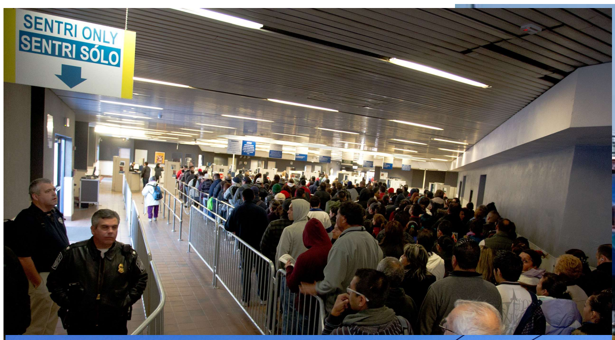
CROSS BORDER MODEL FINAL REPORT



April 2022

SANDAG

RSG
the science of insight

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

The San Diego Association of Governments (SANDAG) serves as the forum for regional decision-making for the San Diego region. SANDAG is required by state and federal law to maintain a transportation modeling system that addresses regional planning needs and can be used for air quality conformity determinations. Regional planning activities, such as the 2025 Regional Plan, and corridor and project level studies, rely on comprehensive, detailed, and validated transportation models produced from this work element. The regional transportation model is a key tool used to help decision-makers prioritize funding. Models provide planners and decision makers with information to help them equitably allocate scarce resources. SANDAG maintains and enhances the regional travel demand model used for this purpose.

The regional travel model is a disaggregate activity-based model implemented in the Java programming language and Python scripts that use the EMME Application Programming Interface (API). The model system includes a resident travel model as well as a number of special market travel models, including the 'cross-border travel model', which represents residents of Mexico (including U.S. citizens) traveling into and within San Diego County. In 2020 RSG was contracted to convert the cross-border travel model from Java into the ActivitySim[1] framework and update the model based on a recently (2019) collected travel survey.

The mission of the ActivitySim project is to create and maintain advanced, open-source, activity-based travel behavior modeling software based on best software development practices for distribution at no charge to the public. The ActivitySim project is led by a consortium of Metropolitan Planning Organizations (MPOs), Departments of Transportation (DOTs), and other transportation planning agencies, which provides technical direction and resources to support project development.

The ActivitySim resident travel model is based on the same basic structure as the current SANDAG resident model. However, there was no equivalent cross-border travel model implemented in ActivitySim. This project represents the first such special market model to be implemented in the ActivitySim framework. As part of the update, the model was enhanced to take advantage of more up-to-date data including the 2019 cross border survey, collected by True North Research, Inc.[2] on behalf of SANDAG.

Cross-border travel not captured by the Cross-Border Model includes:

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

[2] True North Research, Inc., *Cross-Border Travel Behavior Survey Summary Report*, (Encinitas: True North Research, Inc., 2020)

- Residents of San Diego County who travel to/from Mexico. This travel will be represented by the resident travel model.
- Residents of Mexico (including U.S. citizens) who travel into San Diego County and who do not return to Mexico at the end of the day. This travel is represented in the overnight visitor travel model.
- Travel between points of entry through San Diego County to other U.S. destinations. This travel is represented by the external-external travel model.
- Commercial vehicle travel to/from points of entry. This travel is represented by the commercial vehicle model.

The goals of this project have been achieved. They are listed below.

- The model has been implemented in the ActivitySim framework. The open-source software will reduce software maintenance costs and allow SANDAG to leverage feature and software performance investments made by the consortium as they become available. Similarly, other ActivitySim users will be able to implement the Cross-Border Model or selected features of the model for their region. The new Cross-Border model runs in 30-45 minutes depending upon the number of wait time iterations; the old Java model runs over 100 minutes; a runtime savings of 50% or more.
- The Cross-Border Model includes a dynamic wait time model component that estimates a wait time for general purpose lanes, special SENTRI and Ready lanes, and pedestrian crossings based on total lanes at each point of entry and total demand predicted by the model. This component improves model realism and can be used to inform border infrastructure investment decisions.
- The model has been updated to represent Cross-Border travel patterns revealed by recently collected survey data and data from Customs and Border Patrol and has been calibrated to match survey data for all model components.
- The model has been dynamically validated by a series of five sensitivity tests and demonstrates reasonable sensitivities to changes in inputs including increased number of lanes at Otay Mesa, increased total border crossings, reduced transit headways for transit services to/from San Ysidro, increased employment near Otay Mesa, and implementing a toll at Otay Mesa.

2.0 DATA DEVELOPMENT

2.1 SURVEY DATA AND EXPANSION

This chapter describes the data used in model development and the process of creating and applying the expansion weights. As part of the Cross Border Travel Model Update project, expansion weights are created to be applied to the processed 2019 cross border survey data. These expansion factors are used to measure the total number of residents of Mexico crossing into the United States at the San Diego County ports of entry (POE) on an average weekday. The expanded travel data is used to develop the SANDAG cross-border travel model. The expansion factors are also used to understand the total number of United States residents entering through POEs, and their US destination. This data is used to calibrate the internal-external (IE), external-internal (EI), and external-external (EE) travel models. Note that survey data expansion factors, which when applied to the survey data represent the total universe of travel, are different from the weights provided in the initial survey data. The initial survey weights are probability weights, in which each record is scaled up or down, such that when applied to the data the sum of the weights equals the total number of records. These are useful for model estimation, but for calibration targets, we need to understand the magnitude of travel. Therefore, we calculated expansion weights for the survey data. There are three main datasets that are used to create the 2019 cross border survey data expansion weights:

- The 2019 cross border survey, collected by True North Research, Inc.[3] on behalf of SANDAG
- 2019 border crossings volumes by lane type and port of entry (POE) from US Customs and Border Protection (CBP)[4]
- The 2016 border delay study, which collected information on border crossings by POE, lane-type, and residence status[5]

The 2019 cross border survey data consisted of two stages. The recruitment sample included general questions about activities in Mexico prior to crossing, method of crossing (i.e. lane type), planned activities within the United States, and a set of questions on willingness to pay for a future tolled crossing at East Otay Mesa. The second stage consisted of a travel diary which is exclusive to cross-border tours. The travel diary included detailed information about every stop/activity

---

[3] True North Research, Inc., *Cross-Border Travel Behavior Survey Summary Report*, (Encinitas: True North Research, Inc., 2020)
[4] Data provided by US Customs and Border Protection for time period between November 10th and 18th, 2019
[5] HDR, *Economic and Air Quality/Climate Impacts of Delays at the Border: Survey Methodology and Plan.* (San Diego: HDR 2016).

within the United States. Therefore, the recruitment questionnaire of the 2019 cross border survey data captured the following three respondent types:

1. US residents who were excluded from the travel diary portion of the survey
2. Residents of Mexico who participated in the travel diary portion of the survey
3. Residents of Mexico who did not participate in the travel diary portion of the survey

The border crossing volumes by lane type contained observed crossing volumes for every hour of eight consecutive days ranging between November 10th to November 18th of 2019. Both vehicular and pedestrian volumes were provided for each of the three POEs: San Ysidro, Otay Mesa and Tecate. The vehicular volumes were provided for both vehicles and passengers in vehicles. The vehicular count volumes were specified by lane type: Dedicated Commuter Lane (i.e. SENTRI), Ready Lane and Regular Lane. Pedestrian volumes also include bus passengers.

SENTRI lanes offer expedited border crossings to pre-qualified citizens of the United States and Mexico. One must apply for a SENTRI pass, which requires extensive background checks. Residents of Mexico must have a valid US Visa, passport, and contact number in the US.

Shortly after the 2010 cross border survey data was collected, the Ready Lane option was added to both the San Ysidro and Otay Mesa port of entries. Like SENTRI, Ready Lane offers expedited service when crossing into the United States. The Ready Lane is only for crossers who have approved RFID-enabled travel cards. Such travel cards include: the U.S. Passport Card, SENTRI card, the new Legal Permanent Resident "green card" and the new Border Crossing Card.

For modeling purposes, it is important to understand the share of residents of Mexico crossing at each POE by lane type and mode, in order to distinguish trips that are modeled in the cross-border travel model versus trips that are modeled in IE, EI, and EE models. Although vehicle passenger count volumes are available by visa status (Alien/Non-Immigrant, Immigrant, and US Citizen), visa status is not a useful variable for expansion because it is not consistent with resident status. Many US citizens live in Mexico and cross into the US to travel before returning home. To resolve this issue, the data from the 2016 border delay study, which involved a random sample of respondents by POE and lane type, including both Mexico and US residents, is utilized to calculate country of residency distributions by POE and lane type (including passengers). The residency distributions are then applied to the observed crossing volumes.

Average weekday daily crossing volumes (showing vehicle passengers), from the observed 2019 border crossing volumes from CBP data, by POE and lane type are shown below in Table 1.

TABLE 1 AVERAGE 2019 DAILY WEEKDAY PERSON CROSSING VOLUMES BY POE AND LANE TYPE

LANE TYPE	SAN YSIDRO	OTAY MESA	TECATE	TOTAL
SENTRI	23,636	8,219	-	31,855
Ready Lane	27,738	17,639	-	45,377
Regular Lane	16,077	2,438	4,825	23,340
Pedestrians	28,862	10,685	2,672	42,219
Total	96,313	38,981	7,497	142,791

Tecate only provides service to Regular Lane and Pedestrians

The residency distribution by POE and lane type from the 2016 border delay study are shown below in Table 2. As can be seen in the table residents of Mexico are broken out by whether they return the POE at the end of the day (and are therefore modeled explicitly by the cross-border travel model). US residents are broken out by county of residents; San Diego County resident travel is modeled by the resident internal-external model, while non-San Diego County travel is modeled by the external-external model. Note that within each POE and lane type, the percentages add up to 100% (with some rounding error).

TABLE 2 RESIDENCY DISTRIBUTION BY POE AND LANE TYPE

Lane Type	Resident Type		San Ysidro	Otay Mesa	Tecate
SENTRI	**MX**	**Return**	77.4%	80.5%	0.0%
		Not Return	1.9%	2.7%	0.0%
	US	**Non-SD County**	3.3%	2.5%	0.0%
		SD County	17.4%	14.3%	0.0%
Ready Lane	**MX**	**Return**	78.2%	84.9%	0.0%
		Not Return	3.9%	2.3%	0.0%
	US	**Non-SD County**	1.9%	2.5%	0.0%
		SD County	16.0%	10.3%	0.0%
Regular Lane	**MX**	**Return**	65.2%	50.2%	80.6%
		Not Return	2.1%	2.5%	5.9%
	US	**Non-SD County**	9.8%	15.0%	8.4%

		SD County	22.9%	32.3%	5.2%
Pedestrians	MX	Return	87.4%	84.7%	86.1%
		Not Return	2.8%	4.7%	1.4%
	US	Non-SD County	2.9%	1.6%	12.5%
		SD County	6.8%	9.0%	0.0%
Total	MX	Return	76.8%	78.0%	83.1%
		Not Return	2.7%	2.9%	3.6%
	US	Non-SD County	4.7%	4.5%	9.4%
		SD County	15.8%	14.6%	3.9%

Finally, the residency distributions, from Table 2, applied to the observed border crossing volumes, from Table 1, are shown below in Table 3.

TABLE 3 PERSON CROSSING VOLUMES BY POE AND LANE TYPE

Lane Type	Resident Type		San Ysidro	Otay Mesa	Tecate	Total
SENTRI	MX	Return	18,294	6,617	-	24,911
		Not Return	459	222	-	681
	US	Non-SD County	781	205	-	986
		SD County	4,102	1,174	-	5,276
	Total		*23,636*	*8,218*	-	*31,854*
Ready Lane	MX	Return	21,704	14,972	-	36,676
		Not Return	1,075	406	-	1,481
	US	Non-SD County	522	444	-	966
		SD County	4,437	1,816	-	6,253
	Total		*27,738*	*17,638*	-	*45,376*
Regular Lane	MX	Return	10,482	1,225	3,888	15,595
		Not Return	334	61	283	678
	US	Non-SD County	1,572	365	405	2,342
		SD County	3,689	787	249	4,725
	Total		*16,077*	*2,438*	*4,825*	*23,340*

Pedestrians	**MX**	**Return**	25,229	9,046	2,300	36,575
		Not Return	818	503	38	1,359
	US	**Non-SD County**	840	169	334	1,343
		SD County	1,975	967	0	2,942
	Total		*28,862*	*10,685*	*2,672*	*42,219*
Total	**MX**	**Return**	75,709	31,860	6,188	113,757
		Not Return	2,686	1,192	321	4,199
	US	**Non-SD County**	3,715	1,183	739	5,637
		SD County	14,203	4,744	249	19,196
	Total		*96,313*	*38,979*	*7,497*	*142,789*

Expansion Weights

Resident of Mexico "returning" crossing volumes, listed in Table 3, are utilized to create the expansion factors for the travel diary portion of the survey. First, both the 2019 cross border survey and border crossing volumes datasets are categorized into 10 expansion weight categories, which are based on the 3 POEs (San Ysidro, Otay Mesa and Tecate) and 4 crossing types (SENTRI, Ready Lane, Regular Lane, and Passengers). The Tecate POE only provides service to two crossing types: Regular Lane and Passengers.

Next, average daily person crossing volumes are calculated for each of the 10 expansion weight categories. For vehicular crossings, *passenger* volumes are used to calculate expansion factors, so that application of the factors result in *person* tours, as opposed to a mixture of person and vehicle tours. Similarly, total completed survey responses (considering party size) are calculated for each of the 10 expansion weight categories for the 2019 cross border survey. When calculating total completed survey responses, the 2019 cross border survey diary probability weights, which account for sample error, are applied.

Finally, the 2019 cross border survey and border crossing volume datasets are joined using the expansion categories as join keys. The final expansion weights for each weight category consisted of the ratio between the estimated average daily crossing volumes (from the observed border crossing volumes) and the total completed survey respondents (considering party size).

The resulting weights for each of the 10 expansion weight categories are shown below in Table 4. Note that in order to account for the fact that cross-border travel parties can consist of multiple persons, the expanded diary weights for cross-border tours are multiplied by the total number of travelers (i.e. party size) to ensure consistency with the target number of crossings.

TABLE 4 EXPANSION WEIGHTS

EXPANSION WEIGHT CATEGORY	CATEGORY DESCRIPTION	AVERAGE DAILY CROSSINGS	TOTAL COMPLETED SURVEYS	EXPANSION WEIGHT
1	San Ysidro SENTRI	18,294	608	30.1
2	San Ysidro Ready Lane	21,704	625	34.7
3	San Ysidro Regular Lane	10,482	398	26.3
4	San Ysidro Pedestrian	25,229	731	34.5
5	Otay Mesa SENTRI	6,617	183	36.2
6	Otay Mesa Ready Lane	14,972	373	40.1
7	Otay Mesa Regular Lane	1,225	53	23.1
8	Otay Mesa Pedestrian	9,046	282	32.1
9	Tecate Regular Lane	3,888	91	42.7
10	Tecate Pedestrian	2,300	81	28.4
	Total	113,758	3,425	33.2

Border Crossing Wait Time Data In the original cross-border travel model, the wait time is fixed for each POE and border crossing mode. In the revised model, a data-driven method is used to estimate the cross-border wait time model based on observed border crossings at three ports of entry: San Ysidro, Otay Mesa, and Tecate. This model is applied within the Cross-Border Model system, in order to reflect the dynamic relationship between wait time and demand.

Data

Data from San Ysidro, Otay Mesa and Tecate border crossings are studied. Vehicle crossings per hour and lane type (SENTRI, Ready, Standard, Pedestrian) between November 10 and November 18, 2019 was obtained from Customs and Border Protection. Only data weekdays is used, leaving November 11, 12, 13, 14, 15, and 18 for this sample. Border crossing wait times by hour and lane type are also included from CBP for the given date range. The vehicle wait time data set is incomplete, therefore not every observed crossing volume has a recorded crossing wait time. Table 5 shows the number of vehicle volumes by hour from the CBP dataset and the number of corresponding wait times for those records.

TABLE 5: NUMBER OF OBSERVATIONS IN BORDER CROSSING DATA

PORT	VEHICLE VOLUMES	VEHICLE WAIT TIMES	PED VOLUMES	PED WAIT TIME
San Ysidro	432	231	144	77
Otay Mesa	389	292	144	108
Tecate	108	76	108	26

From the available data, the crossing wait times are joined to the crossing volume by hour, lane type, and port of entry. The observed volumes are divided by the number of lanes at the respective port of entry giving the volume/lane. Table 6 shows the number of lanes at each port of entry.

TABLE 6: NUMBER OF LANES BY POE

PORT	NUMBER OF VEHICLE LANES	NUMBER OF PEDESTRIAN LANES
San Ysidro	24	15
Otay Mesa	13	6
Tecate	2	2

The relationship between volume/lane and wait time for each port of entry is shown in Figure 1 through Figure 3.

