specifically in mode choice. Instead we apply terms in mode choice that reflect lower disutility for using bus to access premium transit, and apply technology-specific constants in premium and mix modes to reflect preferences for bus rapid transit, light-rail, and commuter rail, all else being equal. A potential refinement of this approach would be to allow all modes to compete in transit path-building and attempt to address transit user preferences in path parameters as well as in mode choice instead of forcing the path-finder to find separate paths for each transit technology. However, we felt that this was beyond our current scope of work.

The tour mode choice model is based on the round-trip level-of-service (LOS) between the tour anchor location (home for home-based tours and work for at-work sub-tours) and the tour primary destination. The tour mode is chosen based on LOS variables for both directions according to the time periods for the tour departure from the anchor and the arrival back at the anchor. This is one of the fundamental advantages of the tour-based approach. For example, a commuter can have very attractive transit service in the a.m. peak period in the outbound direction, but if the return home time is in the midday or later at night, the commuter may prefer private auto due to lower off-peak transit service. The appropriate skim values for the tour mode choice are a function of the TAZ/MAZ of the tour origin and TAZ/MAZ of the tour primary destination. The mode choice model alternatives and skims are shown Table 39.

TABLE 39: MODE CHOICE ALTERNATIVES AND NETWORK LEVEL-OF-SERVICE VARIABLES

ALTERNATIVE	DESCRIPTION	NETWORK LEVEL-OF-SERVICE (SKIMS)*
Drive-alone	Single occupant auto	Auto time, distance by drive-alone (no HOV lanes) by three value of time bins
Shared 2	Auto with two occupants	Auto time, distance, and cost by shared-ride 2 (no HOV 3+ lanes) by three value of time bins
Shared 3+	Auto with three or more occupants	Auto time, distance, and cost by shared-ride 3+ by three value of time bins
Walk	Walk, skateboard, human powered scooter, wheelchair	Walk distance and time calculated across an all-streets network between MAZs within a certain distance threshold
Bike	Human powered bike	Bike logsums calculated across an all-streets network between MAZs and TAZs within certain distance thresholds
e-scooter	Electric scooter	e-scooter distance and time calculated across an all-streets network between MAZs within a certain distance threshold
e-bike	Electric bike	e-bike distances and logsums calculated across an all-streets network between MAZs and TAZs within certain distance thresholds
Walk-local	Transit by walk-access on both ends of the trip, local bus only in transit path	Walk local skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, walk access time, walk egress time

ALTERNATIVE	DESCRIPTION	NETWORK LEVEL-OF-SERVICE (SKIMS)*
Walk-premium	Transit by walk-access on both ends of the trip, premium transit only in transit path	Walk premium skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, walk access time, walk egress time
Walk-mix	Transit by walk-access on both ends of the trip, local bus and premium transit must be in transit path	Walk mix skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, walk access time, walk egress time
PNR-local	Transit by PNR-access on one end of the trip and walk on the other end, local bus only in transit path	PNR local skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, PNR access (or egress) time, walk egress (or access) time
PNR-premium	Transit by PNR-access on one end of the trip and walk on the other end, premium transit only in transit path	PNR premium skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, PNR access (or egress) time, walk egress (or access) time
PNR-mix	Transit by PNR-access on one end of the trip and walk on the other end, local bus and premium transit must be in transit path	PNR mix skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, PNR access (or egress) time, walk egress (or access) time

ALTERNATIVE	DESCRIPTION	NETWORK LEVEL-OF-SERVICE (SKIMS)*
KNR-local	Transit by KNR-access on one end of the trip and walk on the other end, local bus only in transit path	KNR local skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time
KNR-premium	Transit by KNR-access on one end of the trip and walk on the other end, premium transit only in transit path	KNR premium skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time
KNR-mix	Transit by KNR-access on one end of the trip and walk on the other end, local bus and premium transit must be in transit path	KNR mix skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time
TNC-local	Transit by KNR-access on one end of the trip and walk on the other end, local bus only in transit path	TNC local skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time
TNC-premium	Transit by KNR-access on one end of the trip and walk on the other end, premium transit only in transit path	TNC premium skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time

ALTERNATIVE	DESCRIPTION	NETWORK LEVEL-OF-SERVICE (SKIMS)*
TNC-mix	Transit by KNR-access on one end of the trip and walk on the other end, local bus and premium transit must be in transit path	TNC mix skims: in-vehicle by transit technology, total in-vehicle time, first wait, transfer wait, auxiliary walk time, transit fare, KNR access (or egress) time, walk egress (or access) time
Taxi	Traditional taxi mode	Auto time, distance, and cost by shared-ride 2 (no HOV 3+ lanes) by high value of time bin
TNC-single	Single-pay TNC mode, such as Uber or Lyft	Auto time, distance, and cost by shared-ride 2 (no HOV 3+ lanes) by high value of time bin
TNC-pool	Shared TNC mode, such as Uber Pool, Lyft Line or other micro-transit service	Auto time, distance, and cost by shared-ride 2 (no HOV 3+ lanes) by low value of time bin
School bus	Yellow bus	Auto time, distance, and cost by shared-ride 3+ by low value of time bin

* Notes on skims:

- Auto and transit skims are differentiated by five time of day periods and calculated across a planning network between TAZs.
- Auto skims are additionally differentiated by three value-of-time bins to represent travel time and cost heterogeneity. This is described more fully below.
- Auto costs include tolls on the auto network
- Transit network level-of-service skims are differentiated by access and egress mode as well as the transit technology used in the path. Walk-transit skims are based on walk access and walk egress (walk-transit-walk). Park-and-ride, kiss-and-ride, and TNC skims assume auto is used at the home end of the tour. Skims must be built by direction

for each time period (PNR-transit-walk and walk-transit-PNR, KNR-transit-walk and walk-transit-KNR, TNC-transit-walk and walk-transit-TNC).

- Walk to transit times are calculated from an all-streets network based on the MAZ centroid and the nearest stop(s), as differentiated by broad transit stop types (local, premium).

Travel time and cost heterogeneity

The mode choice model includes recommendations contained in research sponsored by the Strategic Highway Research Program (SHRP) C04 track on pricing and reliability[4]. The final report recommended a number of key features:

- Travel time heterogeneity: Sensitivity to travel time should be represented as a distribution reflecting personal preference and contextual conditions
- Continuous representation of income: Sensitivity to cost should ideally be represented as a continuous function of income rather than a global average
- Vehicle occupancy effects: Sensitivity to cost is also dependent upon the occupancy of the vehicle, but it is not a linear relationship.

Each person in the synthetic population draws a mandatory and a non-mandatory travel time sensitivity parameter from log-normal distribution with a mean of 1.0. The mandatory travel time sensitivity parameter is used to factor skimmed times identified in Table 39 for work and school tours; the non-mandatory travel time sensitivity factor will be used to factor skimmed times for all other tour purposes. These factors are used to represent heterogeneity in travel time sensitivity across the population. So the travel time sensitivity coefficient is:

$$\beta time_{,purpose,person} = \beta time_{purpose} * \ \beta time\_sensitivity_{person}$$

The cost coefficient is calculated based on a cost coefficient that varies by tour/trip purpose, the household income of the traveler, the mode occupancy, and an income exponent:

---

[4] Parsons Brinkerhoff, Northwestern University, Mark Bradley, University of California at Irvine, RSG, University of Texas at Austin, Frank Koppelman, GeoStats, SHRP2 C04 - Improving Our Understanding of How Highway Congestion and Pricing Affect Travel Demand, Transportation Research Board of the National Academies (TRB), The Second Strategic Highway Research Program (SHRP2) Capacity Research, April 2013

$\beta cost_{,purpose,person} = \beta cost_{purpose} / max(income,1000)^{income\_exponent} * occupancy\_factor$

The value-of-time is calculated as:

$\beta time_{,purpose,person} / \beta cost_{,purpose,person} * 60/100$

Because travel time sensitivity varies across the population, the value-of-time (derived from time and cost sensitivity) also varies across the population. The variance in sensitivity to time and cost provides greater sensitivity to pricing, rather than relying on a single value of time for all travelers or a value-of-time that only varies according to trip or tour purpose. It also recognizes the value of time is a function of household income but is not solely determined by household income. Note that the time sensitivity distribution applies to all components of time, not just auto time. After calculating the value of time for each person tour or trip, the tour or trip is associated with the appropriate auto skims for that value of time category. These calculations are performed in the ActivitySim tour and trip mode choice preprocessors.

Transit modes and utilities

The model specifies 4 access modes for transit – walk, park-and-ride (PNR), kiss-and-ride (KNR) and transportation network company or TNC. As mentioned above, skims must be built by direction for each time period; this is different from most trip-based models in which skims are built only in the AM peak period and assume auto can only be used on the access end of the transit trip. In activity-based models, skims are built in both directions for drive access – PNR-transit-walk and walk-transit-PNR; same with KNR.

Journey rules are used to calculate transit fare based on the modes used in the path. There are also overrides for transfer wait times specified at several commuter rail stations to compensate for errors in skimmed transfer wait time resulting from headway-based assignments.

The skimmed walk access and egress times are replaced by MAZ-stop times calculated across an all-streets network and written to an MAZ file that has one row for each MAZ and the time to the nearest stop, differentiated by type of stop (bus versus premium transit).

Walk mode

Walk mode times will be based on MAZ distances within walking distance calculated across an all-streets network.

Bicycle mode treatment and bike comfort level

A bike route choice logsum and bike distance is determined by path choice model across the all-streets network for MAZs and TAZs within bike distance.

Micromobility modes

Micromobility modes include e-scooters and e-bikes. An interesting aspect of micromobility is that these modes can be privately owned or shared. If privately owned, we assume the traveler can use the mode for any tour starting at home for free, with no access time. If shared, we include the cost of the mode in the utility for using it, as well as an access time that varies according to the origin zone. These access times must be specified in the input zonal file.