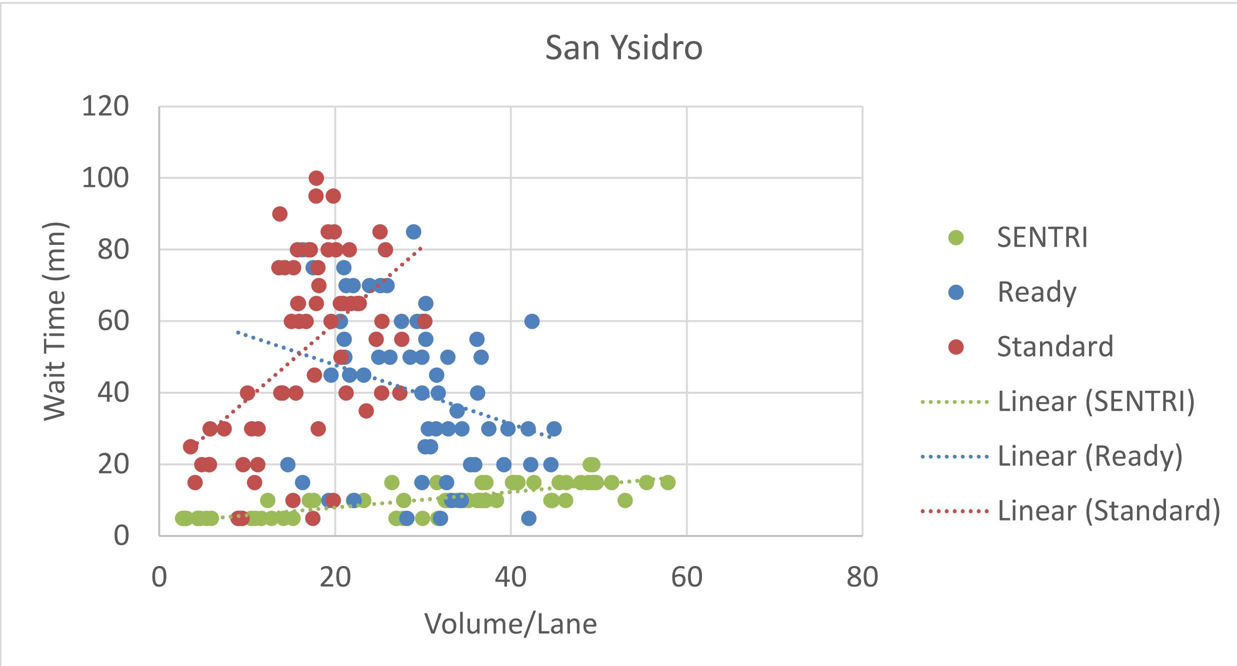


FIGURE 1: SAN YSIDRO BORDER WAIT TIMES BY VOLUME PER LANE

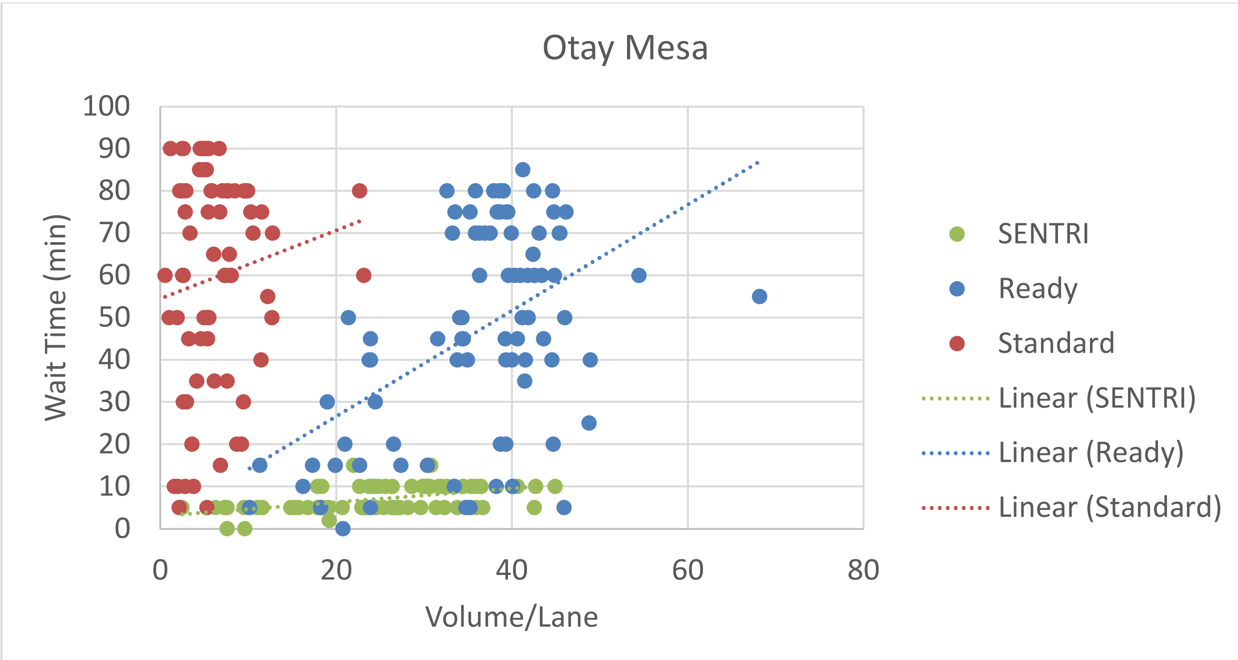


FIGURE 2: OTAY MESA BORDER WAIT TIMES BY VOLUME PER LANE

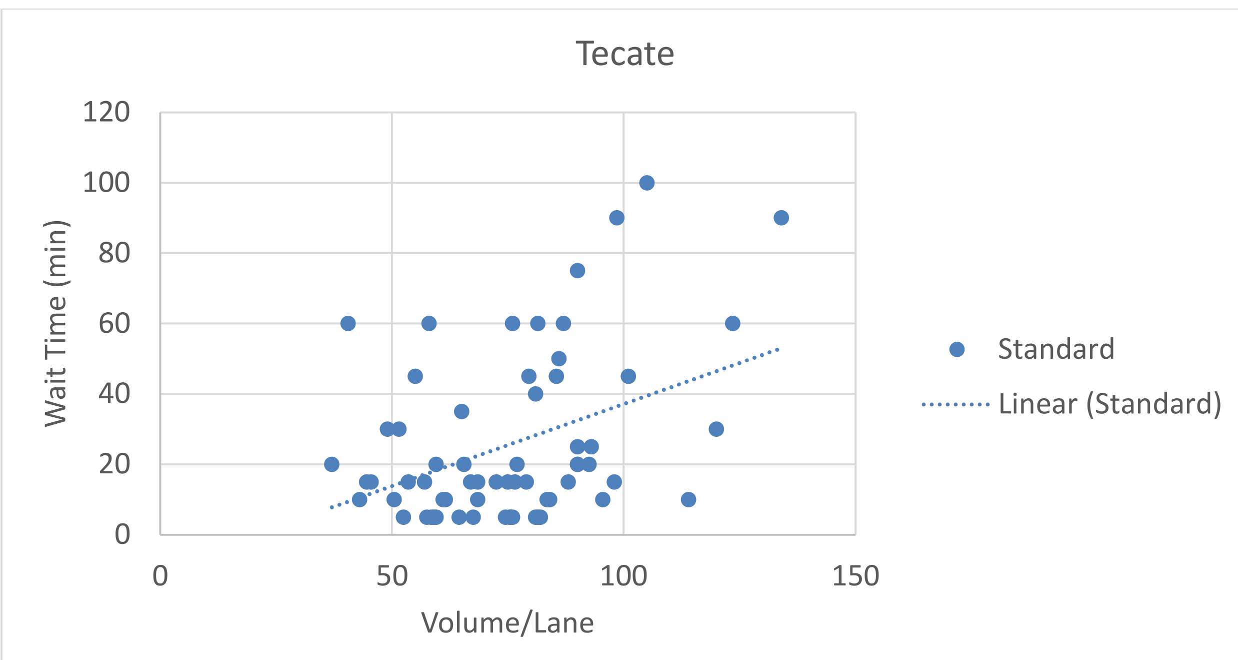


FIGURE 3: TECATE BORDER WAIT TIMES BY VOLUME PER LANE

The linear approximation seems to fit the data well, suggesting a linear regression may be used to fit the data. However, at San Ysidro the Ready lane wait time shows a negatively correlated relationship with volume per lane. This may suggest that as volume at the Ready lanes increases more Ready lanes are opened as a response resulting in the negative correlation. Another explanation is that the CBP does not necessarily prioritize a fast border crossing; their main priority is to enforce the law and ensure the safety and integrity of US border. In addition, the number of border crossers that can be processed at any given time is also a function of the number of customs agents working at each POE. It is possible that for a given time period, the number of agents available may have been a constraint on the rate at which northbound travelers could be processed.

The pedestrian wait times by Volume/Lane also show a positive linear relationship at each port of entry. Pedestrian wait times are shown in Figure 4 through Figure 6.

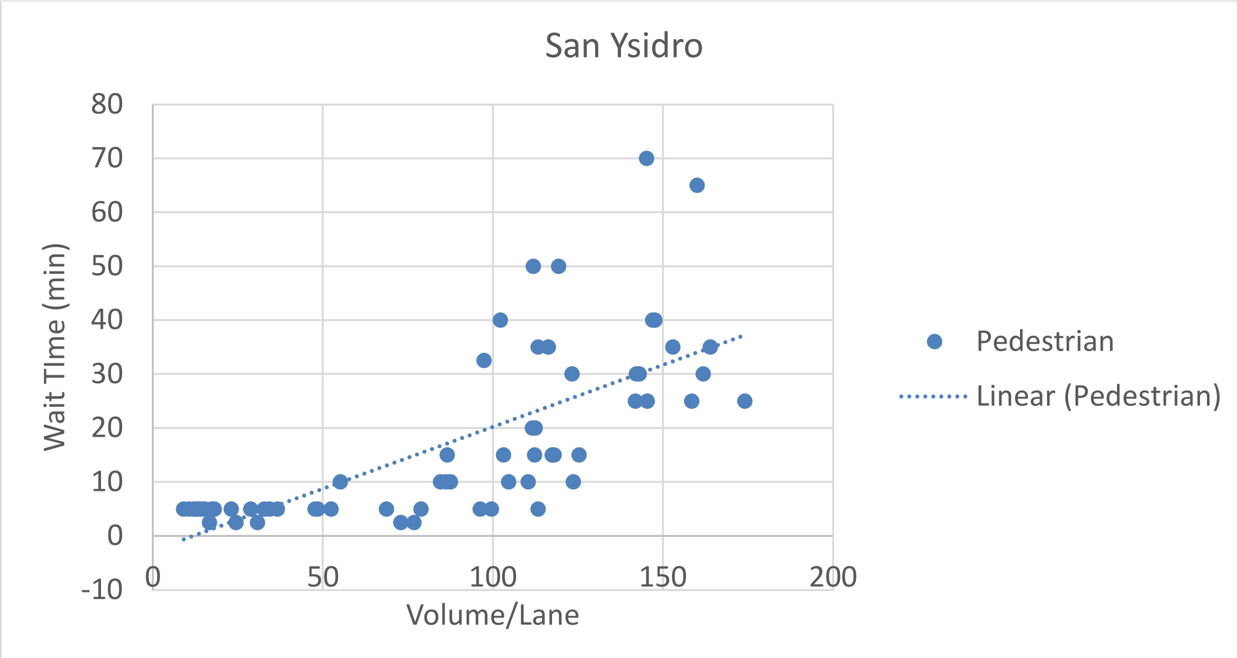


FIGURE 4: SAN YSIDRO PEDESTRIAN WAIT TIME BY VOLUME PER LANE

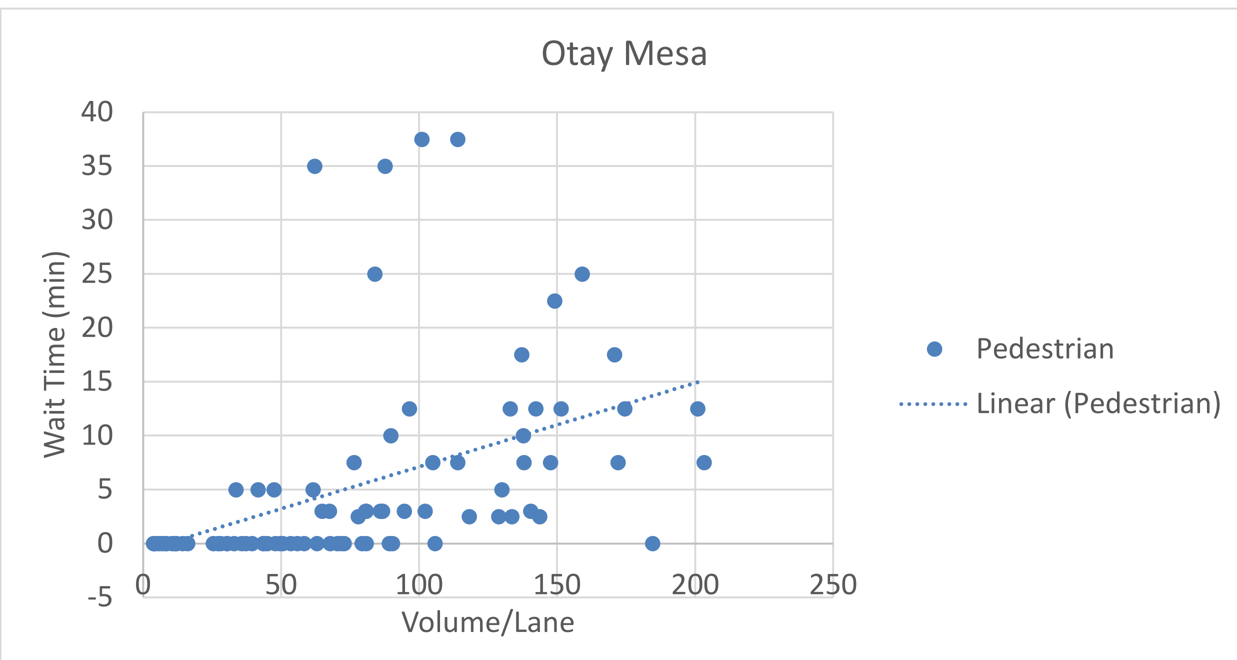


FIGURE 5: OTAY MESA PEDESTRIAN WAIT TIME BY VOLUME PER LANE

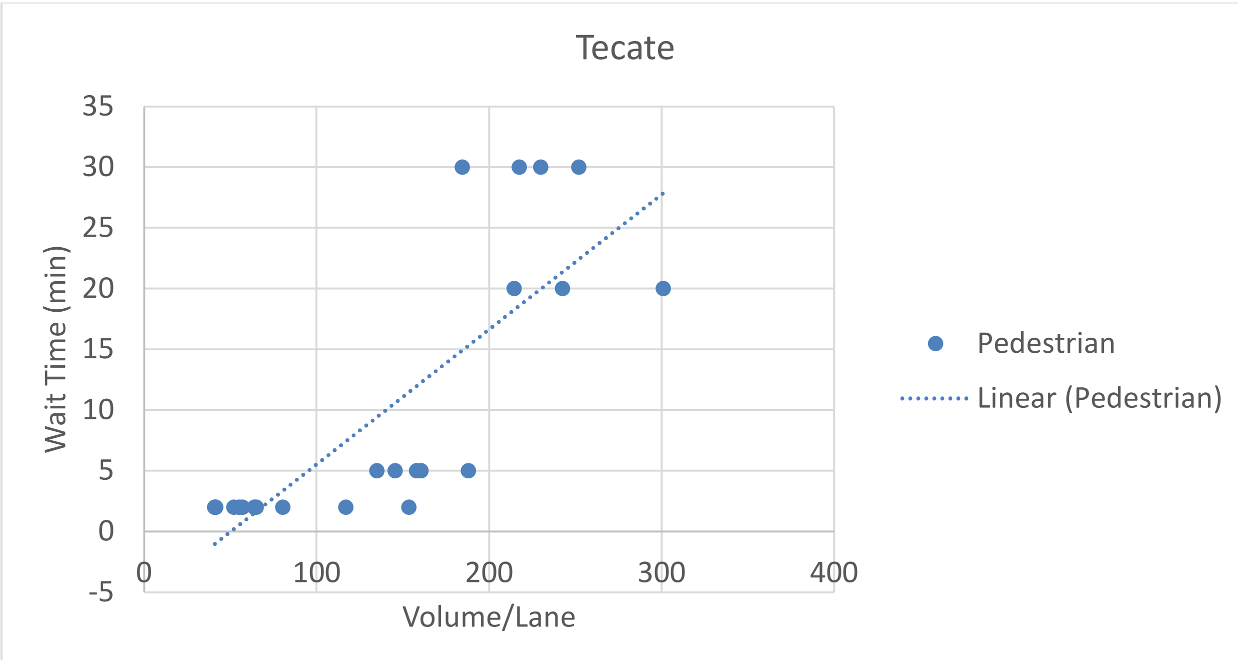


FIGURE 6: TECATE PEDESTRIAN WAIT TIME BY VOLUME PER LANE

The wait times are also evaluated by time of day shown in Figure 7 through Figure 12. The wait times vary strongly in the early AM (midnight to 3am) and evening (8pm to midnight) hours suggesting that time of day factors may be appropriate here noting that the Tecate port of entry does not operate before 5am or after 11pm.

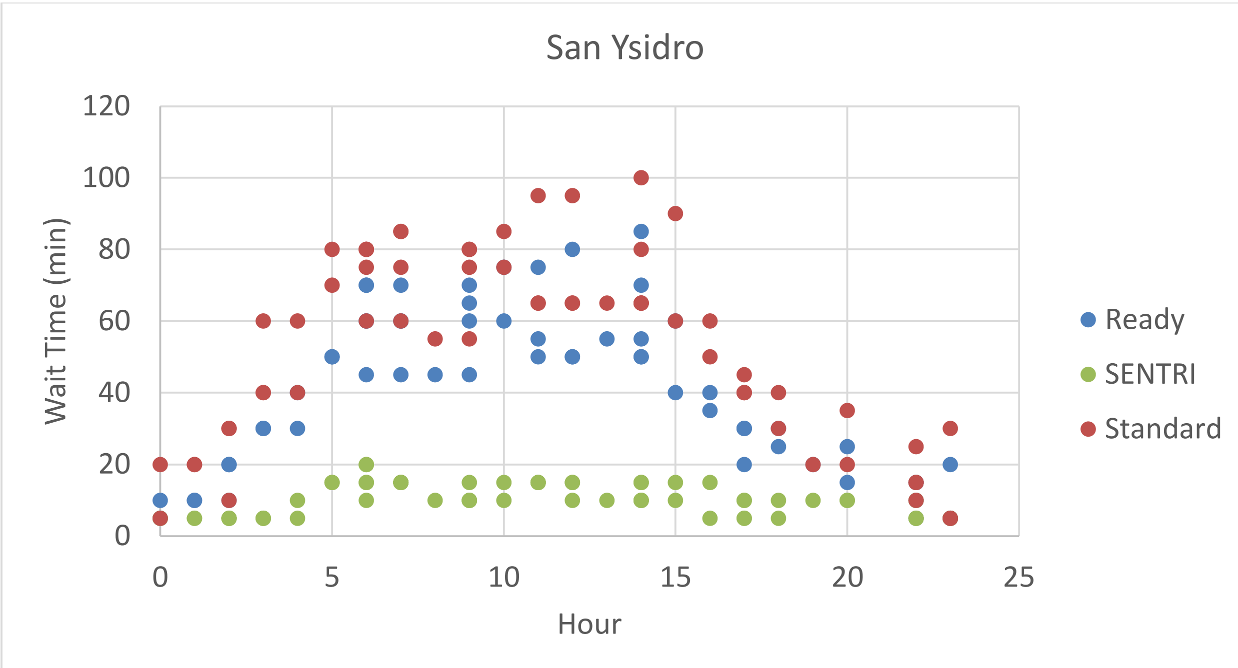


FIGURE 7: SAN YSIDRO WAIT TIMES BY TIME OF DAY

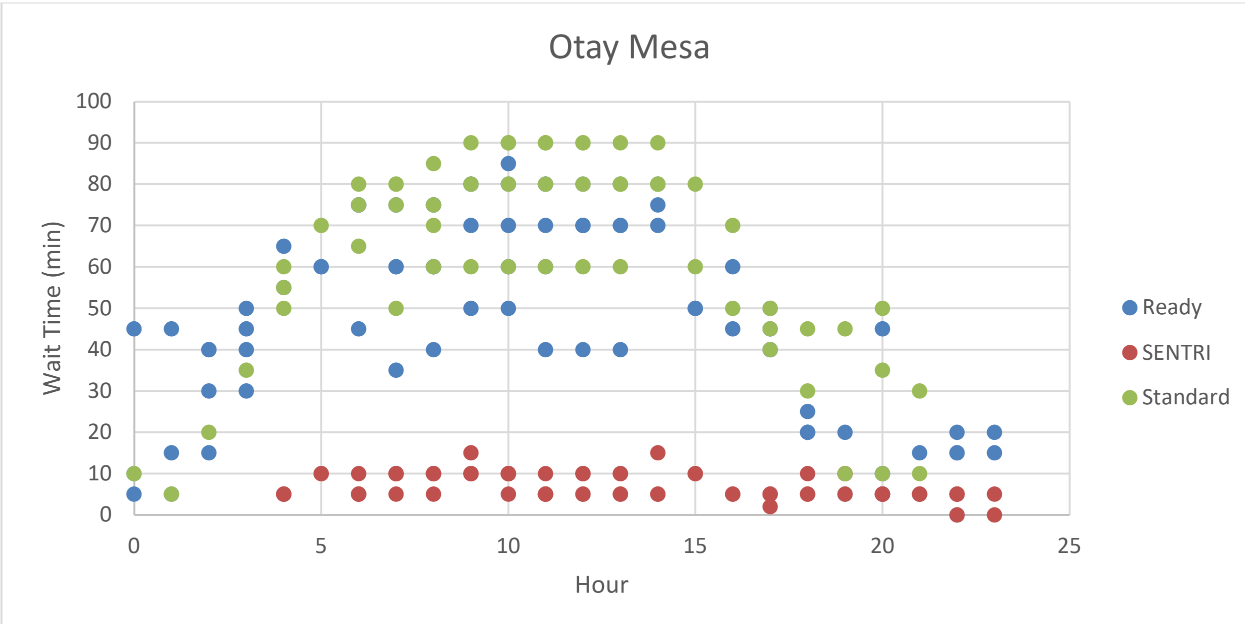


FIGURE 8: OTAY MESA WAIT TIMES BY TIME OF DAY

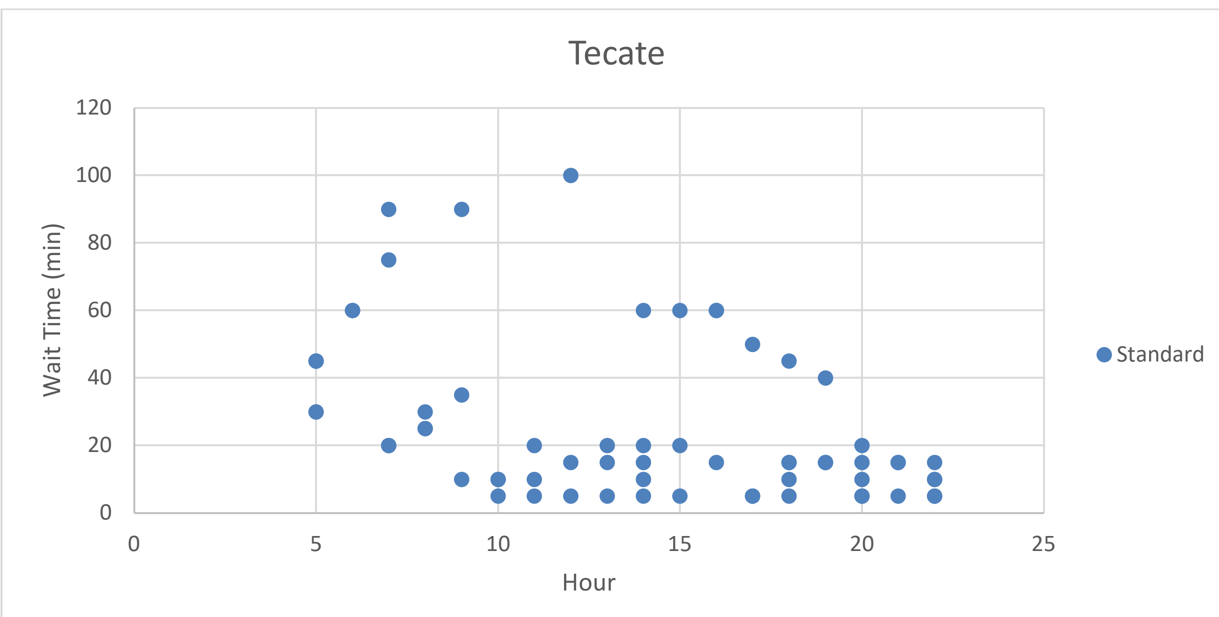


FIGURE 9: TECATE WAIT TIMES BY TIME OF DAY

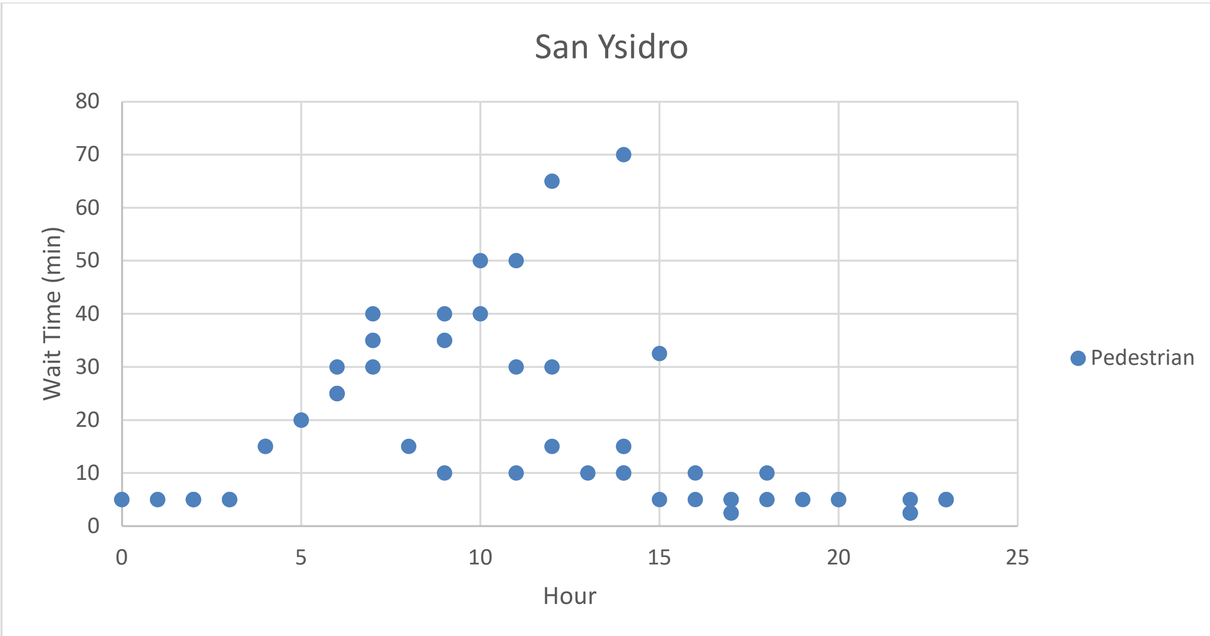


FIGURE 10: SAN YSIDRO PEDESTRIAN WAIT TIME BY TIME OF DAY

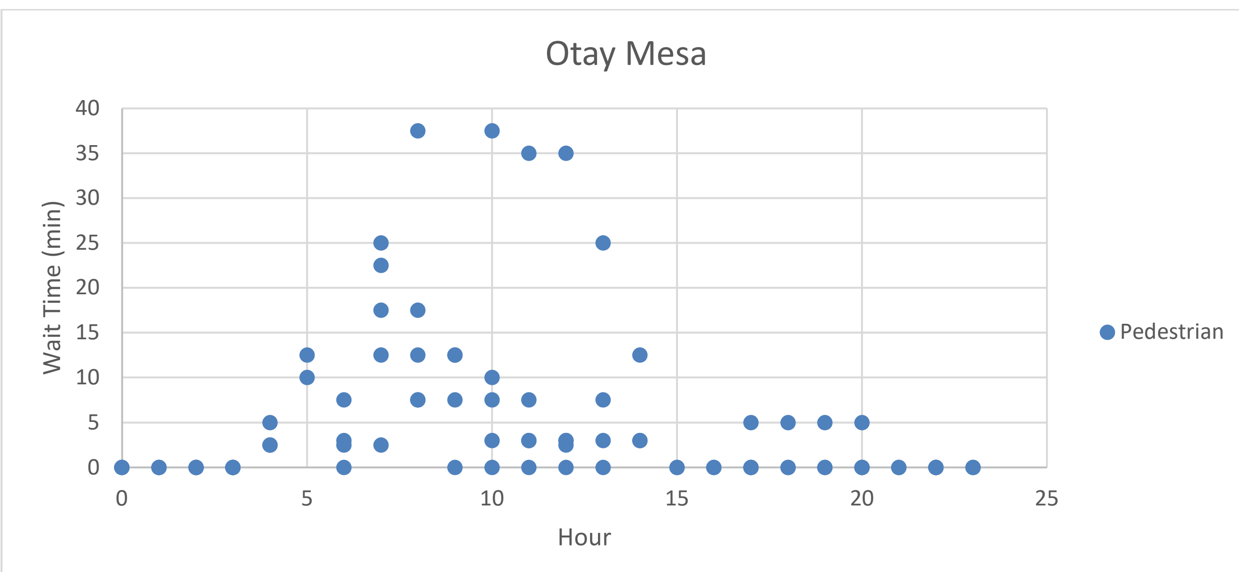


FIGURE 11: OTAY MESA PEDESTRIAN WAIT TIME BY TIME OF DAY

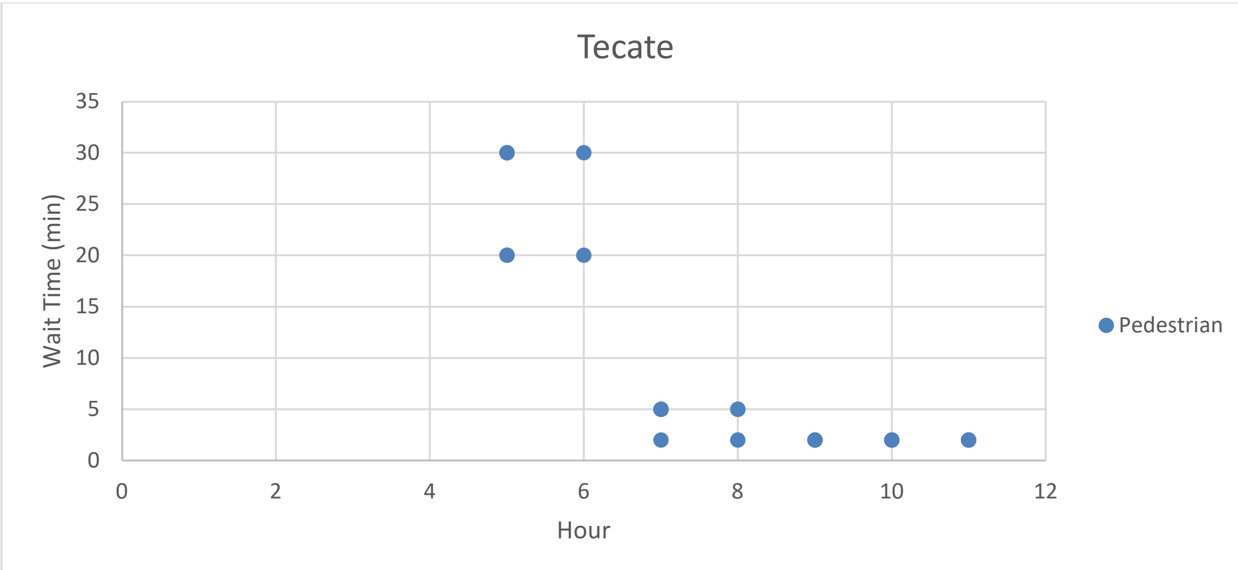


FIGURE 12: TECATE PEDESTRIAN WAIT TIME BY HOUR

2.2 TRAVEL SUMMARIES

This section summarizes observed travel patterns for cross-border tours from the 2019 survey data.

A cross-border tour is a sequence of trips starting and ending at the POE with at least one destination activity within San Diego County and no activities outside of San Diego County. The model assumes symmetry with respect to border crossing location; in other words, the travel party is assumed to depart the same POE that they arrive at. All tours are closed; in other words, they start at the POE, engage in one or more activities in San Diego County, then depart the US at the same POE. Each tour is composed of one primary destination and one or more intermediate stops. An intermediate stop is an activity location other than the primary destination and occurs on either the "outbound" journey (between the POE and the primary destination) or the "inbound" journey (between the primary destination and the POE).