E-scooter time and e-bike times are based on a distance calculated as shortest path across the all-streets network. Figure 9 and Figure 10 show frequency distributions of e-scooter and e-bike distances and speeds from the 2022 household travel survey respectively. Table 40 shows distance and speed statistics from the data; mean speed values are used to calculate utilities for e-scooters and e-bikes, and 90th percentile distances are used to set availability of these modes.

Ridehail modes

Ridehail modes include taxi, single-pay TNC (such as Uber or Lyft) and pooled TNC (Uber Pool, Lyft Line) Pooled TNC is also referred on ‘on-demand transit’ or ‘micro-transit’, and can include minivans and minibuses in addition to passenger cars. We recommend keeping these modes separate because the wait times and availability can vary considerably across the region, and due to fare differences between the modes. Auto skims are used for these modes, with an additional wait time and cost function. ActivitySim provides the ability to relate the wait time according to a distribution based on area type of the origin zone. The utility of pooled TNC includes an extra factor on both wait time and in-vehicle time, due to the need to divert vehicles to pick up the requested ride. The cost of each mode is specified in the constants.yaml file and includes an initial fee (meter-drop), a cost per mile, a cost per minute, and a minimum cost if relevant.

Pooled TNC is a unique mode whose availability has changed in recent years as Uber and Lyft have pulled these services back from most metropolitan areas due to the COVID pandemic. However, there continues to be interest in this mode, particularly in scenarios in which autonomous vehicles are expected to be available. It is possible that shared autonomous vehicles could reduce private auto ownership if the price reduction (due to elimination of labor costs) would make the mode more attractive. Microtransit is also of interest to many regions, as a substitute for more expensive bus service in corridors with low ridership.

Alternative-specific constants

The calculation of calibration targets for trip mode choice are based on the tour mode coding. Currently in ActivitySim, we limit PNR access tours to only one trip in each direction; this is done to ensure that the same parking location is accessible at both ends of the tour. Therefore after tour mode is coded, any non-PNR transit trips on PNR tours must be moved to the tour mode that is consistent with their mode. For example, if there are drive-alone trips on PNR tours, they would be moved to the drive-alone tour mode prior to model calibration. Similar steps are taken for cells with very low counts of trips within a tour mode.

FIGURE 9: E-SCOOTER DISTANCE AND SPEED FREQUENCY DISTRIBUTIONS

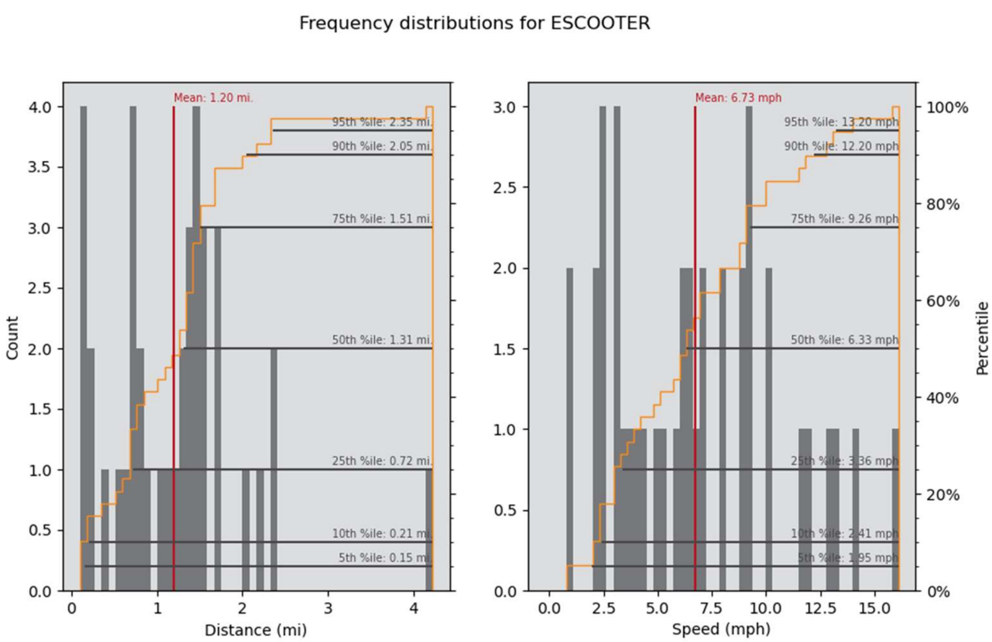


FIGURE 10: E-BIKE DISTANCE AND SPEED FREQUENCY DISTRIBUTIONS

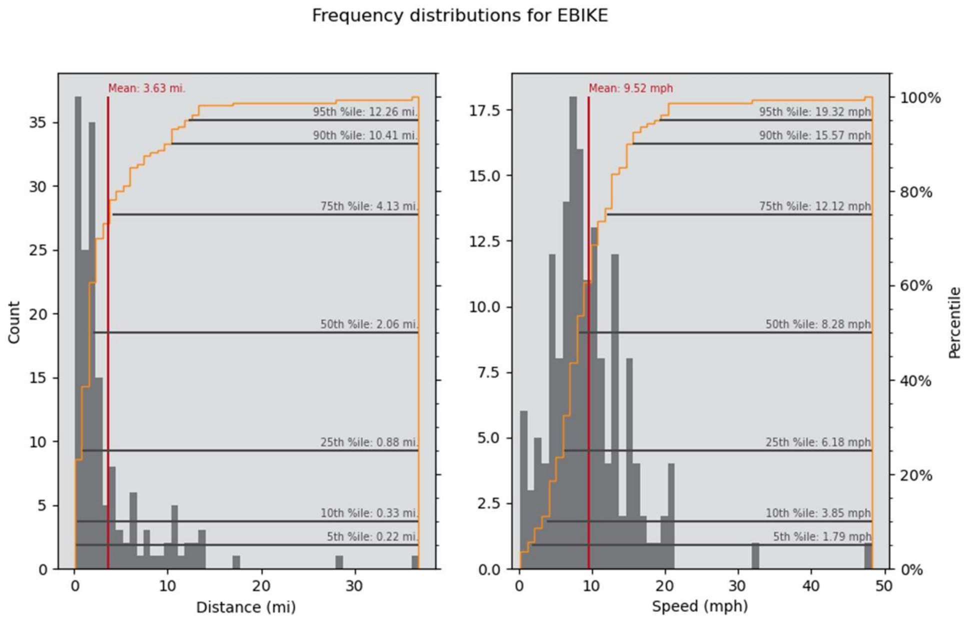


TABLE 40: WALK, BIKE, E-SCOOTER, AND E-BIKE DISTANCE AND SPEED STATISTICS

	MODE	WALK	BIKE	E-SCOOTER	E-BIKE
	Records	5928	487	39	161
Distance (mi)	Mean	0.70	3.61	1.20	3.61
	5th	0.11	0.27	0.15	0.22
	10th	0.14	0.36	0.21	0.33
	25th	0.22	0.70	0.72	0.88
	50th	0.41	1.46	1.31	2.06
	75th	0.81	3.84	1.51	4.12
	90th	1.44	9.75	2.05	10.36
	95th	2.16	14.75	2.35	12.21
Speed (mph)	Mean	3.02	7.78	6.73	10.07
	5th	0.79	1.20	1.95	1.79
	10th	1.22	2.13	2.41	3.87
	25th	1.88	4.33	3.36	6.18
	50th	2.43	7.56	6.33	8.29
	75th	2.95	10.66	9.26	12.21
	90th	4.52	12.99	12.20	15.85
	95th	7.47	15.07	13.20	19.53

7.0 REVISED PARKING CALCULATIONS

Parking-related inputs and calculations were simplified and automated in ABM3. In previous versions of the model, parking cost is specified for three time periods (hourly, daily, and monthly) by MGRA, along with the quantities of each type of parking differentiated by whether the parking in the MGRA is only allowed for trips to the MGRA or can be used for other MGRAs, the number of free hours of parking allowed before the cost takes effect, and a parking district/area code which indicates whether

1) trips with destinations in the MGRA may choose to park in a different MGRA but incur expected parking costs, or if

2) trips with destinations in parking area 1 may choose to park in the MGRA, parking charges might apply (incorporating a buffer around downtown), or if

3) only trips with destinations in the MGRA may park there and incur a parking charge, or if

4) only trips with destinations in the MGRA may park there, and no parking charges are applied.

In total ten different fields were used to track parking cost and supply at an MGRA level. SANDAG found it difficult to update and maintain this data and use the representation to test parking cost and supply changes for plan scenarios. Therefore a simpler method was developed.

The new parking cost and supply inputs to the SANDAG model include the hourly, daily, and monthly parking cost for each MGRA and the total number of parking spaces in each MGRA. These data were created by consolidating parking inventory data collected by SANDAG in 2021 (Figure 11). Parking costs (hourly_cost, daily_cost, monthly_cost) and parking spaces (paid_space) are the actual cost of parking in the MGRA and the actual number of spaces in the MGRA based on the inventory. Two other fields are included in the model inputs: free_spaces and parkarea. Free parking spaces describes the number of on-street free parking spaces in a buffer around parking-constrained areas. It provides an option for travelers with a destination in a parking constrained area to park for free outside the area and walk to their destination.

The parking inventory covered 1498 MGRAs. Table 41 tabulates MGRAs in the data by the comnbination of up to two types of parking cost data available in the inventory. As can be seen from the table, most MGRAs had only hourly costs. Figure 12 shows parking cost inventory data for downtown San Diego for all three time periods. Parking costs for undefined periods were calculated using an imputation method called Multiple Imputation by Chained Equations (MICE). The method works by cycling through each variable, using remaining data as predictors iteratively until convergence is reached. Missing costs were imputed for MGRAs where cost for at least one period (hourly, monthly, or daily) was available. The equations are shown in Figure 13.