Tour Purpose

There are five tour purposes in the new cross-border travel model, as shown in Table 7. They are coded based on the activity purposes engaged by the traveler in the United States, according to hierarchy of activity purposes as follows:

- *Work*: At least one trip on the tour is for working in the US.
- *School*: At least one trip on the tour is made for attending school in the US, and no work trips were made on the tour.

- *Shop*: No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was shopping in the US.
- *Visit*: No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was visiting friends\relatives in the US.
- *Other*: No trips on the tour were made for work, school, or cargo, and the activity with the longest duration on the tour was other (collapsed escort, eat, personal, medical, recreation, sport, and other activity purposes).

It should be noted that the original Cross-Border model represents an additional Cargo purpose, in which at least one activity is made for picking up or dropping off cargo, and no work or school trips were made on the tour. The 2019 cross border survey does not capture cargo at the trip purpose level and is only recorded as a stated primary purpose for crossing into the United States. In the 2010 data, any tour with a cargo trip is recoded as a cargo tour, so the number of cargo tours in the 2010 data is much higher than the number of cargo tours in the 2019 data. Therefore, we only show cargo tours in cases where we compare the 2019 survey data to 2010 data. In cases where we summarize only 2019 data, we group cargo tours with shop tours and do not maintain a separate cargo purpose for travel modeling.

The 2010 and 2019 cross border tour volumes, by tour purpose, are shown below in Table 7, Figure 13 and Figure 14. The tables summarize person tours (tours expanded by participants).

TABLE 7 TOUR PURPOSE

TOUR PURPOSE	NUMBER OF TOURS (RAW)		NUMBER OF TOURS (EXPANDED)		NUMBER OF TRIPS ON TOUR (RAW)		NUMBER OF TRIPS ON TOUR (EXPANDED)	
Survey Year	**2010**	**2019**	**2010**	**2019**	**2010**	**2019**	**2010**	**2019**
Work	233	435	11,055	21,150	686	1,621	31,302	82678
	15%	23%	12%	19%	12%	21%	9%	18%
School	64	58	3,182	3,746	201	242	9,909	17,798
	4%	3%	3%	3%	4%	3%	3%	4%
Cargo	210	6	11,840	250	853	20	48,364	767

TOUR PURPOSE	NUMBER OF TOURS (RAW)		NUMBER OF TOURS (EXPANDED)		NUMBER OF TRIPS ON TOUR (RAW)		NUMBER OF TRIPS ON TOUR (EXPANDED)	
	14%	0%	13%	0%	15%	0%	14%	0%
Shop	690	1,176	46,275	74,698	2,626	5,179	180,411	323,334
	45%	62%	49%	66%	47%	67%	52%	70%
Visit	123	66	8,525	4,169	446	205	30,178	13,286
	8%	3%	9%	4%	8%	3%	9%	3%
Other	197	171	12,853	9,743	738	488	48,175	26,797
	13%	9%	14%	8%	13%	6%	14%	6%
Total	1,517	1,911	93,730	113,756	5,550	7,754	348,338	464,660
	100%	100%	100%	100%	100%	100%	100%	100%

FIGURE 13 EXPANDED PERSON TOURS BY TOUR PURPOSE

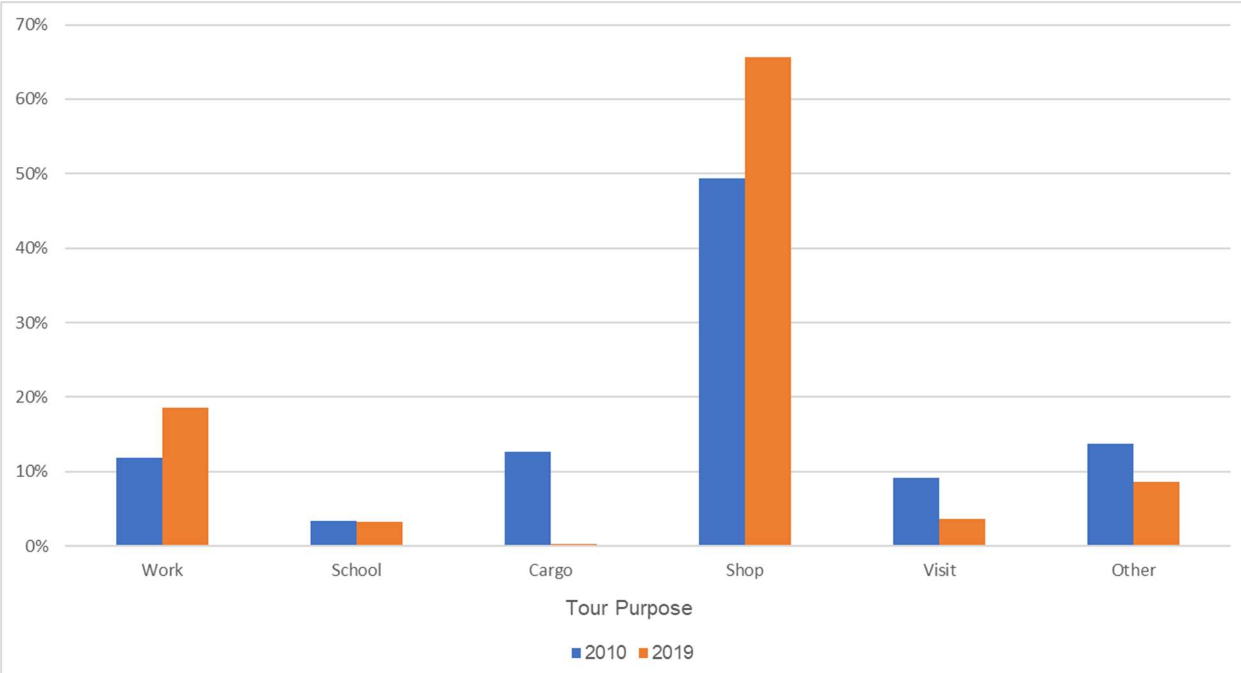
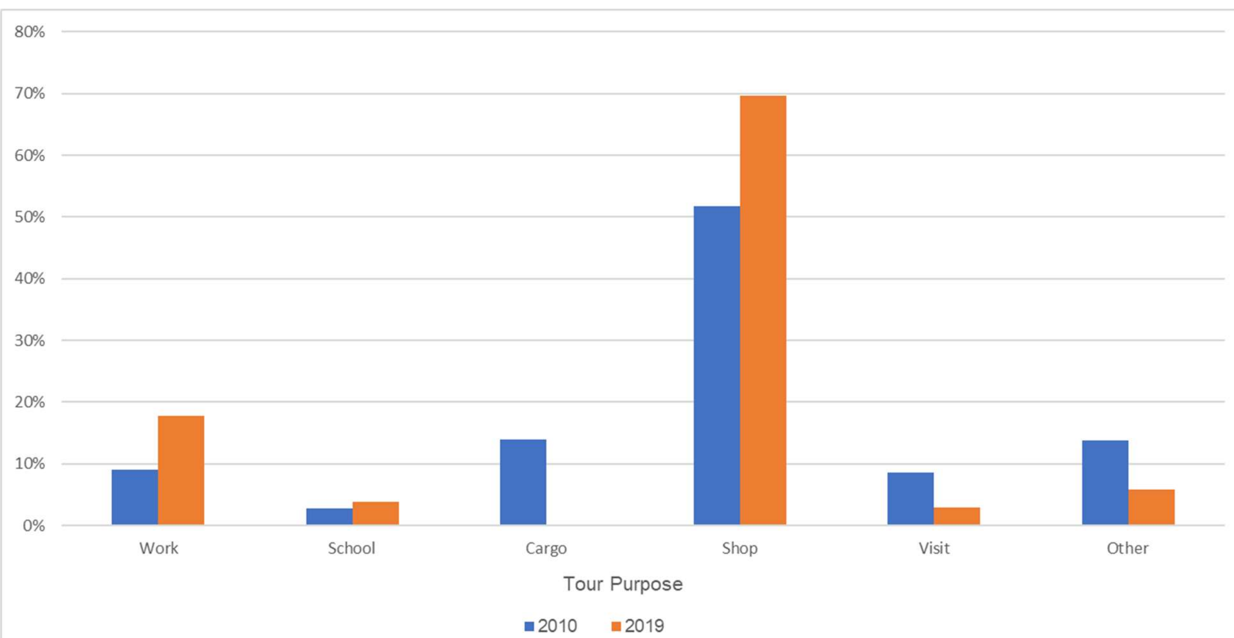


FIGURE 14 EXPANDED PERSON TRIPS BY TOUR PURPOSE

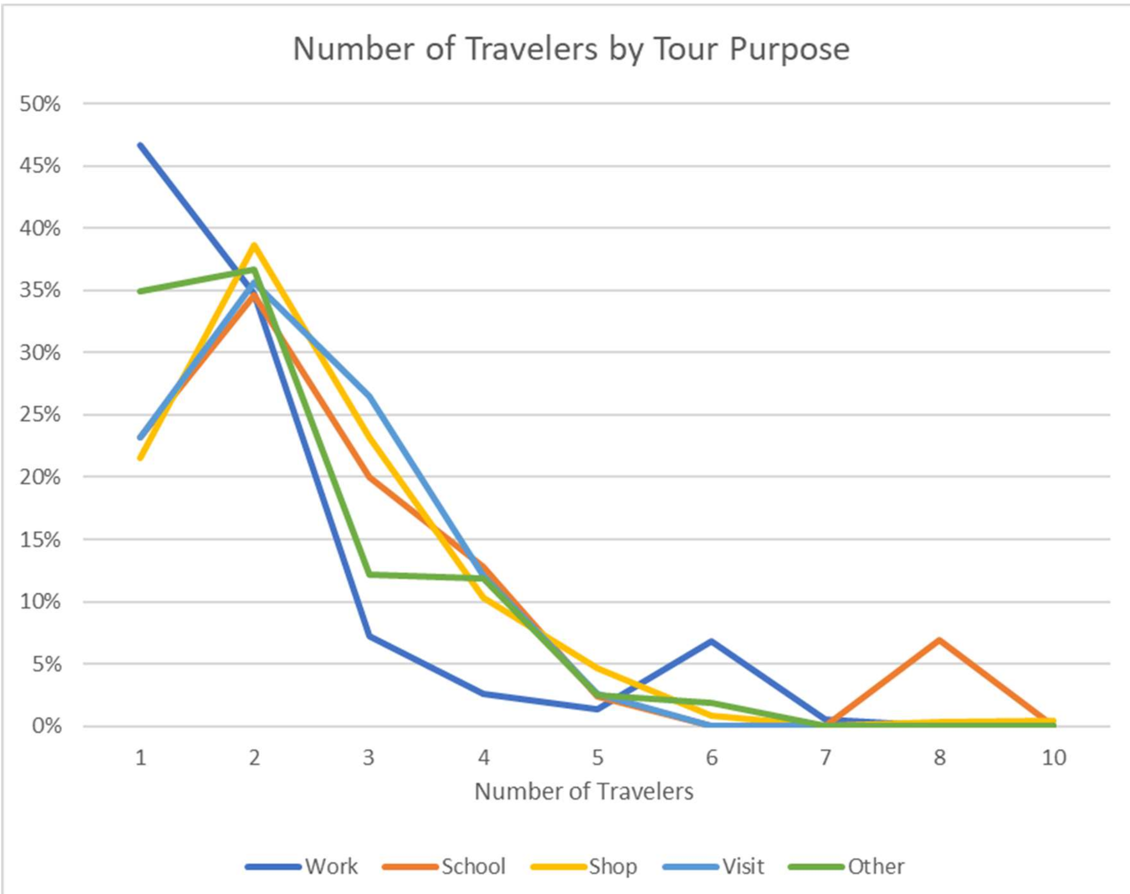


Tours by the total number of outbound and inbound stops by tour purpose is shown below in Figure 15. Outbound stops refer to stops in the outbound direction (towards primary destination) of a travelers' anchor location (i.e. home or work) while inbound stops refer to stops in the inbound direction (towards anchor location). In the case of cross-border tours, outbound stops refer to those stops that are made in the time between (after having crossed the border (anchor location) into the United States and before arriving at their primary destination). Inbound stops refer to those stops that are made between the primary destination and the border. The figures below show that in the outbound direction, other tours had the largest share of tours with no stops, visit tours following with the next highest share. On the other hand, school tours had the largest share of three and four stops. In the inbound direction, visit tours had the largest share of tours with no stops, with other tours being a close second. Shopping had the largest share of 2 or 3 stops in the inbound direction, while school had the highest share of 1 and 4 inbound stops. Overall, there are more stops in the inbound direction.

FIGURE 15 PERCENT OF TOURS BY NUMBER OF OUTBOUND AND INBOUND STOPS, BY TOUR PURPOSE

Figure 16 shows tours by number of participants by tour purpose. The figure shows that work tours had the largest share of tours with only one participant while school, shop, and visit tours had the largest share of tours with more participants, each tour type having a very similar distribution for high occupancy tours. The average number of participants is about 2.3.

FIGURE 16 TOURS BY NUMBER OF PARTICIPANTS, BY TOUR PURPOSE



Tour Mode

The tour mode is the mode used to cross the border, which conditions the mode used for all trips on the tour, including the trip from the border crossing to the first destination in the United States. The tour modes are listed in Table 8 and are defined by whether the border is crossed via auto or by foot and the occupancy if by auto. The tables and summaries in this section also segment the tour by pass type (no pass or standard, SENTRI, or Ready).

The number of tours by tour mode, pass type, POE and survey year are shown below in Table 8. Additionally, Figure 17 charts the data for San Ysidro, Otay Mesa and Tecate. Note that Shared-ride 2 and Shared-ride 3+ tour modes are showing total number of travelers rather than total number of parties.

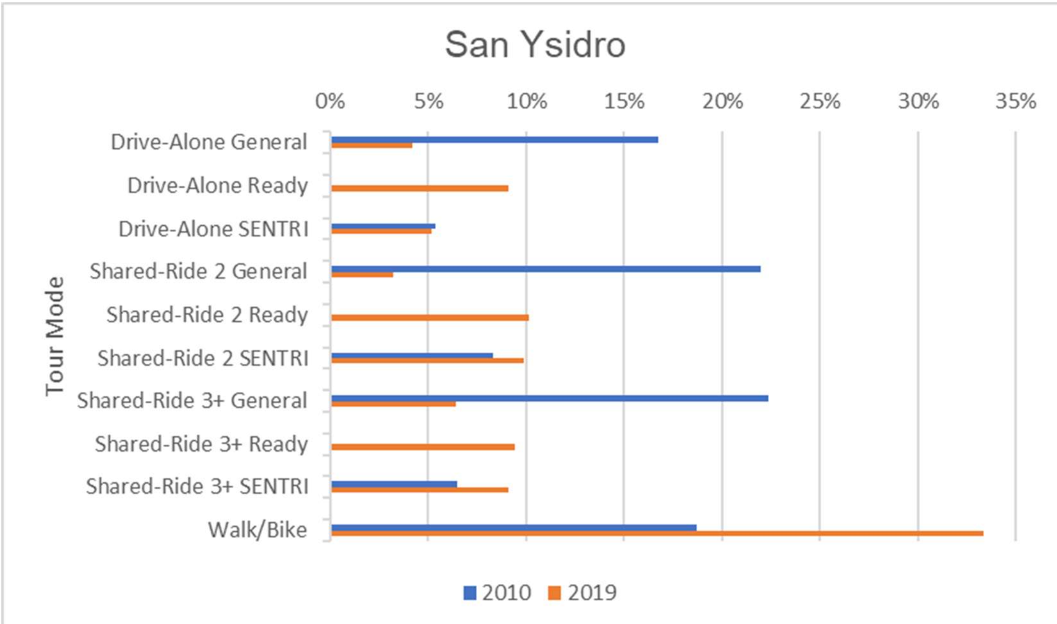
The data shows that there is an increase in overall crossing volumes across all three POEs. It is also interesting to note that with Ready Lanes in place, there is a decrease in general lane crossing volumes for both San Ysidro and Otay Mesa (no Ready Lane at Tecate). Further, the data is consistent with the True North summary report[6] in showing that general lanes received the lower share of crossing volumes as compared to SENTRI and Ready Lane.

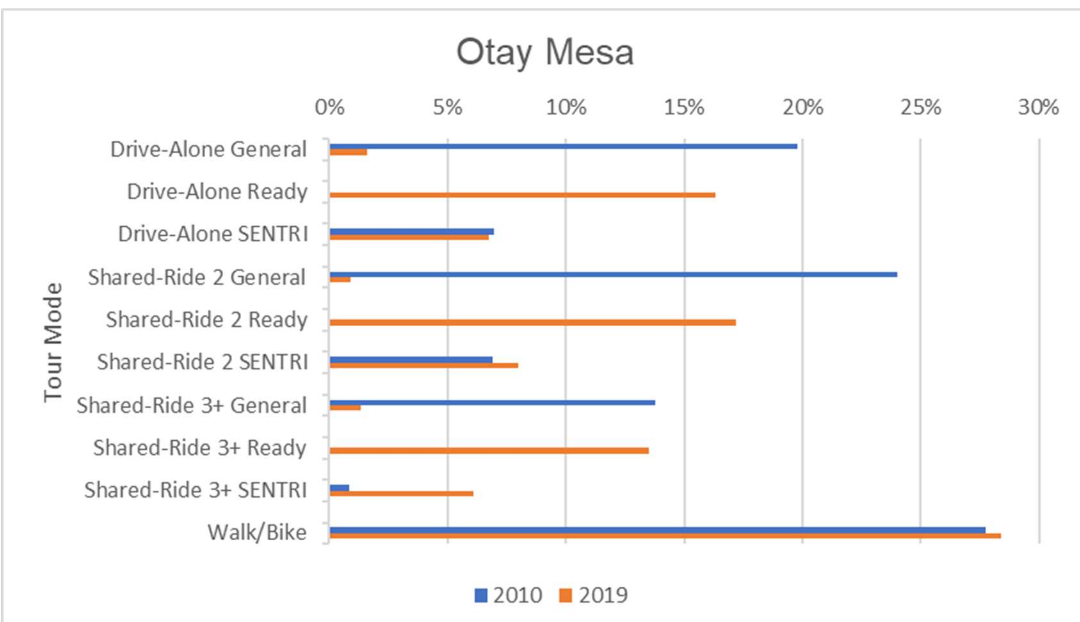
TABLE 8 TOUR MODE AND PASS TYPE BY PORT OF ENTRY

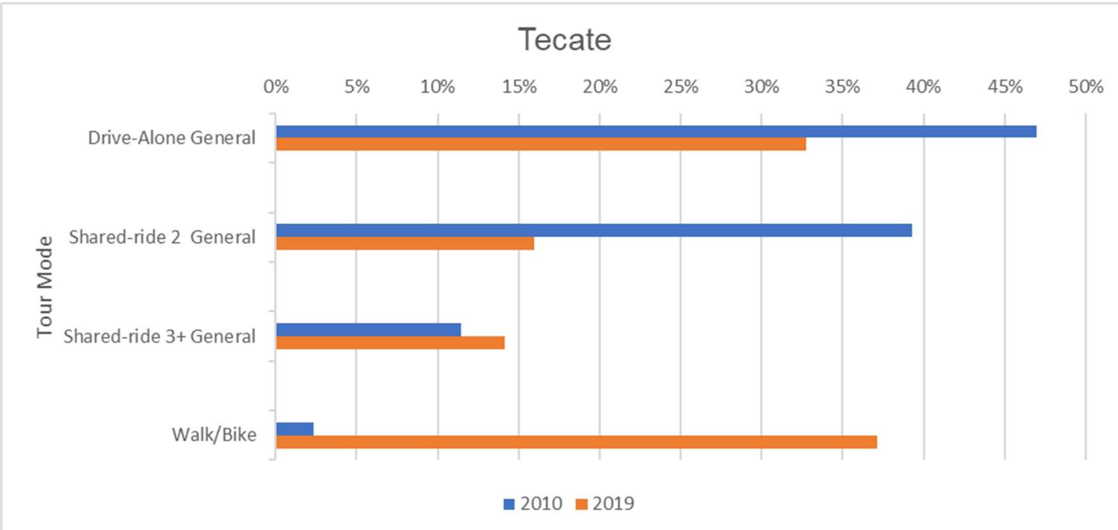
TOUR MODE AND PASS TYPE	NUMBER OF TOURS							
	San Ysidro		Otay Mesa		Tecate		Total	
	2010	2019	2010	2019	2010	2019	2010	2019
Drive-Alone Standard	11,970	10,062	3,623	5,703	1,854	2,027	17,446	17,792
	17%	13%	20%	18%	47%	33%	19%	16%
Drive-Alone SENTRI	3,840	3,912	1,269	2,145	-	-	5,109	6,057
	5%	5%	7%	7%	0%	0%	5%	5%
Drive-Alone Ready	-	6,885	-	5,194	-	-	-	12,079
	-	9%	0%	16%	0%	0%	0%	11%
Shared-ride 2 Standard	15,720	10,154	4,398	5,759	1,551	987	21,669	16,900
	22%	13%	24%	18%	39%	16%	23%	15%
Shared-ride 2 SENTRI	5,966	7,506	1,259	2,538	-	-	7,224	10,044
	8%	10%	7%	8%	0%	0%	8%	9%
Shared-ride 2 Ready	-	7,693	-	5,472	-	-	-	13,166
	0%	10%	0%	17%	0%	0%	0%	12%
Shared-ride 3+ Standard	16,005	11,970	2,520	4,735	451	874	18,976	17,579
	22%	16%	14%	15%	11%	14%	20%	15%
Shared-ride 3+ SENTRI	4,620	6,875	158	1,934	-	-	4,778	8,809
	6%	9%	1%	6%	0%	0%	5%	8%
Shared-ride 3+ Ready	-	7,126	-	4,306	-	-	-	11,431
	0%	9%	0%	14%	0%	0%	0%	10%
Walk	13,362	25,229	5,074	9,046	92	2,299	18,527	36,575
	19%	33%	28%	28%	2%	37%	20%	32%
Total	71,482	75,709	18,302	31,860	3,947	6,188	93,730	113,756

[6] True North Research, Inc., *Cross-Border Travel Behavior Survey Summary Report*, (Encinitas: True North Research, Inc., 2020), 43

FIGURE 17 SHARE OF TOURS BY TOUR MODE AND POE







Trip Mode

The 2019 cross border travel diary did not collect number of vehicle occupants and the trip mode variable did not differentiate between shared 2 and shared 3+ modes. Therefore, in order to differentiate shared 2 from shared 3+ trips, we use the number of tour participants (number of persons in party crossing the border). We assume that the trip mode is shared 2 if the respondent reported crossing the border alone or with one other person and the reported trip mode is carpool. It is possible in such cases that a person crossing in a vehicle with a non-household member might respond that they crossed 'alone' even though there is someone else in the vehicle. In such cases, we let the carpool variable define the trip mode as shared ride 2. If they report a travel party with three tour participants we assume that the trip mode is shared 3+.

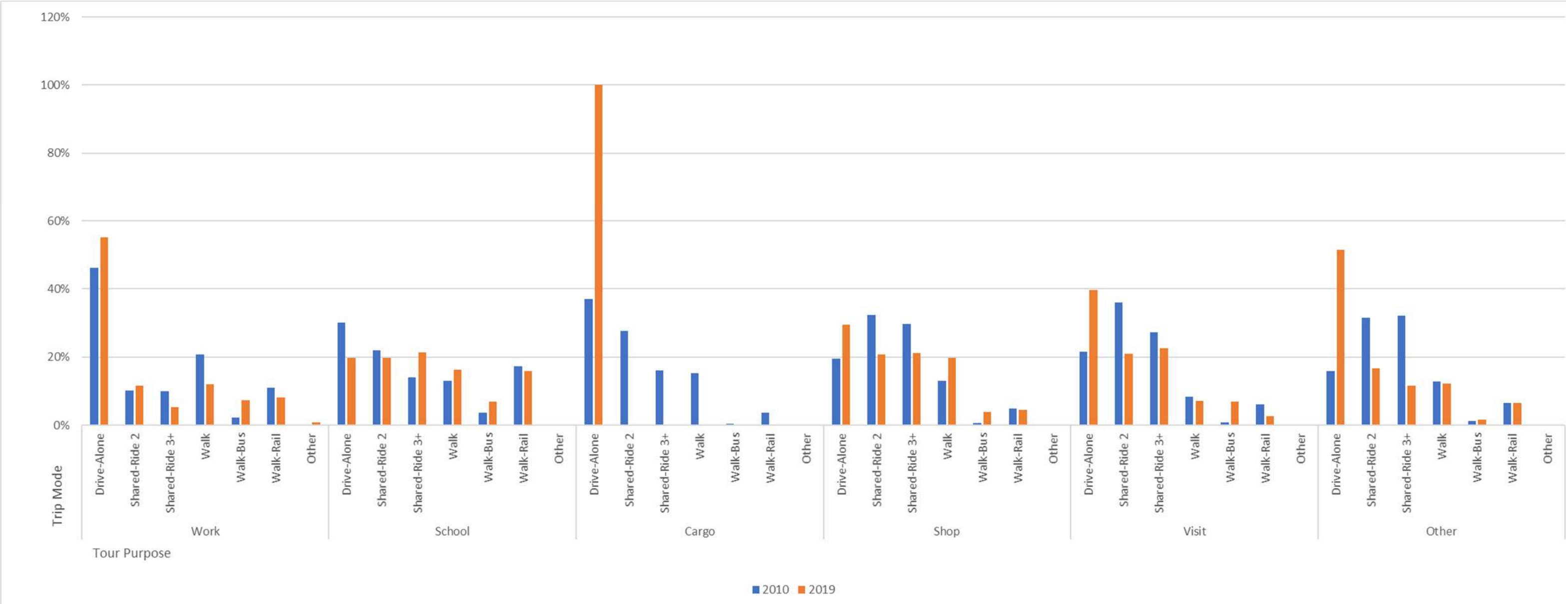
The trip modes used in the cross-border travel model are the same modes available in the resident activity-based model, as shown in Table 9.

Trips by tour purpose, trip mode and survey year are shown below in Table 9 and Figure 18. When comparing the 2010 and 2019 data, the summaries show that there is an increase in number of trips across work, school, and shopping modes while there is a decrease in trips from visit, cargo and other tour purposes. In terms of share of trips per trip mode, the share of drive-alone trip shares increased for all tour purposes except for school. Walk-bus trip shares increased across all purposes, while walk-rail trip shares decreased.

TABLE 9 TRIPS BY TOUR PURPOSE AND TRIP MODE

TRIP MODE	TOUR PURPOSE													
	Work		School		Cargo		Shop		Visit		Other		Total	
	2010	2019	2010	2019	2010	2019	2010	2019	2010	2019	2010	2019	2010	2019
Drive-Alone	14,439	45,554	2,990	3,506	17,895	767	35,150	95,419	6,499	5,280	7,642	13,773	84,614	164,300
	46%	55%	30%	20%	37%	100%	19%	30%	22%	40%	16%	51%	24%	35%
Shared-Ride 2	3,174	9,554	2,167	3,514	13,381	-	58,510	67,257	10,887	2,790	15,151	4,486	103,269	87,601
	10%	12%	22%	20%	28%	0%	32%	21%	36%	21%	31%	17%	30%	19%
Shared-Ride 3+	3,097	4,290	1,396	3,818	7,803	-	53,485	68,707	8,205	2,990	15,475	3,111	89,461	82,916
	10%	5%	14%	21%	16%	0%	30%	21%	27%	23%	32%	12%	26%	18%
Walk	6,478	9,871	1,283	2,907	7,338	-	23,432	64,030	2,521	958	6,158	3,264	47,209	81,031
	21%	12%	13%	16%	15%	0%	13%	20%	8%	7%	13%	12%	14%	17%
Walk-Bus	672	5,982	366	1,224	201	-	976	12,453	255	922	616	438	3,086	21,018
	2%	7%	4%	7%	0%	0%	1%	4%	1%	7%	1%	2%	1%	5%
Walk-Rail	3,443	6,786	1,708	2,828	1,745	-	8,858	14,614	1,810	345	3,134	1,725	20,698	26,298
	11%	8%	17%	16%	4%	0%	5%	5%	6%	3%	7%	6%	6%	6%
Other	-	642	-	-	-	-	-	854	-	-	-	-	-	1,496
	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Total	31,302	82,678	9,909	17,798	48,364	767	180,411	323,334	30,178	13,286	48,175	26,797	348,338	464,660
	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

FIGURE 18 TRIPS BY TOUR PURPOSE AND TRIP MODE



Time of day

The share of tours by 30-minute time bin entering the US (arrivals) and departing the US (departures) are shown in Figure 19. Time periods are defined with period 1 starting at 3:00 AM (3:00 to 3:30 AM) and period 48 starting at 2:30 AM (2:30 to 3:00 AM).

FIGURE 19 SHARE OF TOUR ARRIVALS AND TOUR DEPARTURES BY 30-MINUTE TIME BIN

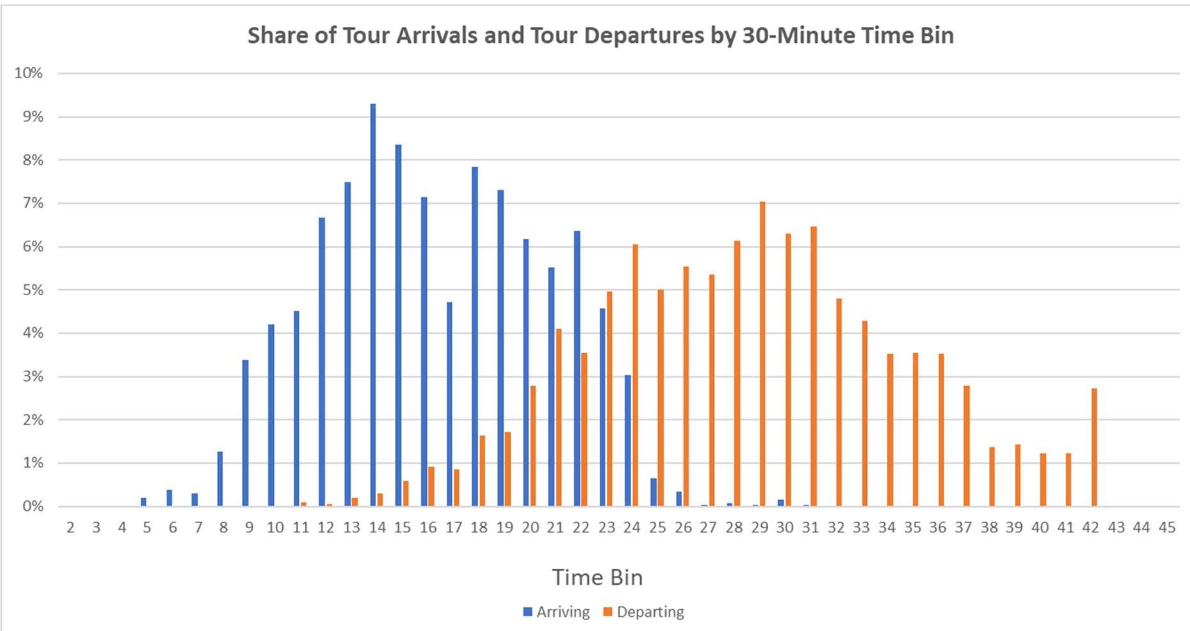
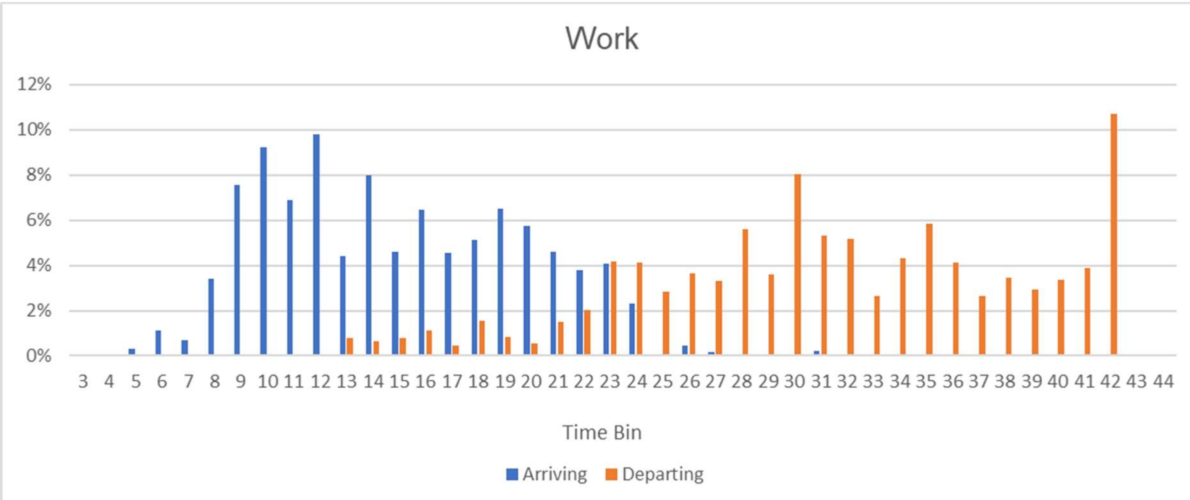
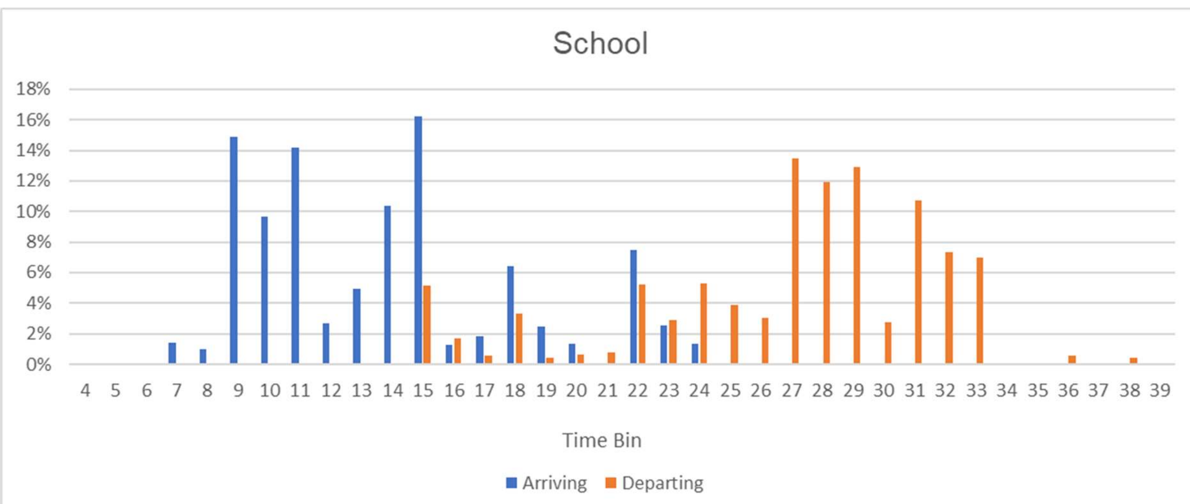
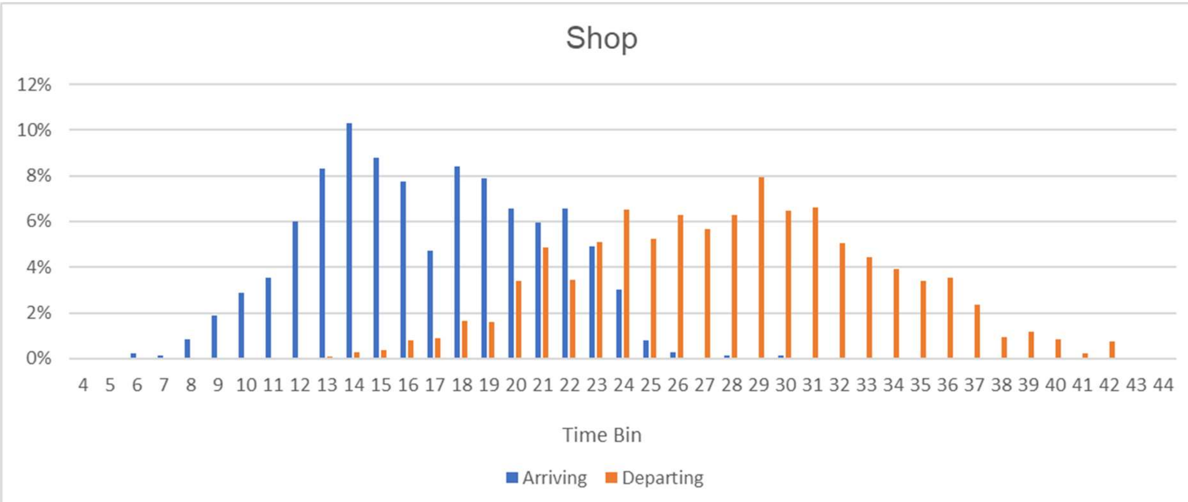