FIGURE 11: PARKING INPUT DATA CONSOLIDATED FROM PARKING INVENTORY

Fields:
mgra
on_street_free_spaces
on_street_paid_spaces
on_street_hourly_cost_during_business
on_street_hourly_cost_after_business
off_street_free_spaces
off_street_paid_public_spaces
off_street_paid_private_spaces
off_street_paid_public_hourly_cost_during_business
off_street_paid_public_hourly_cost_after_business
off_street_paid_public_daily_cost
off_street_paid_public_monthly_cost
off_street_paid_private_hourly_cost_during_business
off_street_paid_private_hourly_cost_after_business
off_street_paid_private_daily_cost
off_street_paid_private_monthly_cost
off_street_residential_spaces

Values:
- Cost
- Spaces

Segments:
- *pricing:*
  - paid
  - free
- *rate:*
  - hourly
  - daily
  - monthly
- *location:*
  - on-street
  - off-street
- *ownership:*
  - public
  - private
- *tod:*
  - business hours
  - non-business hours

Consolidated data table:
- **mgra**
- **free_spaces** = total free spaces
- **paid_space** = total paid spaces
- **hourly_cost** = average hourly cost weighted by spaces
- **daily_cost** = average hourly cost weighted by spaces
- **monthly_cost** = average hourly cost weighted by spaces

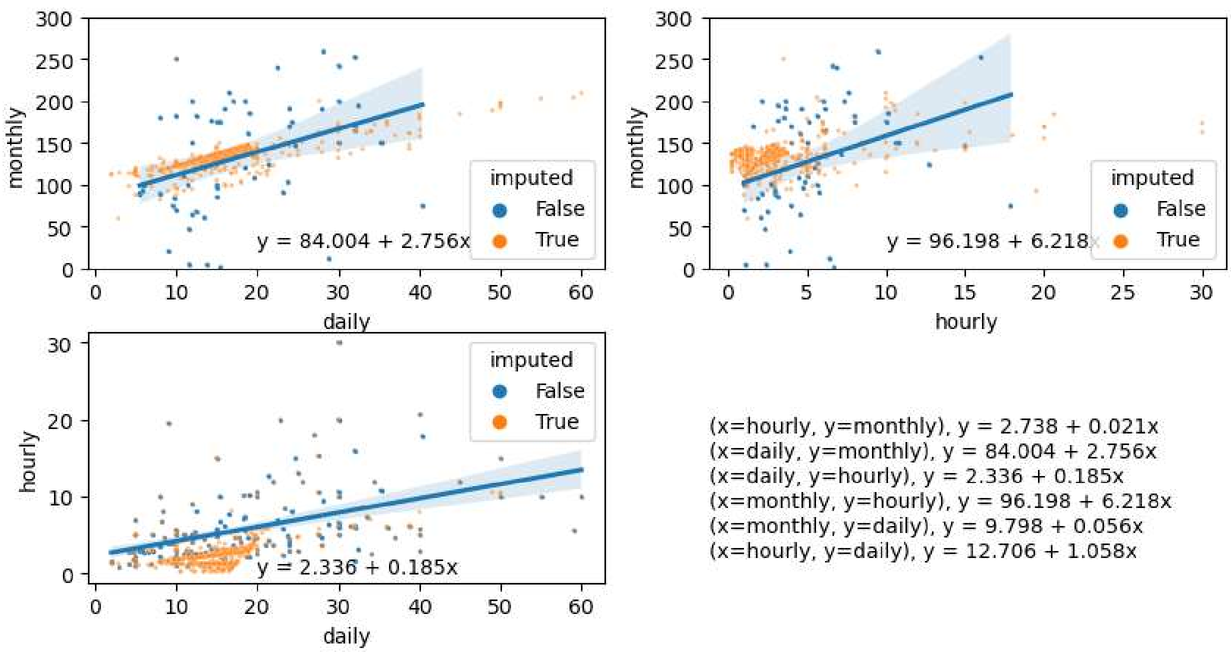
TABLE 41: NUMBER OF MGRAS BY PARKING COST DATA

	Hourly	Daily	Monthly	Total
Hourly	523	205	67	795
Daily	205	228	67	500
Monthly	67	67	69	203
Total	795	500	203	1498

FIGURE 12: PARKING COST INVENTORY DATA IN DOWNTOWN SAN DIEGO

Hourly Costs | Daily Costs | Monthly Costs

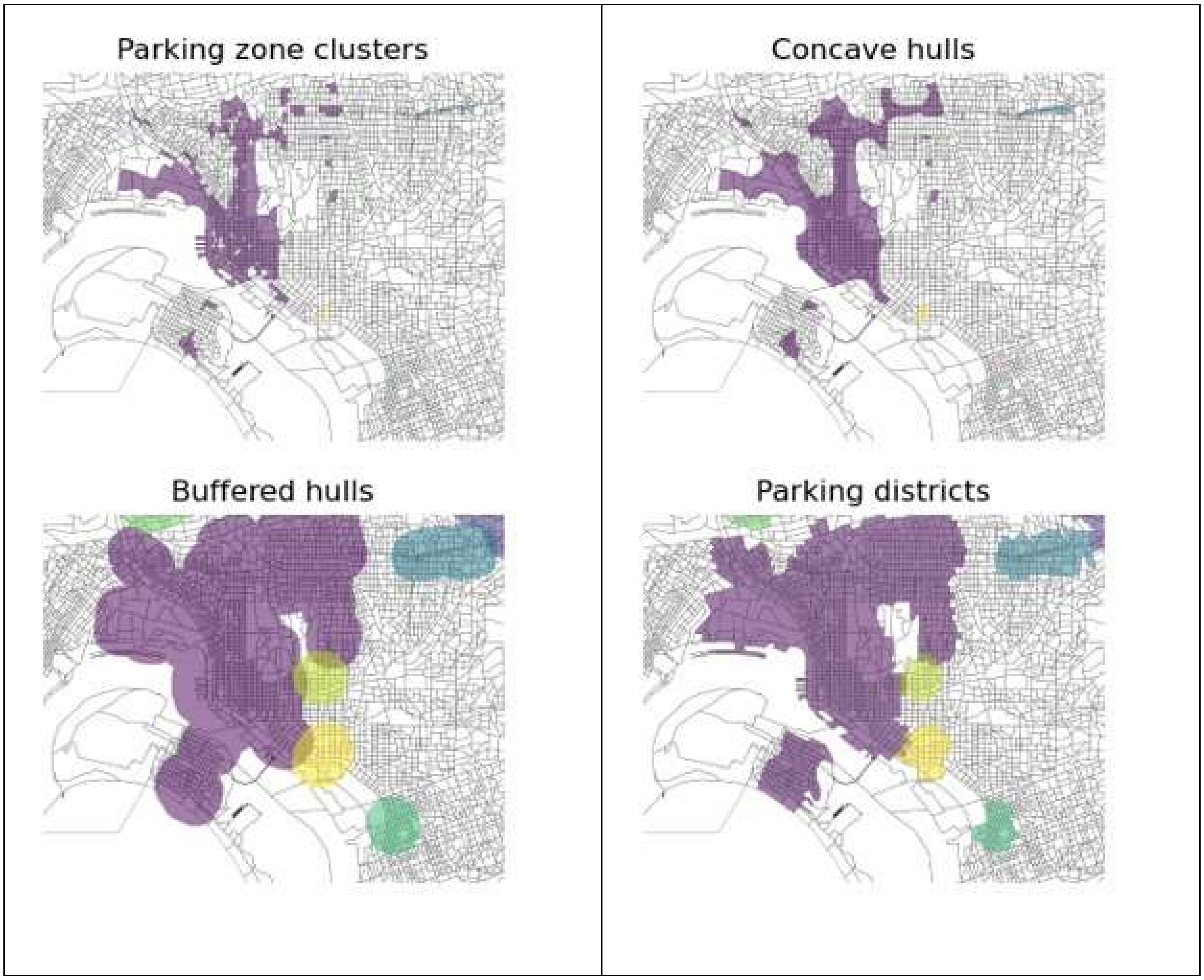
FIGURE 13: PARKING COST IMPUTATION EQUATIONS

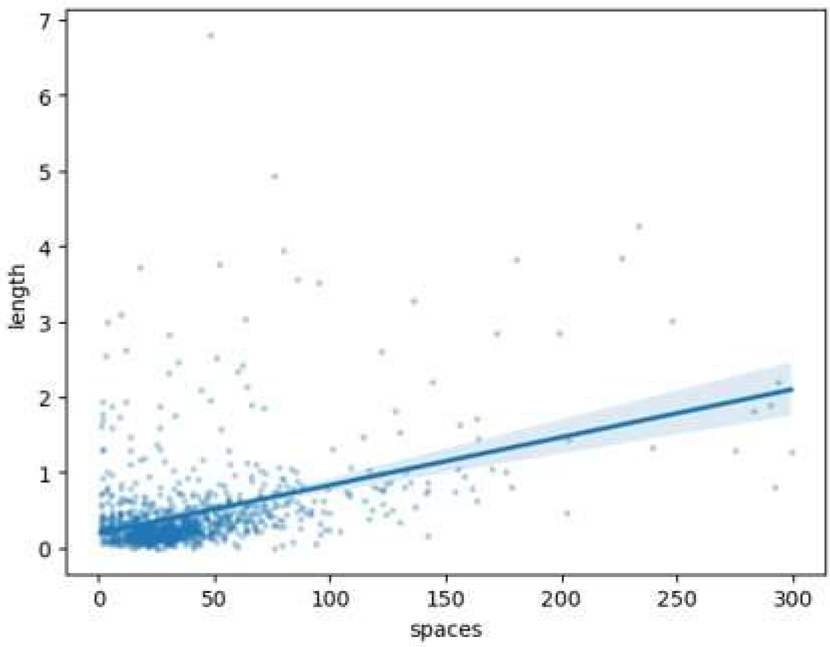


The parking area code was created using geospatial clustering algorithms. As noted above, parking area codes define which destination MGRAs travelers would be expected to pay a parking cost, even if there is no public parking in the specific destination MGRA. In other words, if a traveler goes to an MGRA in downtown San Diego, they may park in a nearby MGRA, pay for parking, and walk to their destination. The cost that they expect to pay is a weighted average of all nearby parking facilities, weighted by the number of spaces in the parking lot but discounted by the distance to their actual destination. Zones either ‘belong’ to this parking-constrained area, a buffer around the area, or are outside the area. Deterministic geospatial clustering and computational geometry algorithms were used to group MGRAs into parking “districts” to then compute the expected parking cost. The geospatial algorithm involves four steps: 1) clustering, 2) concave hull formation, 3) adding a buffer boundary, and 3) spatial joining zones into districts.