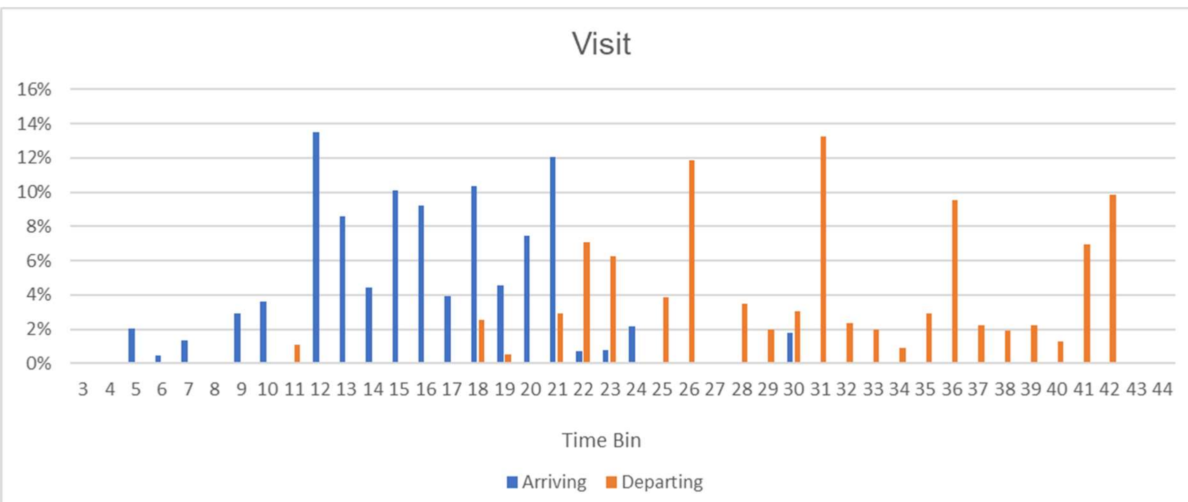
Figure 20 below breaks down the number of tour arrivals and departures further by tour purpose. Note that the percentages for each chart (i.e. tour mode) in Figure 20 add up to the total tours for each respective tour mode and not total tours across all tour modes.

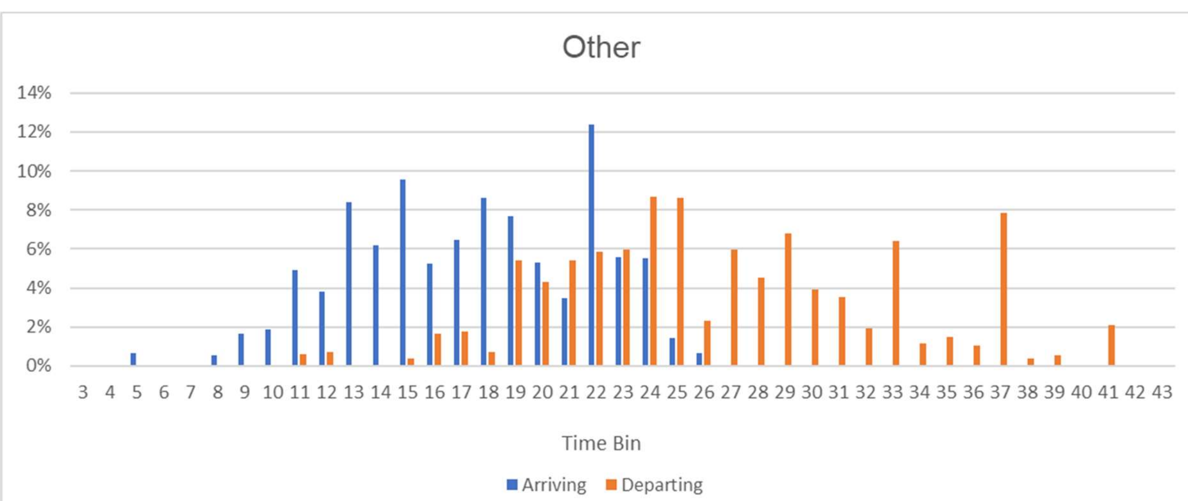
FIGURE 20 SHARE OF TOUR ARRIVALS AND TOUR DEPARTURES BY 30-MINUTE TIME BIN AND TOUR PURPOSE





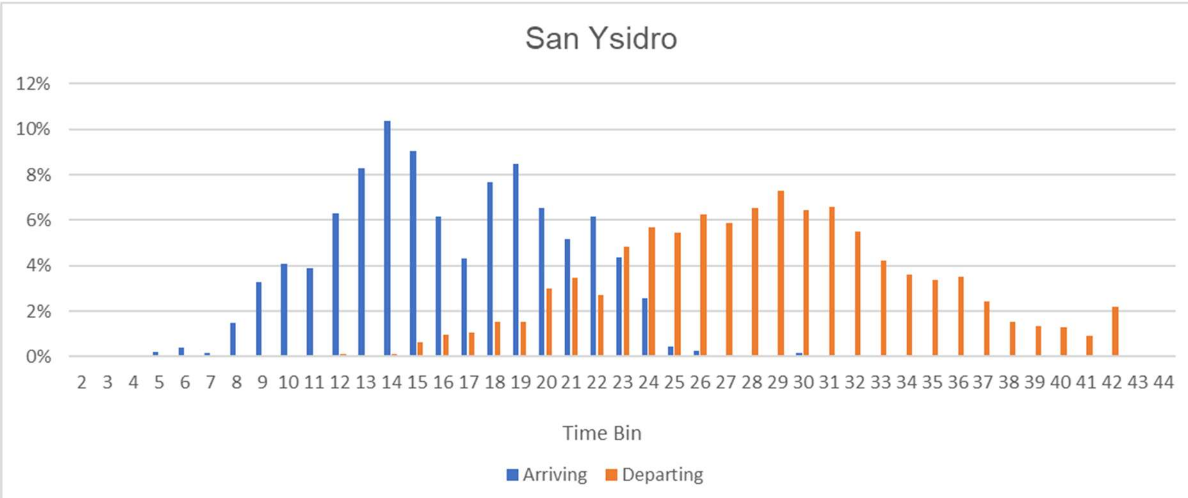

Shop
12%
10%
8%
6%
4%
2%
0%
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
Time Bin
Arriving
Departing

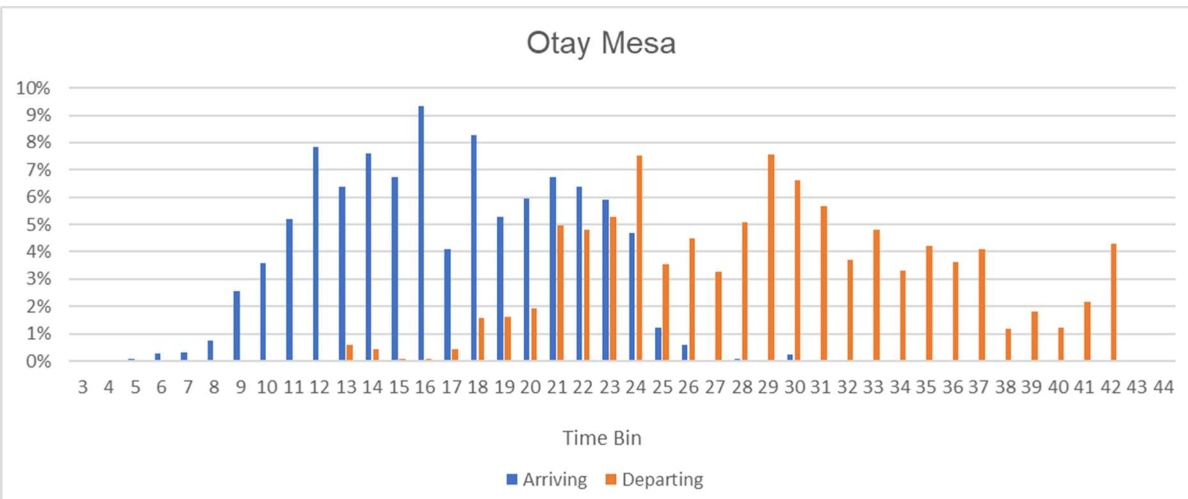

Visit
16%
14%
12%
10%
8%
6%
4%
2%
0%
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
Time Bin
Arriving
Departing

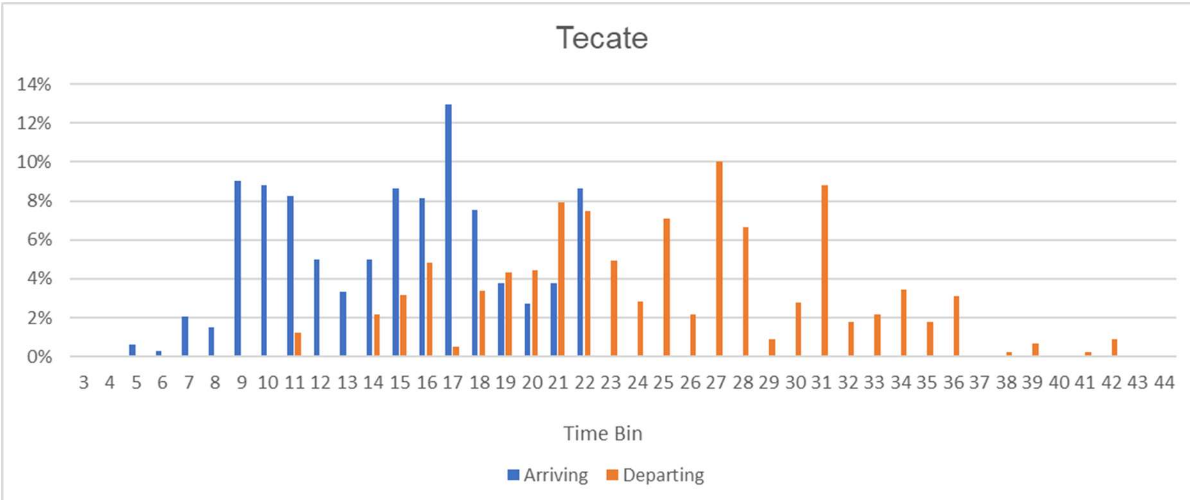

Other
14%
12%
10%
8%
6%
4%
2%
0%
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
Time Bin
Arriving
Departing

Finally, Figure 21 below breaks down the number of tour arrivals and departure by POE.

FIGURE 21 SHARE OF TOUR ARRIVALS AND TOUR DEPARTURES BY POE






Tecate
14%
12%
10%
8%
6%
4%
2%
0%
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
Time Bin
Arriving
Departing

3.0 MODEL DESIGN

3.1 MODEL DIMENSIONS

Purpose Definitions

There are five activity purposes in the cross-border travel demand model.

- Work: Any activity involving work for pay.
- School: Pre-k school, K-12, college/university, or trade school.
- Shop: Shopping at retail, wholesale, etc.
- Visit: Visiting friends or family
- Other: A broad category including eating out, medical appointments, recreational activities, etc.

Note that home activities are not listed, since we do not model activities south of the border.

Mode Definitions

The model has the following mode types at the trip level:

- Drive-alone: Single occupant private vehicle
- Shared 2: A private vehicle with exactly two passengers
- Shared 3+: A private vehicle with three or more passengers
- Walk: Walk mode
- Bike: Bike mode
- Walk-transit: Walk access to transit. There are three sub-types of transit: Local only, premium only, local + premium (which includes both local and premium services in the transit path)
- Taxi: Door-to-door taxi trip
- Single-pay TNC: Door-to-door TNC trip with a single payer (e.g. UberX)
- Shared-pay TNC: Stop-to-stop TNC trip with potentially multiple payers (e.g. UberPool)

We also model tour mode, which is the mode used to cross the border. These modes include drive-alone, shared 2, shared 3+ and walk. We assume that anyone crossing by bus or taxi is similar to walk, since they do not have access to a personal vehicle for the rest of their travel in San Diego County.

We also classify border crossings by lane type: general purpose, SENTRI, and Ready. We assume that the use of these lanes is related to the border crossing party; we attribute each party with SENTRI or Ready availability. The proportion of total border crossing parties with access to SENTRI and Ready lanes are based on observed survey data, pooled across all stations. This data is used to simulate the availability of the lane to the travel party. Each lane crossing type is related to the wait time that the travel party experiences at each border crossing station by mode, according to the wait time calculations described in Section 0.

Treatment of Time

Every trip is allocated to one of 48 time bins. Each of the 48 time bins represents a half-hour time slot in the day, where the day is assumed to begin at 3:00 AM. The time bins are broken up as indicated below in Table 10.

TABLE 10 TRAVEL MODEL TIME PERIODS

TIME PERIOD(S)	DESCRIPTION
1	Between 3:00 AM and 3:30 AM
2 through 47	Every half hour time slot of the day
48	Between 2:30 AM and 3:00 AM

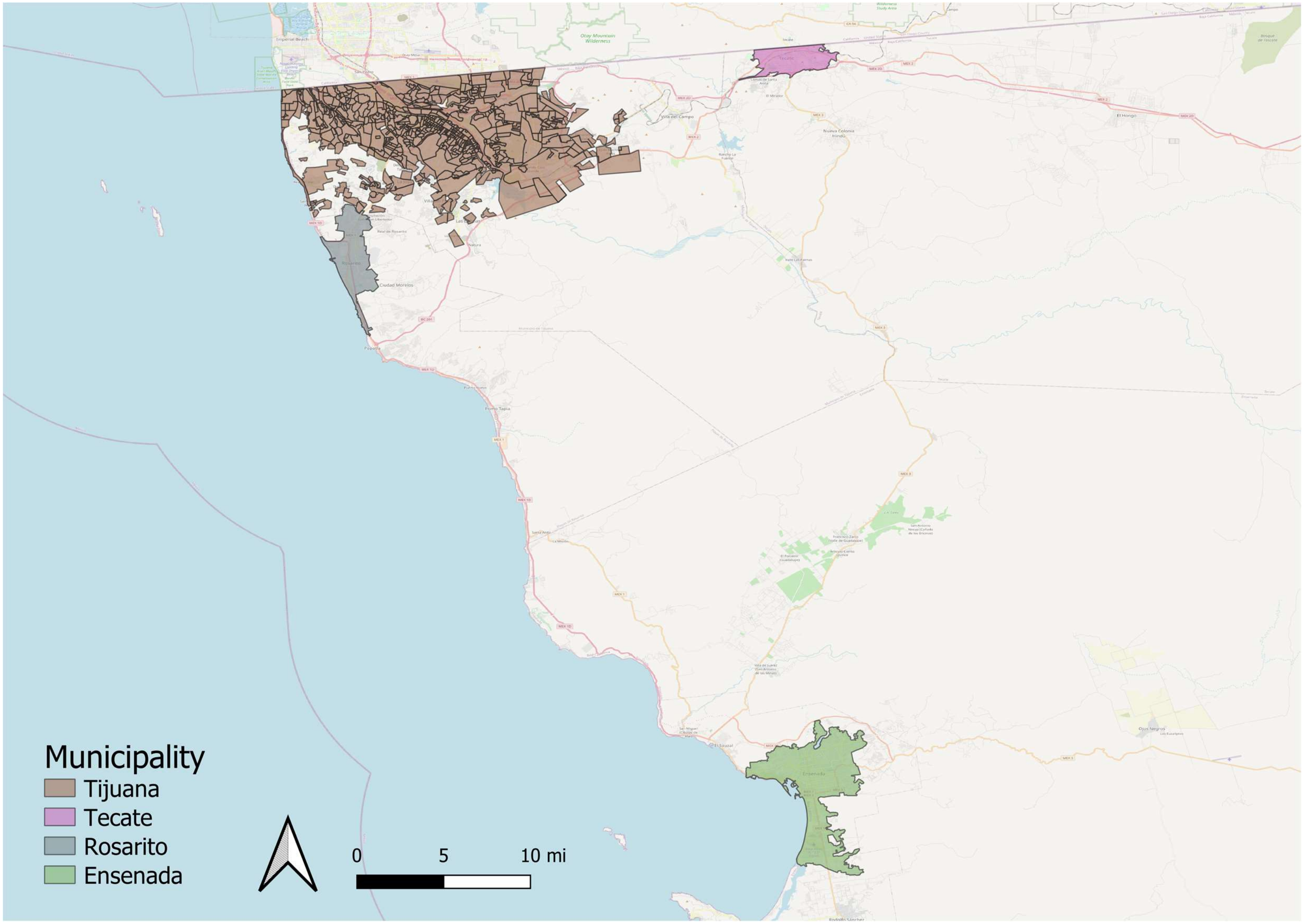
It should be noted that the SANDAG resident travel model currently allocates trips to 1 of 40 time bins, where all time bins, except the first and last, represent half hour time slots in the day. The first and last aggregate time between 12:00 AM and 5:00 AM. The ABM3 SANDAG resident travel model treatment of time is in a consistent 48-time bin structure.

Travel skims are also consistent with resident travel demand models. There are a total of five skims corresponding to the following time periods: Early AM, AM Peak, Midday, PM Peak and Evening.

Spatial distribution of travel

Every trip end in San Diego County is allocated to an MGRA. .

FIGURE 22: POPULATION SOUTH OF THE US BORDER



3.2 MODEL DESCRIPTION

Below is a general description of the model structure, followed by a more complete description for each model component shown in Figure 23. The more detailed description includes the number of model segments, the number of alternatives, the decision-making unit (where the choice is attributed in the model), and the model form.

1. Tour Enumeration: A list of person-tours is created by first cross-multiplying the input total person tours with the share of tours by pass type, then expanding tours by pass type to tours by pass type and purpose.
2. Tour Level Models

    2.1 Time-of-day Choice: Each person-tour is assigned an outbound and return half-hour period.

    2.2 Primary Destination and Station Choice: Each border crossing person-tour chooses a primary destination MGRA and border crossing station.

    2.3 Border Crossing Mode Choice: Each person-tour chooses a border crossing tour mode.

3. Wait Time Model

    3.1. Wait time model: Calculate wait time based on demand at each POE from model 2.2

    3.2. Convergence check: If max iterations reached (currently 3), goto Stop and Trip level models, else goto Model 2.2

4. Stop and Trip Level Models

    4.1 Stop Frequency Choice: Each person-tour is assigned number of stops by half-tour (outbound, return).

    4.2 Stop Purpose Choice: Each stop is assigned a stop purpose (consistent with the tour purposes).

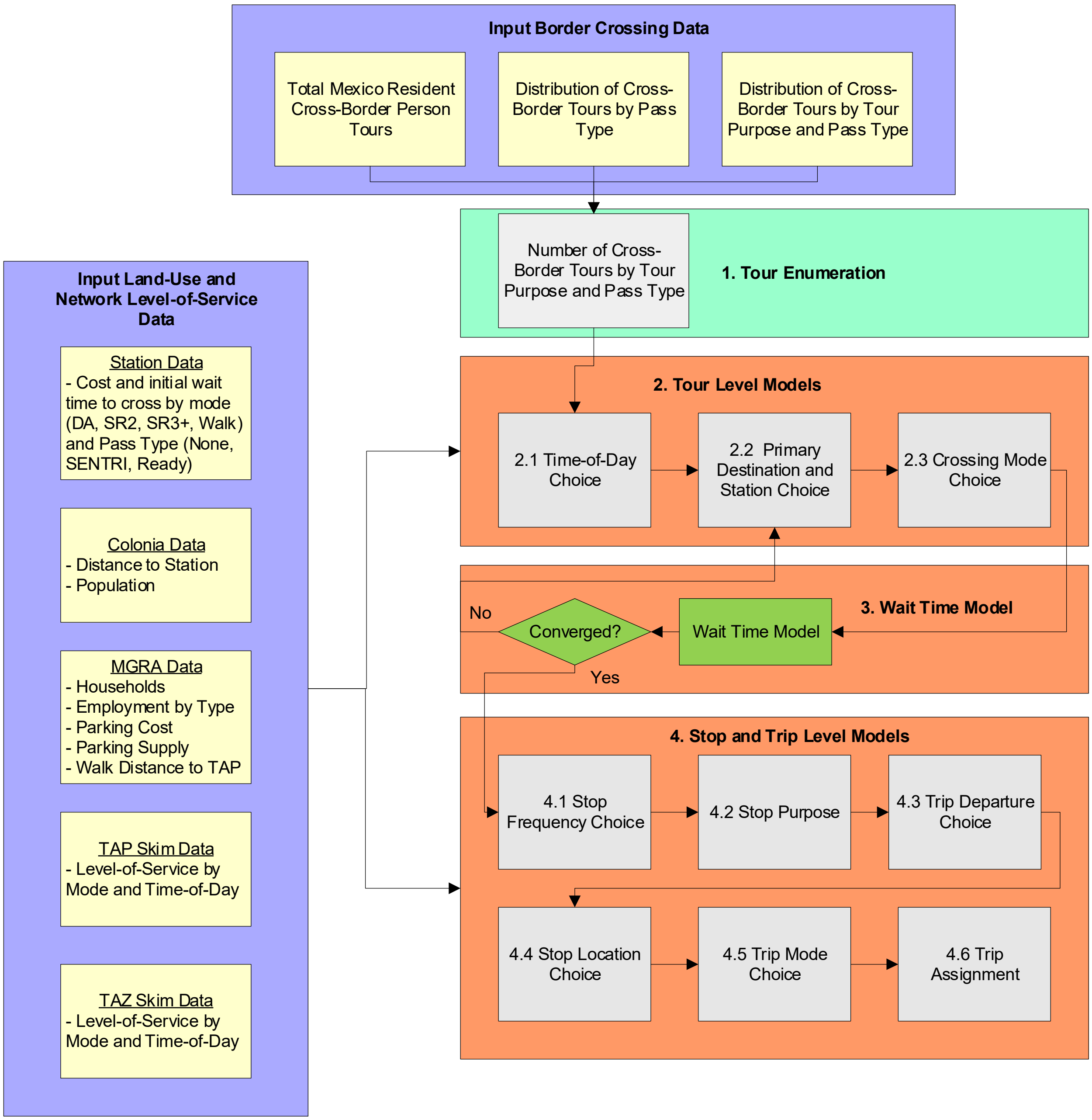
    4.3 Trip Departure Choice: Each trip is assigned a half-hourly time period.

    4.4 Stop Location Choice: Each stop chooses an MGRA location.

    4.5 Trip Mode Choice: Each trip is assigned a trip mode.

4.6 Trip Assignments: Trips are assigned to networks, along with resident and other special market trip tables, and skims are created for the next iteration of the model.

FIGURE 23 CROSS-BORDER TRAVEL MODEL



Tour Level Models

Tour Time-of-Day Choice

Number of Models: 5 (Work, School, Shop, Visit, Other)

Decision-Making Unit: Person-tour

Model Form: Lookup from probability distributions

Alternatives: 1176 (periods^2 + periods)/2)

A time-of-day choice model is applied to each tour with a known primary destination and border crossing station. The model is a lookup from observed distributions by tour purpose and has a temporal resolution of one-half hour that is expressed in 1176 half-hour entry/return time alternatives.

The tour scheduling model is placed after destination choice and before mode choice in the border crossing model, similar to the placement in the CT-RAMP resident model structure.

Primary Destination and Station Crossing Choice

Number of Models: 5 (Work, School, Shop, Visit, Other)

Decision-Making Unit: Person-tour

Model Form: Multinomial Logit

Alternatives: MGRAs and border crossing stations

The primary destination and border crossing choice model is a joint choice model of tour primary destination in the US and border crossing station. Due to the number of alternatives in the model, sampling is used to select a sub-set of primary destination MGRAs and border crossing pairs. The sampling procedure is based upon a simplified destination choice model that considers:

- The weighted distance from the TAZ to all border crossing stations S
- The time and cost that it takes to cross each station S
- The accessibility of the border crossing station to persons in Mexico

There are three main utility components of the sample of alternatives model, as follows:

$$U_{o,s} = \alpha * Dist_{o,s} + \ln{[Size_o]}$$

$$U_{d,s} = \beta * Dist_{s,d} + \ln{[Size_d]}$$

$$U_{m,s} = \gamma * Time_{m,s} + \delta * Cost_{m,s}$$

Where:

o = Tour origin in Mexico

O = Total number of tour origins

d = Tour primary destination in the US

D = Total number of tour primary destinations

s = Border crossing station

S = Total number of stations

m = tour mode (used to cross border)

M = Total number of tour modes

$Size_o$ = Size of origin zone o (population)

$Size_d$ = Size of primary destination zone d (*f*(population, employment))

$Dist_{o,s}$ = Distance from tour origin zone o to station s

$Dist_{s,d}$ = Distance from station s to tour primary destination zone d

$U_{o,s}$ = Utility of origin o from station s

$U_{d,s}$ = Utility of primary destination d from station s

$U_{m,s}$ = Utility of crossing station s by mode m

$Time_{m,s}$ = Time required to cross station s by mode m

$Cost_{m,s}$ = Cost required to cross station s by mode m

α, β, γ, δ = Coefficients to be estimated

The station choice utility [U_s] portion of the sample of alternatives model depends on the accessibility of each station to all potential origin zones in Mexico (e.g. population in Tijuana weighted by distance), and the time and cost required to cross each station, as follows:

$$U_s = \ln\left[\sum_{o=1}^{O} e^{U_{o,s}}\right] + \ln\left[\sum_{m=1}^{M} e^{U_{m,s}}\right]$$

Note that these utilities can be pre-calculated and stored by tour purpose and station.

The utility of choosing a primary destination and station pair can be calculated by adding the utility of the station to the utility of the station to destination:

$$U_{sd} = U_s + U_{d,s}$$

The simplified destination and station choice model is used to select a subset of 50 destination/station pairs. To speed calculations, the TAZ is sampled first since distance is the only measure of impedance used to represent accessibility of primary destination to station, and distance is represented at the TAZ level. Zone size in this case is equal to the sum of the sizes of the MGRAs within the TAZ, by tour purpose. Once the TAZ is sampled, an MGRA within the TAZ can be chosen based on the pre-calculated probability of the MGRA within the TAZ, which is based on the MGRA proportion of the TAZ size. Note that the total number of alternatives in the model is TAZs * stations (currently 13,800).

Once the sample of destination\station-pairs is chosen, a trip mode choice logsum from Model 4.5 is computed for each sampled destination\station-pair and tour mode. The tour mode choice logsum is then estimated and replaces the distance term in the station\destination utility ($U_{s,d}$) and a choice is made of actual destination and station from the sampled alternatives. The times and costs in the trip mode choice logsums is based upon representative departure and arrival periods for each purpose. The station portion of the utility (U_s) remains unchanged from the sampling model.

Border Crossing Mode Choice

Number of Models: 5 (Work, School, Shop, Visit, Other)

Decision-Making Unit: Person-tour

Model Form: Multinomial Logit

Alternatives: Tour Mode

This model chooses tour mode based on a known tour destination, border crossing station, and entry/return time-of-day. The mode choice is based upon a simplified trip mode choice logsum representing the accessibility of relevant modes for each border crossing mode as well as the time and cost of crossing the station by border crossing mode. The alternatives in the model are Drive-alone, shared-2, shared 3+, and walk. These alternatives compete equally with each other (multinomial logit). The utility of each alternative is based on:

- The trip mode choice logsum calculated for each border crossing mode between the POE and the primary destination for the scheduled departure time and arrival time at the POE
- The wait time by mode from the wait time model, which for auto modes is a function of pass holding (none, SENTRI, or Ready)
- An optional auto toll specified for each POE

- A set of alternative-specific constants

Stop Level Models

Stop Frequency Choice

Number of Models: 5 (Work, School, Shop, Visit, Other)

Decision-Making Unit: Person-tour

Model Form: Lookup table

Alternatives: 16 (0, 1, 2, 3 stops per half-tour)

The stop frequency choice mode is a lookup table of probabilities based upon tour purpose and duration. Probabilities are calculated from observed data.

Stop Purpose Choice

Number of Models: 5 (Work, School, Shop, Visit, Other)

Decision-Making Unit: Stop

Model Form: Lookup table

Alternatives: 5 (Work, School, Shop, Visit, Other)

The stop purpose choice model is a lookup table of probabilities based upon tour purpose and number of stops on tour. Probabilities are calculated from observed data. The purpose-segmentation is based on the tour purpose, if implemented.

Trip Departure Choice

Number of Models: 1

Decision-Making Unit: Trips other than first trip and last trip on tour

Model Form: Lookup from Probabilities

Alternatives: Number of half-hour time after outbound period for outbound trips/number of half-hour time periods before return period for return trips

Each trip is assigned to a trip departure time period. The first and last trips of the tour are set to the entry/return time periods from tour time of day choice model, respectively. Each intermediate trip departure time is calculated from a lookup table of probabilities that consider the number of remaining half-hour periods in the tour from the last scheduled trip and whether the stop is made on the outbound or return direction.

Stop Location Choice

Number of Models: 1 (with stop-purpose-specific size terms)

Decision-Making Unit: Stop

Model Form: Multinomial Logit

Alternatives: MGRAs

The stop location choice model predicts the location of stops along the tour other than the primary destination. The stop-location model is structured as a multinomial logit model using MGRA attraction size variable and route deviation measure as impedance. The alternatives are sampled from the full set of MGRAs, based upon the out-of-direction distance to the stop and the size of the MGRA. The sampling mechanism is also subject to certain rules based on tour mode. All destinations are available for auto tour modes, so long as there is a positive size term for the MGRA. Intermediate stops on walk tours must be within 3 miles of both the tour origin and primary destination MGRAs. The sampling for intermediate stops on walk-transit tours is based upon the MGRAs that are within walking distance of the boarding or alighting stops at the tour origin and primary destination.