First, nearby MGRAs with parking costs are clustered with each other using an affinity matrix of computed straight-line distance between each zone’s centroid to deterministically cluster by proximity using a maximum walking distance of 0.5 miles. This results in a set clusters of contiguous zones where each MGRA with parking cost data is no further than 0.5 miles to its nearest neighbor. To then include nearby MGRAs without parking cost data an envelope, also called a concave hull, around each cluster is formed using Delaunay triangulation. A perimeter is formed around the parking cost MGRA zones where the maximum radii of any triangle formed between MGRA polygons cannot exceed the maximum radii of 0.5 miles. Using a maximum radius of 0.5 miles provides an appropriate “walking scale” envelope. A smaller radius would tighten the envelope around the parking cost MGRAs until it forms a minimum spanning tree, and a larger radius would expand the envelope until it forms a convex hull. Each concave hull is then increased by a buffer size of 0.5 miles to include external park and walk cases. The final expanded concave hulls now represent the parking “districts” from which expected parking costs can be calculated. Any MGRA that falls within the parking district is then spatially joined to the district and used to calculate the distance-based expected parking cost and any MGRA outside of the district is excluded. Figure 14 shows the parking areas created for the area around downtown San Diego. The MGRAs in the concave hull are those for which an expected cost is calculated (parking area 1). The MGRAs in the buffered hull but not belonging to the concave hull are those for which we calculate on-street parking to include in expected costs (parking area 2). And all other MGRAs have no expected parking cost and no on-street parking is included for them (parking area 3).

FIGURE 14: PARKING AREAS





Expected parking costs are the average parking costs that a traveler might expect to pay when they park either in the MGRA or somewhere close to the MGRA and walk to the MGRA from their parking location. Expected parking cost calculations take into account the potential for parking for free to the extent that travelers seek out free on-street parking and walk to their destination which has an expected cost. The parking cost equation is shown below, and graphically in Figure 15.

$$\text{expected cost} = \frac{\sum Cost\ * Spaces\ *\ e^{dist_{walk}\,*\,\beta_{walkdist}}}{\sum Spaces\ * e^{dist_{walk}\,*\,\beta_{walkdist}}}$$

The coefficient used on walking distance in the above equation is -0.03. Figure 16 shows the results of the hourly parking cost calculations for downtown San Diego.

FIGURE 15: PARKING COST GRAPHIC

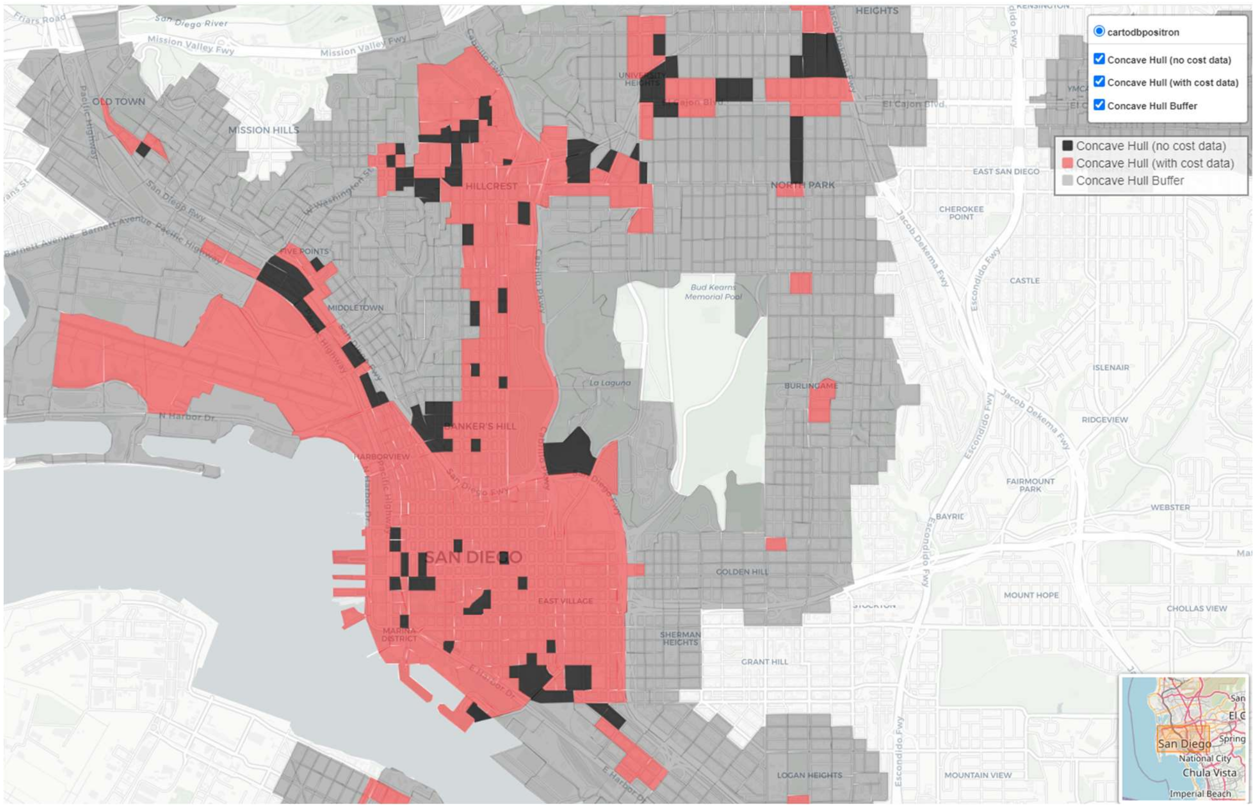
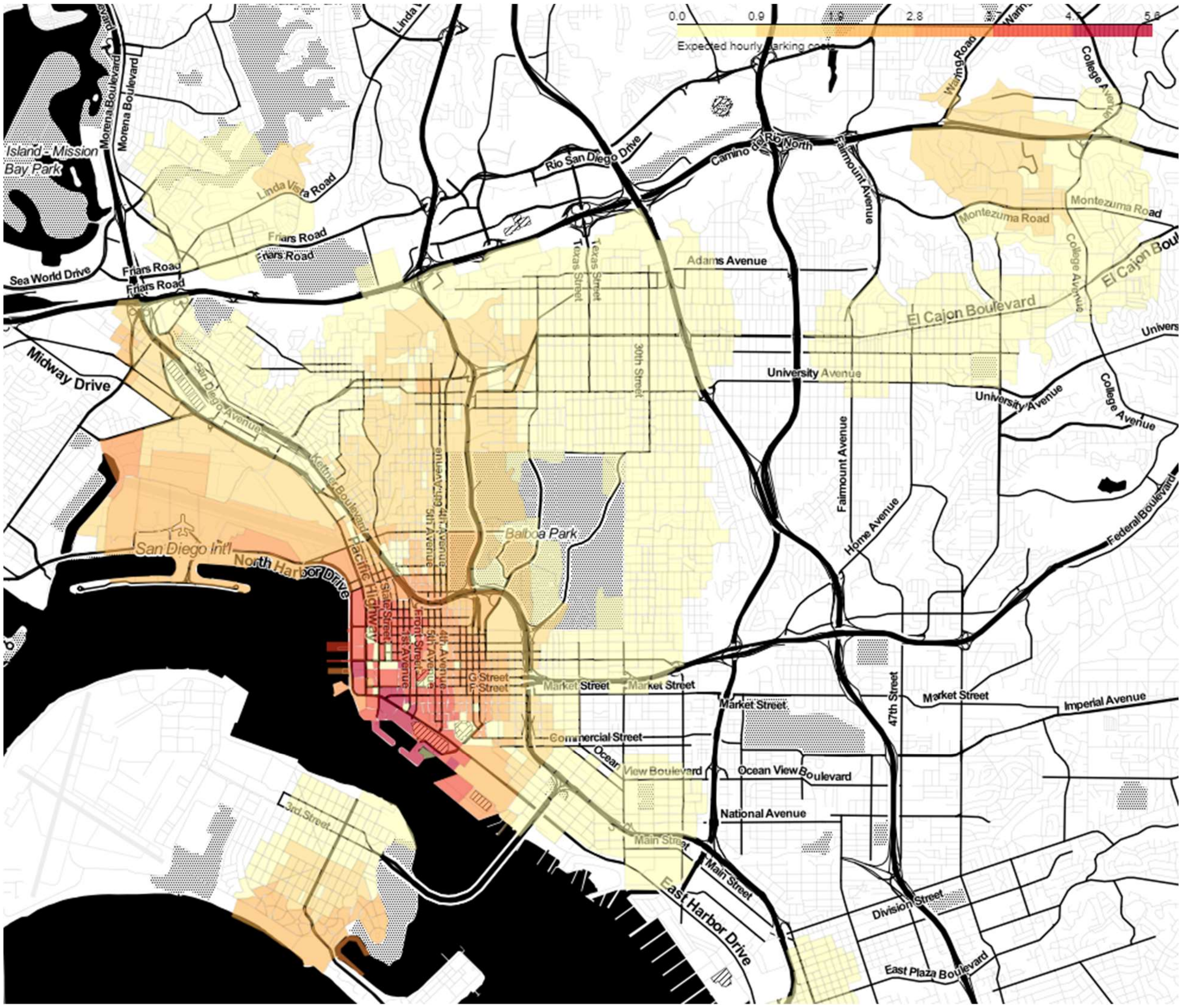


FIGURE 16: EXPECTED PARKING COST CALCULATION



8.0 SPECIAL MARKET MODEL UPDATES

All special market models have been converted to ActivitySim format. The cross-border model was converted under a separate project and documentation on that model was prepared separately. The special market models converted to ActivitySim under this project include the airport model, with separate implementations for San Diego International Airport and the Cross-Border Express airport, and the overnight visitor model. The airport model covers ground-access trips to and from the airport for the purpose of serving airport passengers, and the overnight visitor model generates tours for visitors staying overnight in hotels, motels, bed and breakfasts, short-term rental vacation homes, campgrounds, and staying with friends and family. This section describes those models.

8.1 AIRPORT MODELS

The airport sub-model of the SANDAG ABM3 development has been converted to the ActivitySim platform. This model simulates trips to and from the airport for residents, visitors, and external travelers. These trips are generated by arriving or departing passengers and are modeled as tours within the ActivitySim framework. A postprocessing script generates return trips for passengers who are dropped off. The existing CT-RAMP airport trip distributions will be used in their CT-RAMP input format, which allows for the use of future year distributions created by SANDAG.

In the ActivitySim framework, airport "tours" are equivalent to airport "trips" in the previous model. These tours are assigned an origin at the airport MGRA. During the stop frequency step of ActivitySim, a trip is assigned to the appropriate leg of the tour (either to or from the airport) while the opposite leg is not assigned any trips (referred to as the ‘dummy leg’). Passengers who are departing the county and traveling to the airport are considered "inbound," while arriving passengers are considered "outbound" as they are traveling from their tour origin to a non-airport destination. It is important to note that, to work within the ActivitySim framework, the airport trips must be modeled as tours, rather than being generated directly as in the previous model.

PreProcessing Step

A preprocessor is built in to the *airport_model.py* script which enumerates airport trips and generates trip scheduling distributions for each chooser segment. The *preprocessing.yml* file includes configurations and file locations needed for this step. The preprocessor uses the prepared CT-RAMP inputs for the airport model as shown in Table 42.

TABLE 42: CTRAMP AIRPORT INPUTS

FILE NAME	DESCRIPTION
Airport_purpose.{SAN,CBX}.csv	Probability distribution for trip purposes
Airport_party.{SAN,CBX}.csv	Probability distribution for party size
Airport_nights.{SAN,CBX}.csv	Probability distribution for number of nights the trip lasts
Airport_income.{SAN,CBX}.csv	Probability distribution for household income
Airport_arrival.{SAN,CBX}.csv	Probability distribution for arriving trip schedules
Airport_departure.{SAN,CBX}.csv	Probability distribution for departing trip schedules
Airport_employee_park.SAN.csv	Distribution of parking stalls at employee parking lots and share of employees that travel to the terminal and share of employees that take transit.

Tour enumeration

The tour enumeration replicates the process from the CT-RAMP model with the following formula:

$$airport\ trips = \frac{(Enplanements - \ Connections)}{annualizationFactor * AveragePartySize} * 2$$

Half the tours are assigned to be arriving and the other half departing. Monte Carlo draws from the probability distributions in Table 42 assign each tour with a purpose, party size, number of nights, and income.

Employee trips from off-site employee parking is generated by information in the employee parking file. Each offsite parking lot lists the number of parking stalls and the share which then transfer to the airport. The number of employee tours is simply the number of parking stalls multiplied by the share going to the airport. Additionally, there is a share which is expected to take public transit which is also tagged in the tour enumeration step of the pre-processor. Employee tours are expected to return so for each tour to the airport a corresponding tour is created from the airport to the off-site parking.

Scheduling probabilities

The airport model uses the scheduling distributions when enumerating airport tours to assign tour attributes. The scheduling probabilities are converted to ActivitySim format in the pre-processor. ActivitySim expects a dual tour departure and arrival probability (for the airport trip leg and the dummy leg). Since only one trip occurs on an airport ‘tour’ in this application, departing trips (inbound trips) are assigned a tour start time of 1 and the probabilities are mapped to the end period in the distribution. Similarly, arriving passengers (outbound passengers) are assigned a tour start of the probability distribution with a tour end period of 48.

ActivitySim Prerequisites

The pre-processor also generates the necessary pre-requisites for ActivitySim which include a household, person, and land use file. The household and person file are generated by assuming there is one household per tour, and one person per household. Unique person and household ids are assigned to match the unique tour ids.

The land use file is almost an exact replica of the existing land use file except the household income bins are re-assigned based on the synthetic population to match the size terms used in the airport model by income bin.

ActivitySim Sub-Models

This section describes each ActivitySim sub-model that is run to implement the airport model.

Tour Scheduling Probabilistic

The tour scheduling model uses a probabilistic draw of the scheduling distribution prepared by the pre-processor. This model assigns start and end times to the tour. This is important because it will also serve as the schedule model for the final airport trips. In ActivitySim the trips are scheduled based on the tour schedule. If there is only one trip per leg on the tour (such as our case here) the trip is assigned the tour start/end time.

Non-Mandatory Tour Destination Choice

The destination choice model chooses the non-airport end of the airport trips. Each tour is set with an origin at the airport MGRA. The tour destination model of ActivitySim is used to choose the non-airport end of the trip. The utility equation includes the travel distance, and the destination size terms. ActivitySim destination choice framework requires a mode choice log sum. A fake tour mode choice log sum was created which generates a value of zero for every destination using the 'tour_mode_choice.csv' and 'tour_mode_choice.yml' file. This is a work around to prevent ActivitySim from crashing and not having to include the tour mode choice log sum in the destination choice model.

The important difference to note in the implementation of this model is that the ActivitySim version first samples alternatives and then chooses a destination while the original model chooses the destination directly. In this case, the sampling UEC is the same as the destination choice UEC except for a sampling of alternatives correction factor applied in the destination choice UEC.

Stop Frequency

The stop frequency model is where the trip table is first created in the model flow. The pre-processor tags each tour with a direction of 'inbound' or 'outbound' according to whether the tour is a departing or arriving passenger. The stop frequency model will generate the tour legs and the intermediate stops using the model specification. For this implementation the model is specified so that inbound tours are tagged with zero outbound trips and -1 inbound trips (and the opposite is true for outbound tours: -1 outbound trips and 0 inbound trips). The 0 signifies that no intermediate stops are made; this leg of the tour will only have one trip. The -1 signifies that no trip is made at all on that leg. Using the -1 allows us to create 'half-tours' where only one leg of the tour is recorded as a trip.

Trip Scheduling

The trip scheduling model assigns depart times for each trip on a tour. ActivitySim requires trip scheduling probabilities however, these are not used in this implementation since there is only one trip on any given tour leg. This means the trips will be assigned the tour scheduling times which were determined in the tour scheduling model. The trip scheduling probabilities file is just a dummy file.

Trip Mode Choice

The trip mode choice model determines the airport arrival mode. The arrival modes are shown in Table 43.

TABLE 43: TRIP ARRIVAL MODES

ARRIVAL MODE	DESCRIPTION
Park Location 1	Party drives personal vehicle to an on or off-site parking location.
Park Location 2	
Park Location 3	
Park Location 4	
Park Location 5	
Curb Location 1	Party is dropped off or picked up by another driver at designated curbside location.
Curb Location 2	
Curb Location 3	
Curb Location 4	
Curb Location 5	
Park and Escort	Party is driven in personal vehicle, parks on-site at the airport and is escorted to/from airport.
Rental Car	Party arrives/departs by rental car.
Shuttle Van	Party takes shuttle van.
Hotel Courtesy	Party takes hotel courtesy transportation.
Ridehail Location 1	Party takes ridehail to ridehail pick-up/drop-off location
Ridehail Location 2	
Taxi Location 1	Party takes taxi to taxi pick-up/drop-off location
Taxi Location 2	
Walk Local	Party walks to transit access.

ARRIVAL MODE	DESCRIPTION
Walk Premium	
Walk Mix	
KNR Local	
KNR Premium	Party kiss and rides to transit access.
KNR Mix	
TNC Local	
TNC Premium	Party takes TNC to transit access.
TNC Mix	
Walk	Walk mode is available only to employees who walk from off-site parking to the airport terminal.

The trip mode choice yaml file contains detailed variables associated with each trip mode. For example, each parking location is given an MGRA location, a walk-time, a wait-time, and a cost. If a parking location MGRA is set to -999 it is assumed to be unavailable and will not be in the choice set. The pre-processor in this step stores all values of skims from the trip origin to each of the access modes destinations along with any associated costs. Costs include parking fees per day, access fees, fares, and rental car charges.

Employees are not fed into the trip mode choice model. Instead, if a transit share is specified in the employee park file, that percentage of employees will be assigned 'Walk Premium' mode in the pre-processor. Otherwise, employees are all assigned 'Walk' mode from the employee parking lot to the terminal.

Airport Returns

Airport trips where the party is dropped of curbside or parked and escorted are assumed to also have the driver make a return trip to the non-airport location. This procedure is done as a post-processing step after mode choice and before trip tables are written out. An 'airport_returns.yml' file takes a user setting to determine which trip modes will include a return trip. These trips records are flagged and duplicated. The duplicated trips swap the origin and destination of the original trip and are assigned a unique trip id. These trips are tagged with ‘trip_num =2’ so they are easily sorted in any additional processing (such as for writing trip matrices).

Write Trip Matrices

The write trip matrices step converts the trip lists into vehicle trip matrices. The matrices are segmented by trip mode and value of time bins. The vehicle trip modes in the matrices include SOV, HOV2, and HOV3+. Table 44 shows each trip mode category classification. Value of time segmentation is either low, medium, or high bins based on the thresholds set in the model settings. No occupancy factor is used when creating these trip tables as each tour is associated with one household and one persons. The party size is only determined by the attributed value from the pre-processor. Therefore each record still represents one vehicle trip and the occupancy is determined by the party size and the mode.

TABLE 44: TRIP MATRICES CATEGORIES

TRIP MODE	CLASSIFICATION
Drive Alone	First leg trips with a party size of one who’s trip mode is either rental car or parking. Return trips where the trip mode is park and escort or curbside drop-off.
Shared 2	Trips of party size 2 if the trip mode is parking or rental car. First leg trips of party size 1 if the trip mode is park and escort, ridehail, taxi or curb drop-off.
Shared 3+	Trips of party size 3 if the trip mode is park or rental car. First leg trips of party size 2 or more if the trip mode is ride hail, taxi, curb drop-off, or park and escort.
Walk	Trip mode is walk. (Currently only available to employees).
Walk Transit	Trip mode is walk to transit (local, premium, or mixed service).
is_drive_transit	Trip mode is KNR or TNC access to transit (local, premium, or mixed service).

Base Year Model Results

The results of the ActivitySim model generally show a good match to the results given by the CT-Ramp model.

Destination Choice

The structure of the destination choice model varies a little compared to the CT-Ramp model in that ActivitySim samples alternatives and then chooses a non-airport destination as opposed to choosing from the full choice set of destinations. Even given this difference, since the sampling UEC and the destination choice UEC are nearly identical, the final model results match with the CT-Ramp results very well. Figure 17 and Figure 18 show the destination choice of the non-airport end of the trips by destination pseudomsa.

FIGURE 17: SAN DIEGO AIRPORT MODEL DESTINATION CHOICE

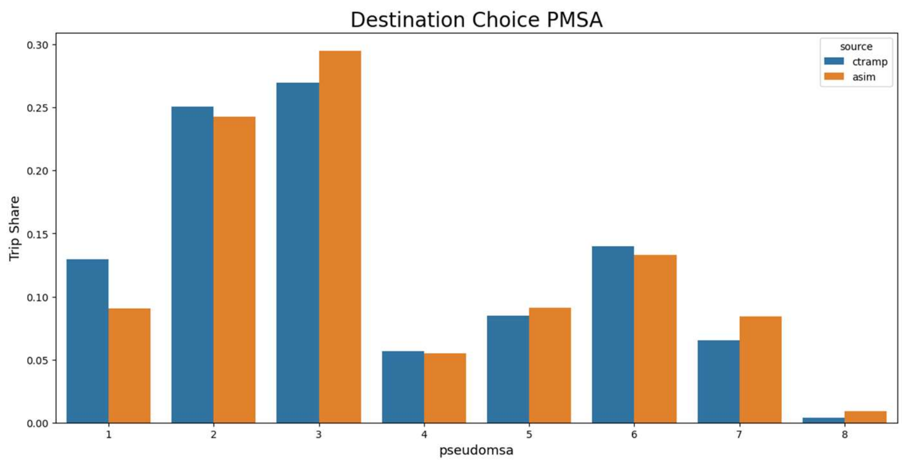
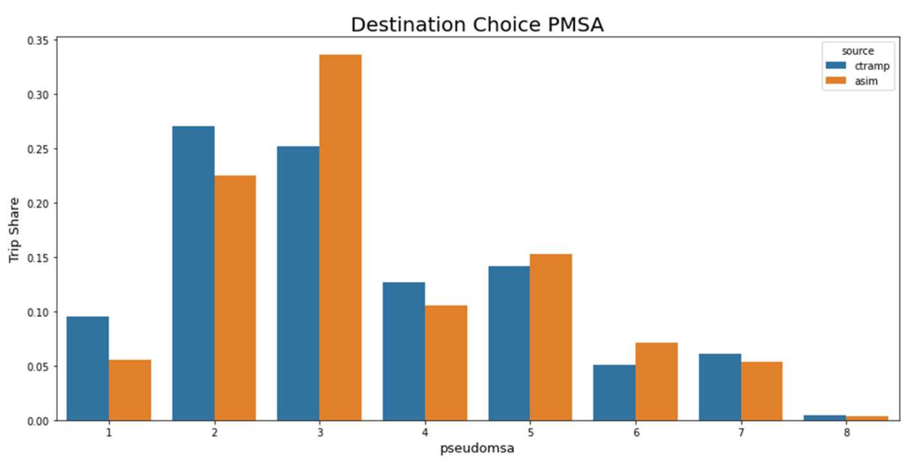


FIGURE 18: CROSS-BORDER EXPRESS AIRPORT MODEL DESTINATION CHOICE



Trip Mode (Arrival Mode) Choice

The structure of the arrival mode choice model also closely follows the structure of the CT-Ramp model.

FIGURE 19: SAN DIEGO AIRPORT MODEL ARRIVAL MODE CHOICE

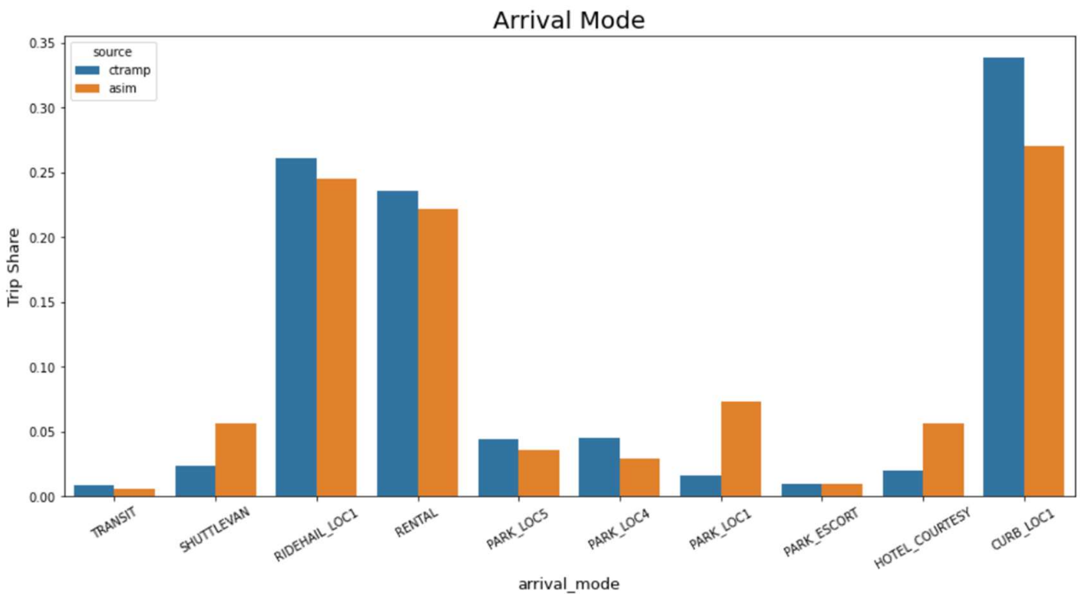
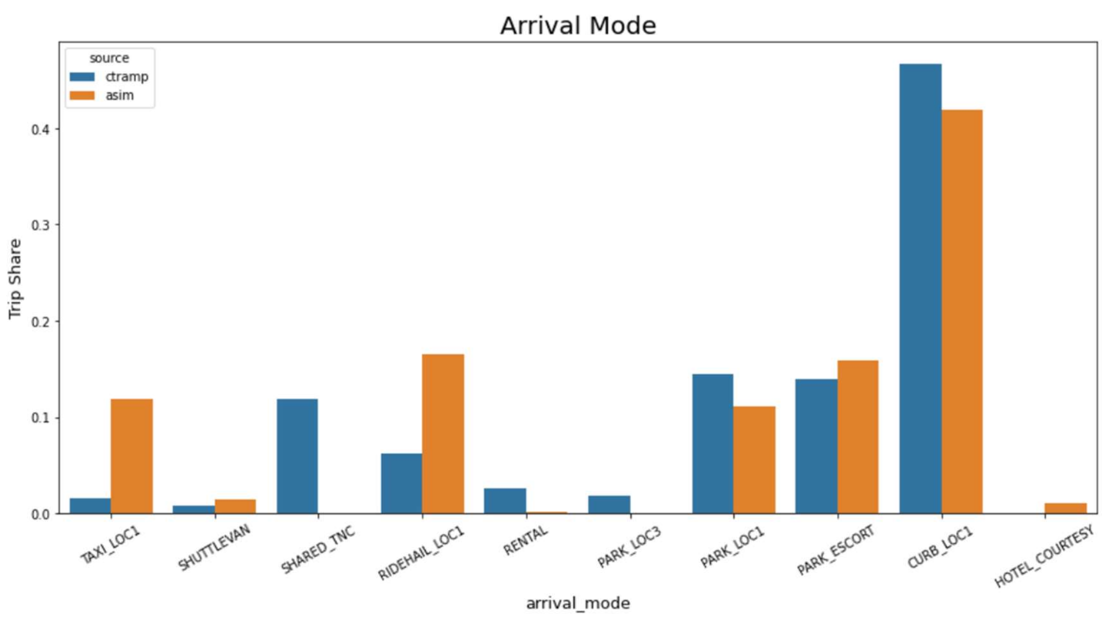


FIGURE 20: CROSS-BORDER EXPRESS AIRPORT MODEL ARRIVAL MODE CHOICE



8.2 VISITOR MODEL

The visitor sub-model of the SANDAG ABM3 development has been converted to the ActivitySim platform. This model simulates the demand of visitor travel in San Diego with special consideration for major attractions with treatment of visitors by visitor type and party size. The visitor sub-model makes use of the full set of modes within San Diego County, including auto trips by occupancy, transit trips, and non-motorized trips.

The visitor model is organized into three python files: *visitor_model.py*, *visitor_convert_configs.py*, and *visitor_tour_enum.py*. The *visitor_model.py* file is the main code and point of entry to the visitor model run. The *visitor_convert_configs.py* is an optional one-time preprocessing step to convert the CT-RAMP formatted inputs into ActivitySim format. If the inputs are already in ActivitySim format, there is no need to run this helper step. The *visitor_tour_enum.py* generates tours from probabilistic distributions which are then used by the native ActivitySim model.

PreProcessing

There is a one-time preprocessing step which performs two actions. First it reads the CT-RAMP formatted inputs and converts them to ActivitySim, and second it copies the necessary ActivitySim configurations from the common ABM3 configurations and stores a subset specifically for the visitor model. The preprocessing step in *visitor_convert_configs.py* is imported into and run from *visitor_model.py* code. The preprocessor uses the prepared CT-RAMP inputs for the visitor model as shown in Table 45.

TABLE 45: CT-RAMP VISITOR MODEL INPUTS

FILE NAME	DESCRIPTION
visitor_autoAvailability.csv	Automobile availability probability distribution by tour purpose.
visitor_income.csv	Income probability distribution by visitor type.
visitor_partySize.csv	Party size probability distribution by tour type.
visitor_stopFrequency.csv	Stop frequency and duration probability distribution by inbound and outbound tours.
visitor_stopPurpose.csv	Stop purpose probability by tour purpose, stop number, and inbound and outbound tours
visitor_tourTOD.csv	Tour time of day probability distribution by depart and return time of day.
visitor_businessFrequency.csv	Tour purpose probability distribution by number of work visitor tours.
visitor_personalFrequency.csv	Tour purpose probability distribution by number of personal visitor tours.
visitor_outboundStopDuration.csv	Stop duration probability distribution by number of remaining stops on outbound tour.
visitor_inboundStopDuration.csv	Stop duration probability distribution by number of remaining stops on inbound tour.
tod_conversion.csv	Crosswalk conversion table to convert CT-RAMP's 40 time periods to ActivitySims 48 time-period system.

Tour enumeration

The support module *visitor_tour_enum.py* is called from the main *visitor_model.py* module, which runs tour enumeration to directly synthesize visitor tours for subsequent modeling. Visitor travel parties are created by visitor segment based upon input data for hotels, households, and visitor travel type. The travel parties are given household income attributes and tours by purpose are generated for each party. Each tour is given attributes of auto availability and party size. Tours are generated from each MGRA, which becomes the tour origin, and its destination is modeled in ActivitySim.

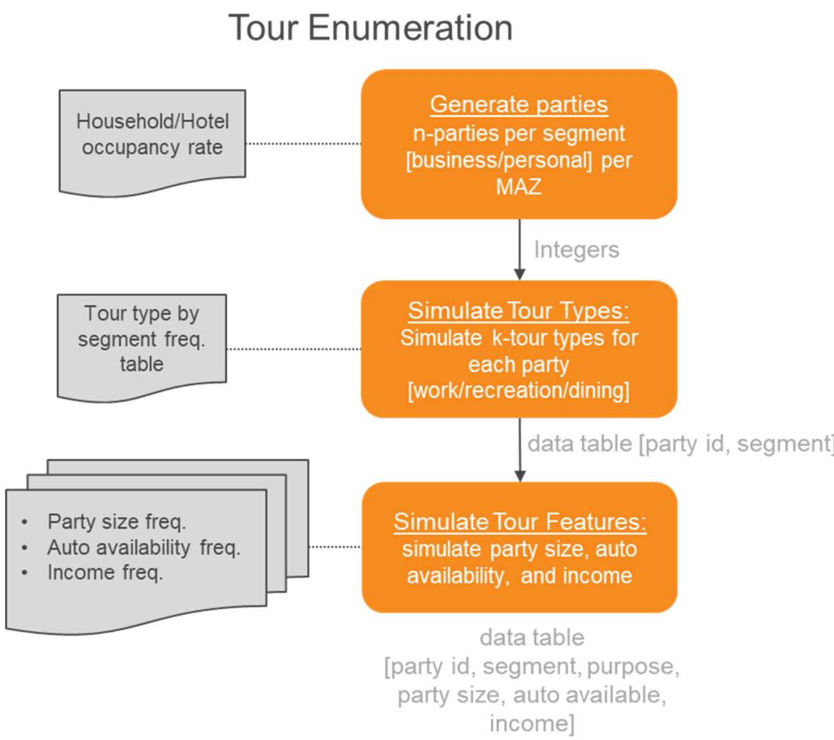


FIGURE 21: VISITOR TOUR ENUMERATION SCHEMATIC

Travel Party Generation

First, visitors are generated for two visitor segment types with distributions coming from Visitor Bureau data and a previously collected visitor survey:

- **Business**: Self-identified as business traveler, or self-identified as 'Both Business and Personal' but took at least one 'business' purpose trip on travel day
- **Personal**: Self-identified as personal traveler, or self-identified as 'Both Business and Personal' but took no business purpose trips on travel day. A few self-identified Personal travelers have reported Work tours.

The model generates visitor parties by segment by applying separate occupancy rates to hotels and households, which were obtained from the San Diego Convention and Visitor Bureau. The model then applies separate distributions of visitor parties by segment to hotel visitor parties and household visitor parties separately. Visitor parties are then attributed with household income based on the distribution of parties by visitor segment and income. The party size and auto-availability are attributed on a tour-by-tour basis since these attributes can change depending on tour and day.

Tour Generation

Tours are then probabilistically generated for visitor parties with probabilistically assigned party size, auto-availability, and income. There are three tour purposes:

- **Work**: Business travel made by business visitors.
- **Dining**: Travel to food establishments for both business and personal visitors.
- **Recreational**: All other non-work non-food related activities.

ActivitySim Prerequisites

After the CT-RAMP inputs are converted to ActivitySim format and Tour Enumeration is completed where visitor parties and their tours are generated, their activity is then modeled natively in ActivitySim. The pre-processor also generates the necessary pre-requisites for ActivitySim which include a household, person, and land use file. The household and person file are generated by assuming there is one household per tour and the party size is analogous to household size. Unique person and household ids are assigned to match the unique tour ids.

The land use file is almost an exact replica of the existing land use file except the household income bins are re-assigned based on the synthetic population to match the size terms used in the visitor model.

ActivitySim Sub-Models

This section describes each ActivitySim sub-model that is run to implement the visitor model. The probabilistic approach is like the airport and cross border models, but simpler regarding its handling of inbound and outbound tours. The model consists of:

Probabilistic Tour Scheduling Model

The tour scheduling model uses a probabilistic draw of the scheduling distribution prepared by the pre-processor. This model assigns start and end times to the tour. In ActivitySim the trips are scheduled based on the tour schedule.

Non-Mandatory Tour Destination Choice

The destination choice model chooses the non-airport end of the airport trips. Each tour is set with an origin party origin MGRA. The tour destination model of ActivitySim is used to choose the tour destination of the outbound tour. The utility equation includes the travel distance, and the destination size terms. ActivitySim destination choice framework requires a mode choice log sum.

Stop Frequency

The stop frequency model is where the trip table is first created in the model flow. The pre-processor tags each tour with a direction of 'inbound' or 'outbound' according to whether the tour is departing or arriving. The stop frequency model will generate the tour legs and the intermediate stops using the model specification.

Trip Scheduling

The trip scheduling model assigns depart times for each trip on a tour probabilistically using the predefined departure and duration probability distribution tables.

Trip Mode Choice

The trip mode choice model determines the mode. This model utilizes the native ActivitySim mode choice model. The trip mode choice yaml file contains detailed variables associated with each trip mode. For example, each parking location is given an MGRA location, a walk-time, a wait-time, and a cost. If a parking location MGRA is set to -999 it is assumed to be unavailable and will not be in the choice set. The pre-processor in this step stores all values of skims from the trip origin to each of the access modes destinations along with any associated costs. Costs include parking fees per day, access fees, fares, and rental car charges.

Base Year Model Results

Tour enumeration results

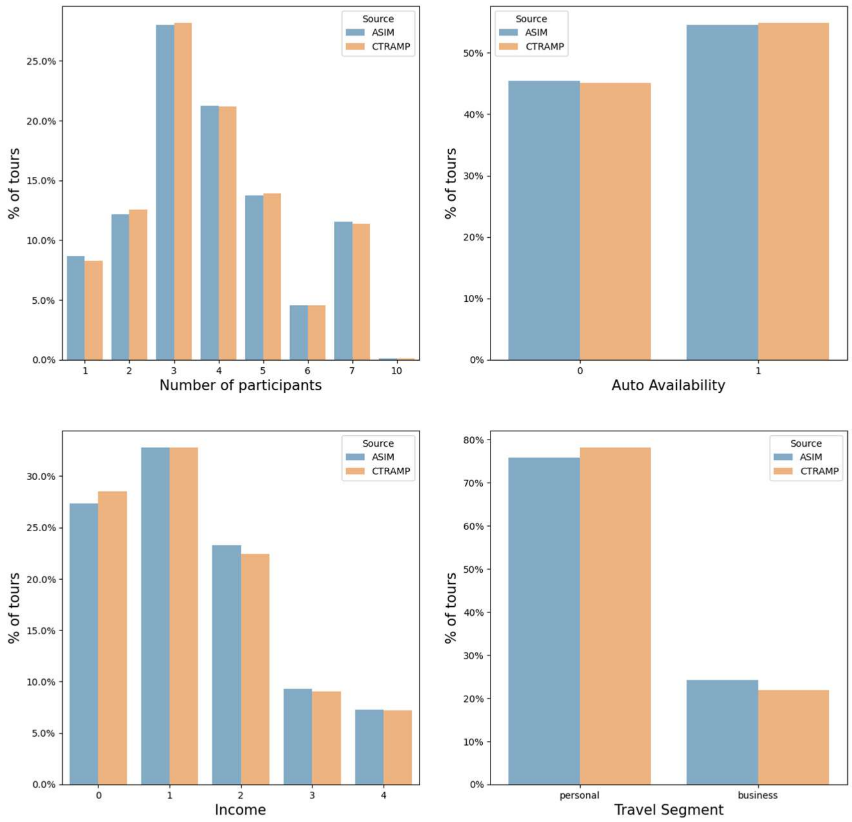


FIGURE 22: VISITOR TOUR ENUMERATION RESULTS

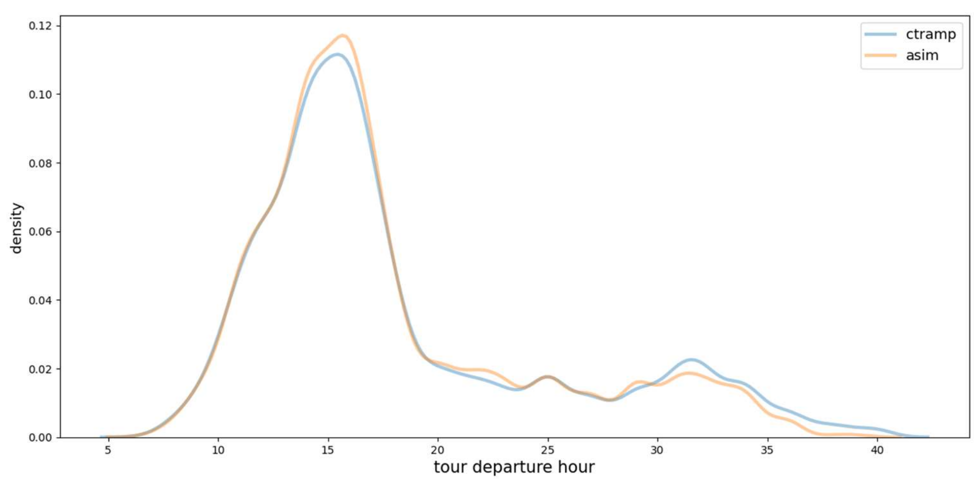


FIGURE 23: VISITOR TOUR SCHEDULING RESULTS

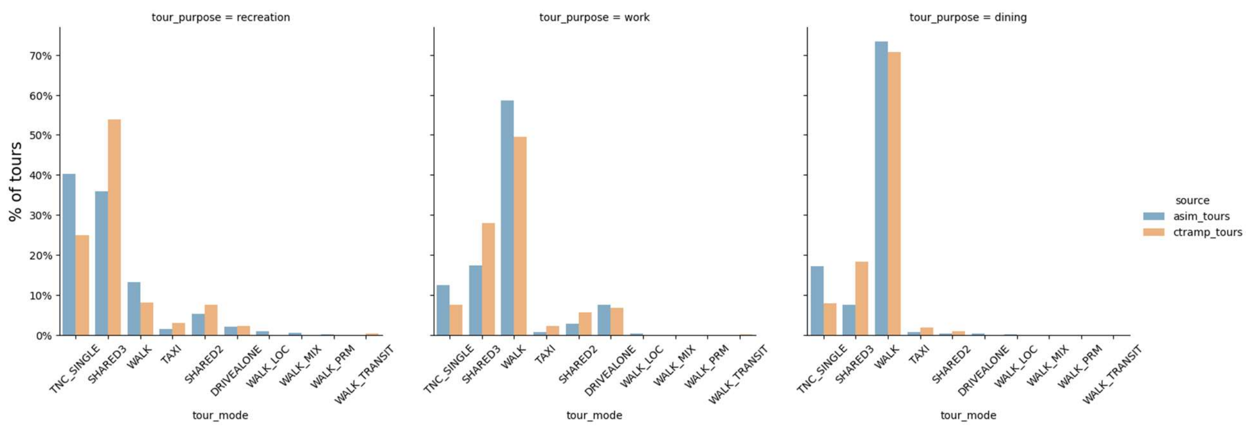


FIGURE 24: VISITOR TOUR MODE CHOICE RESULTS

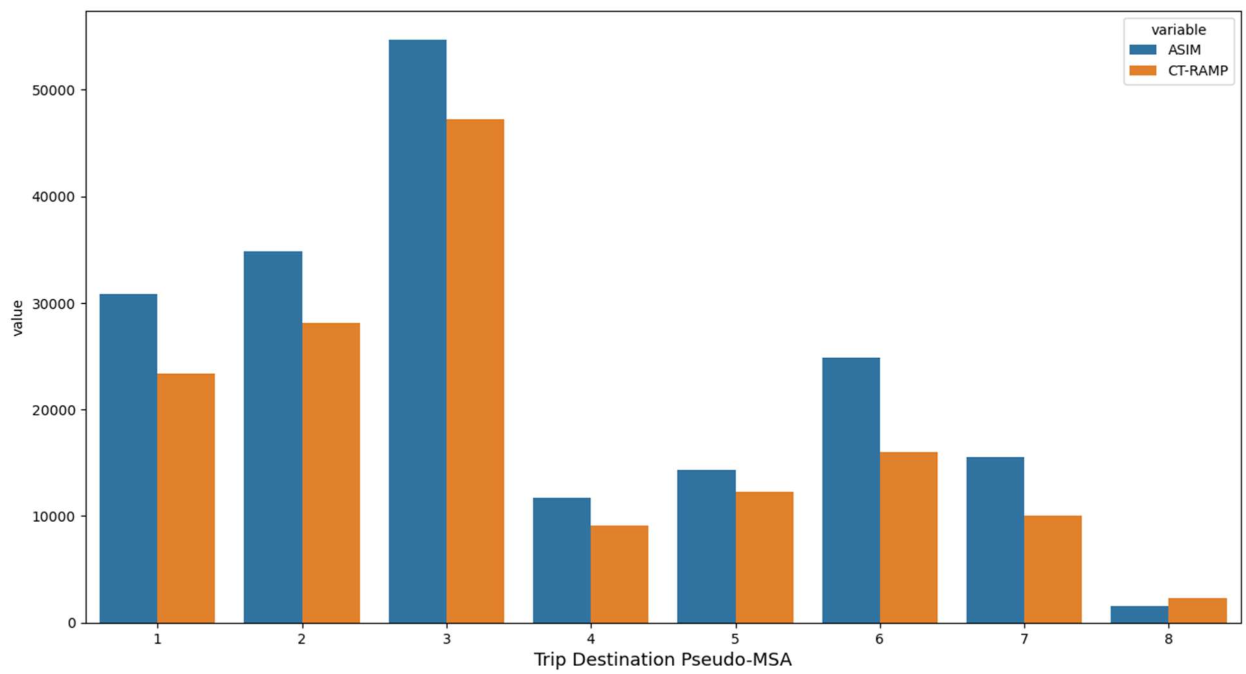


FIGURE 25: VISITOR TRIP DESTINATION RESULTS

9.0 SCENARIO MANAGER

ActivitySim models generally rely on a slew of model parameters in order to forecast the demand in the current or future year scenario. A number of these parameters including auto operating cost, taxi and TNC fare, micromobility cost, and AV ownership penetration are usually assumed to change by forecast year or scenario. Manually changing these parameters requires the model user to know where each parameter is located, and individually changing them according to the scenario forecast values. A scenario manager, therefore, can be a convenient and efficient tool to automate this process. We implement a scenario management solution in the SANDAG activity-based model, as described below.

The ActivitySim Scenario Manager is a python script that reads in a CSV input file containing the parameter values for each scenario, and updates the associated parameters in the ActivitySim config files. An example of the input parameter CSV file is shown in Table 46, where each row is associated with a specific scenario year/name. The parameter names used here can either be identical to the parameter names used in ActivitySim, or different. In case the parameter names are different, a separate file is used to map the parameters names between the input CSV and ActivitySim config files.

TABLE 46 AN EXAMPLE OF A SCENARIO MANAGER INPUT FILE

Year	AOC fuel	AOC maintenance	Taxi baseFare	Taxi costPerMile	Taxi costPerMinute
2012	13.5	6.3	1.78	1.87	0.08
2014	12.9	6.3	1.78	1.87	0.08
2015	19.5	6.2	1.78	1.87	0.08
2016	10.7	5.6	1.78	1.87	0.08
2017	10.8	5.5	1.78	1.87	0.08

The bulk of parameters in ActivitySim are stored in a configuration file named constants.yaml, although depending on the ActivitySim setup, some models' specific parameters may also be set in that model step's YAML file directly. An example of the setup of a constants.yaml is shown in Figure 26.

FIGURE 26: AN EXAMPLE OF AN ACTIVITYSIM YAML FILE CONTAINING MODEL PARAMETERS


```yaml
## ActivitySim
## See full license in LICENSE.txt.

scenarioYear: 2022
NO_EXTERNAL: 0
PRE_COVID: 0
max_local_walk_dist: 0.85
max_prm_walk_dist: 1.2
max_mix_walk_dist: 1.2
walk_speed: 3
max_local_walk_time: 17
max_prm_walk_time: 24
max_mix_walk_time: 24
HHT_NONFAMILY: [4, 5, 6, 7]
HHT_FAMILY: [1, 2, 3]
PSTUDENT_GRADE_OR_HIGH: 1
```


In running the scenario manager, the script will receive the scenario year and model run directory from the user, read in the scenario parameters from the input CSV file, and updates the ActivitySim YAML files accordingly. This script is run at the beginning of each model run to ensure the correct parameter values throughout the full run.