The intermediate stop location choice model works by cycling through stops on tours. The level-of-service variables (including mode choice logsums) are calculated as the additional utility between the last location and the next known location on the tour. For example, the LOS variable for the first stop on the outbound direction of the tour is based on additional impedance between the tour origin and the tour primary destination. The LOS variable for the next outbound stop is based on the additional impedance between the previous stop and the tour primary destination. Stops on return tour legs work similarly, except that the location of the first stop is a function of the additional impedance between the tour primary destination and the tour origin. The next stop location is based on the additional impedance between the first stop on the return leg and the tour origin, and so on.

Trip Mode Choice

Number of Models: 1

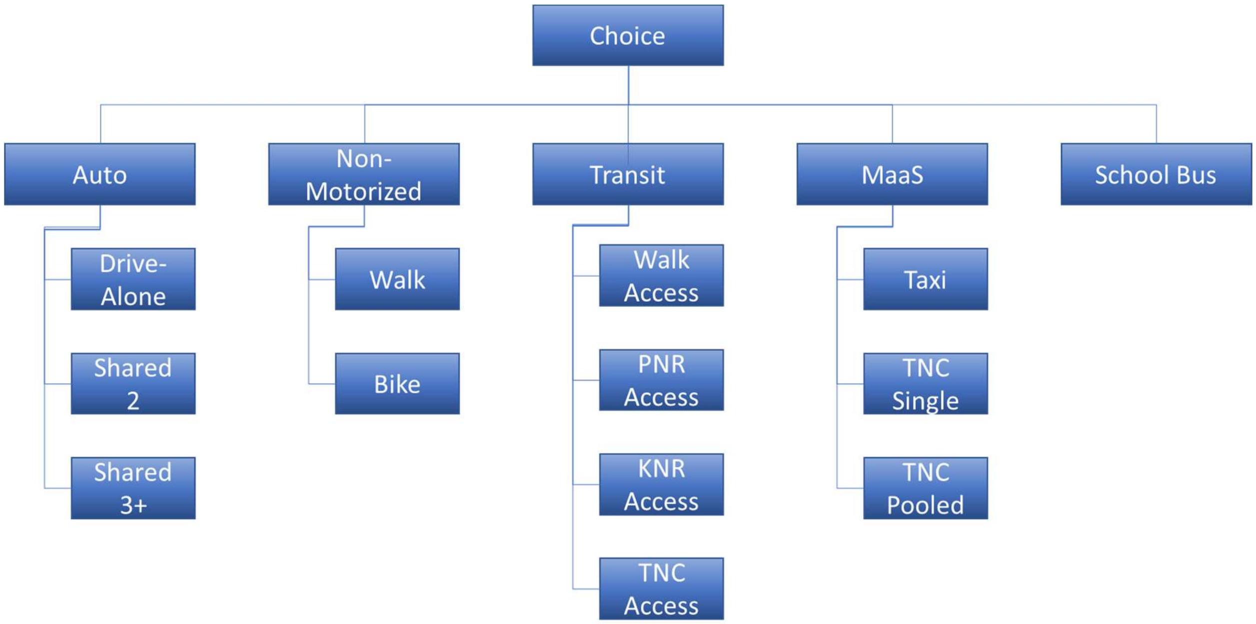
Decision-Making Unit: Trip

Model Form: Nested Logit

Alternatives: 13 Trip Modes

A trip mode is chosen for each trip on the tour. Trip modes are consistent with the resident travel model, as shown Figure 24 below, though certain modes (bike, PNR-transit, KNR-transit, TNC-transit, and school bus) are unavailable for cross-border tours. The utility of each mode is a function of the time and cost of the mode for the period that the trip occurs in and is influenced by the mode used to cross the border.

FIGURE 24 CROSS-BORDER TOUR TRIP MODE CHOICE MODEL FOR TRAVEL IN US



Trip Assignment

Number of Models: 2

Decision-Making Unit: Trip

Model Form: SOLA for Auto and Emme Strategic Pathbuilder for Transit

Alternatives: NA

Cross-border trips are added to US resident trips, commercial vehicle trips, other internal-external (IE) and external-external (EE) trips, and special market trips. The total trips are then assigned to networks. Travel time and cost skims are created from the assignment and fed to the next iteration of the models in a feedback loop until the travel times input to the model are consistent with the times that are output by the model.

3.3 MODEL INPUTS

The model system requires the following exogenously specified inputs (note that three additional data sets - Border Crossing Data, Station Data, and Mexico land-use data) are required in addition to the data currently input to the resident activity-based models):

- *Border Crossing Data*: The total number of border crossings to be modeled in the Cross-Border Model, and the distribution of border crossings by tour purpose and household income. Exogenous specification of these model inputs provides sensitivity of the model system to assumed changes in the purpose and income of border crossings, but recognizes limitations of both the travel survey data, and the model itself.
  - *Constraints on Border Crossing Data and Modeling*: The travel survey is a cross-sectional dataset collected in 2019; therefore, it cannot be used to develop a macro-economic model that measures economic and other influences on the number of residents of Mexico who work in the United States, or the propensity of residents of Mexico to purchase goods and services in the United States.
  - *Opportunities for Border Crossing Data and Modeling*: However, the analyst can choose to change the input total number of border crossings, and/or the distribution of border crossings by purpose and income, and observe the change in number of crossings by station, and travel demand within San Diego County. The total base-year border crossings and distributions of person-crossings by tour purpose and income are derived from the border crossing survey data. For future-year model runs, these inputs can be held constant or varied according to other data sources.
- *Station data*: Time and cost of crossing each station by time-of-day, mode (Drive-alone, Shared-ride 2, Shared-ride 3+, Walk), and Standard\SENTRI\Ready (for auto modes). This provides model sensitivity to differences in time/cost of crossing at each station, and also provides the ability to test potential future scenarios where HOVs are provided for faster crossing times.
- *Mexico (AGEB and Localidad) Land-Use data*: The population in each geographic unit south of the US border and the distance from the Colonia to each station. The number of border crossings originating in each area is assumed to be linearly correlated to its population and inversely proportional to its distance to the US\Mexico border. Population estimates are collected by the Instituto Nacional de Estadistica y Geographifia (INEGI) at the level of a basic geostatistical area (Area Geostadistica Basica, or AGEB, roughly equivalent to U.S. Census Tracts). AGEBs and Colonia largely overlap within Tijuana city boundaries (though there is no coherent spatial nesting scheme), and AGEB population estimates are redistributed to Colonia based on a proportional area operation. Outside of Tijuana, the origins are distributed to a Localidad, or locality. These units are similar to the Census Designated Place in the US. Areas outside of Tijuana include Rosarito, Tecate and Ensenada, as shown in Figure 22
- *MGRA data*: The population and employment (by type) in each MGRA, parking cost and supply, etc. This data provides sensitivity to land-use forecasts in San Diego County. These are the same data sets as are used in the resident activity-based model.

- *TAP skim data*: Transit network level-of-service between each transit access point (transit stop). This provides sensitivity to transit network supply and cost. These are the same data sets as are used in the resident activity-based model.
- *TAZ skim data*: Auto network level-of-services between each transportation analysis zone. This provides sensitivity to auto network supply and cost. These are the same data sets as are used in the resident activity-based model.

4.0 MODEL ESTIMATION

The data described in the previous section is used to re-estimate three components of the cross-border travel model: the tour origin-destination and mode choice model, the intermediate stop frequency model, and the stop destination choice model. This chapter describes the estimation process and results.

4.1 ESTIMATION IN ACTIVITYSIM

ActivitySim is a disaggregate activity-based travel model in which a synthetic population is run through each model component. The software builds a choice model that is specific to each household and person, taking into account the attributes of the synthetic population, the choice outcomes of previous models in the model system, and logsums from downstream model components. As each model is run, the choice for that model is recorded for the decision-maker before moving on to the next model.

The ActivitySim software was recently enhanced with an 'estimation mode'. This feature makes it possible to run a survey population through the software with the same attributes as the synthetic population, in which the observed choices for each decision-maker override the simulated choices from the model. Because the models are constructed for each decision-maker according to their attributes and the observed choices for the decision-maker, the explanatory variables for each model (including logsums from downstream models) can be saved to disk and used to re-estimate the model.

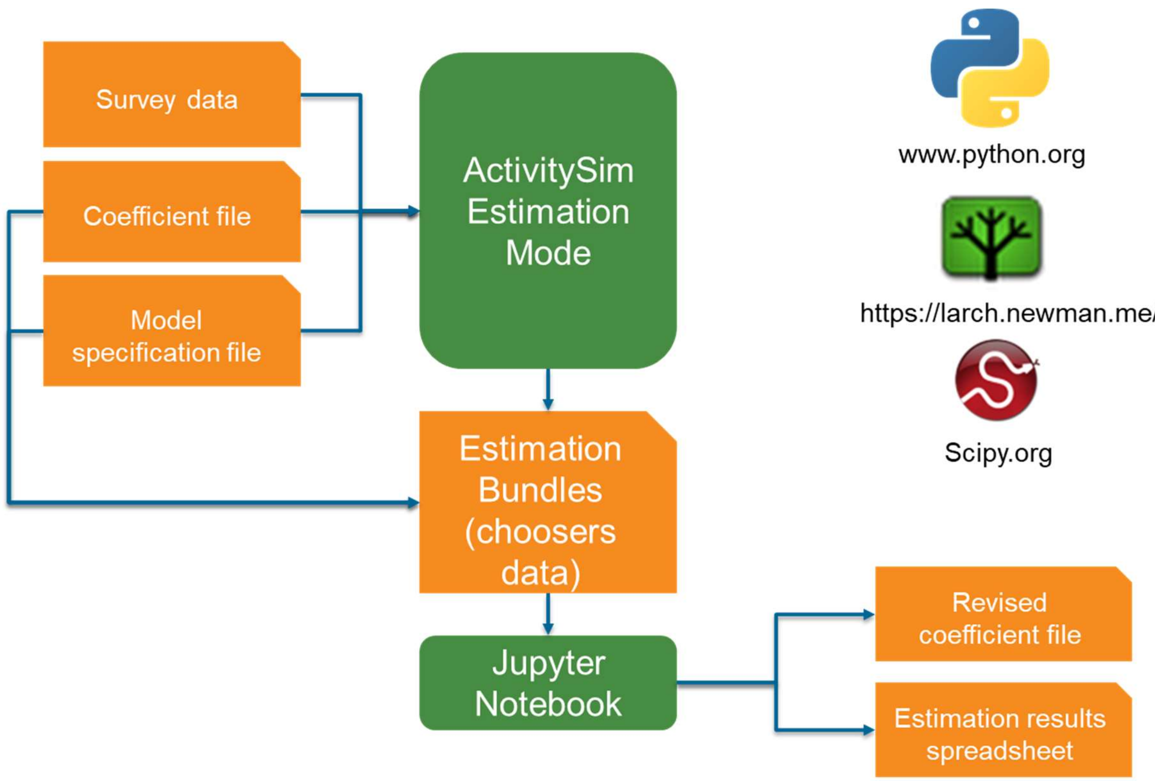
These output files are referred to as 'estimation data bundles'. Each estimation data bundle (EDB) consists of a table of data where rows are decision-makers (households, persons, tours, trips, etc.) and columns are data for each alternative to be used in utility equations. This data along with the ActivitySim input coefficient file(s) and model specification file is read by a Jupyter Notebook[7] that re-estimates the model specification in Larch[8]. Larch is a logit model estimation package in Python that is built on top of the Python Scipy[9] package. The EDBs can also be easily post-processed to data formats required by other logit model estimation packages such as ALOGIT[10].

The model estimation process in ActivitySim is shown in Figure 25. Survey data is input to ActivitySim, as a replacement for the synthetic population, and the outputs from any upstream model components run before the component to be estimated. ActivitySim also requires an input coefficient file for each model, as well as an input model specification. When ActivitySim is run in estimation mode, it outputs an estimation data bundle for each model component. The estimation

[7] https://jupyter.org/
[8] https://larch.newman.me/index.html
[9] https://www.scipy.org/
[10] ALOGIT Model Estimation Package: http://www.alogit.com/index.htm

data bundle is read by the Jupyter notebook, along with the input model specification and coefficient file. The estimation process run in Jupyter writes out a new coefficient file in ActivitySim format, with estimated coefficients, as well as an Excel spreadsheet that describes estimation results. When ALOGIT is used for model estimation, the ActivitySim configuration files need to be updated manually as per the updated model specification and coefficients.

FIGURE 25: ACTIVITYSIM ESTIMATION PROCESS



4.2 DISTRIBUTIONS

The model uses several distributions derived from the cross border survey data. These include stop frequency, stop purpose, inbound and outbound stop duration, tour purpose by pass type, and lane shares by purpose. These distributions were prepared in the CT_RAMP format previously used in the model as the ActivitySim preprocessor was already developed to convert the CT-RAMP inputs to the required format. Distributions are based on expanded weights of survey tours.

The stop purpose distributions are segmented by tour purpose, inbound/outbound direction, and stop number (1, 2, or 3). Stop purpose distributions are shown in the appendix, Table 29. Stop purposes on tours with greater than 3 stops in either direction are not included in the distributions.

Outbound stop duration is segmented by tour purpose, and the upper and lower bounds of remaining half hour periods after the last scheduled trip. The stop duration is then segmented by likelihood the stop departure is in the last inbound/outbound trip period plus 1 to 11 periods as shown in the appendix, Table 30. Tours with greater than 3 stops are included in the distributions up to the probabilities of the third stop.

Inbound stop duration is segmented by tour purpose, and the upper and lower bounds of remaining half hour periods after the last scheduled trip. The stop duration is then segmented by likelihood the stop departure is in the tour arrival period minus 1 to 7 periods as shown in the appendix, Table 31. Tours with greater than 3 stops are included in the distributions up to the probabilities of the third stop.

Tour purpose distributions are derived for Sentri, Ready, and Non-pass holding tours. It is assumed that only tours using the Sentri/Ready lanes are pass holders. This distribution is shown in Table 11.

TABLE 11: TOUR PURPOSE DISTRIBUTIONS

TOUR_PURP_CBM_STR	SENTRI	READY	OTHER
Work	14%	16%	23%
School	3%	1%	5%
Shop	73%	70%	59%
Visit	2%	4%	4%
Other	8%	9%	8%

4.3 WAIT TIME MODEL

To correct for the negative correlation of the Ready lane wait times to volume/lane at San Ysidro, a constant value for wait time is used equal to the mean of the wait time data. The observed data by lane type is then fit to a linear model which assumes San Ysidro is the base case with a constant term for each additional port of entry and a volume term for additional ports of entry.

EQUATION 1

$$Wait\ Time_{Lane\ Type} = C_1 + C_2 OM(0,1) + C_3 TC(0,1) + Volume/lane * (C_4 + C_5 OM(0,1) + C_6 TC(0,1))$$

Where *OM(0,1)* and *TC(0,1)* are boolean terms to indicate the port being modeled. Since SENTRI and Ready lanes do not exist at Tecate, the C_3 and C_6 terms are not used in these fits. Table 12 shows the results of this regression fit for each lane type. A low P Value indicates a greater significance and terms with a P-value greater than 0.066 are not used in this fit. For example, since Ready lane wait times have been fixed to a constant value at San Ysidro, the volume coefficient, C_6, is not significant to the fit (P=1). Whereas in the standard lane, wait time is positively correlated with volume/lane as one would expect (higher volume leads to longer wait times) with negative POE-specific terms at Otay Mesa and Tecate indicating these ports have faster wait times with the same volumes compared to the reference case of San Ysidro.

TABLE 12: INITIAL MODEL COEFFICIENTS

LANE	R^2	TERM	COEF	P VALUE
Standard	0.358	Generic Const (C_1)	16.793	0.052
		OM-specific Const (C_2)	37.694	0
		TC-specific Const (C_3)	26.242	0.066
		Generic Volume (C_4)	2.138	0
		OM-specific Volume (C_5)	-1.329	0.112
		TC-specific Volume (C_6)	-1.672	0.001
Ready	0.259	Generic Const (C_1)	40.079	0
		OM-specific Const (C_2)	-38.516	0
		Generic Volume (C_4)	2.12E-12	1
		OM-specific Volume (C_5)	1.253	0
SENTRI	0.580	Generic Const (C_1)	3.632	0
		OM-specific Const (C_2)	-0.593	0.6
		Generic Volume (C_4)	0.217	0

		OM-specific Volume (*C5*)	-0.054	0.168
Pedestrian	0.534	Generic Const (*C1*)	-2.692	0.249
		OM-specific Const (*C2*)	2.0268	0.495
		TC-specific Const (*C3*)	-2.8802	0.531
		Generic Volume (*C4*)	0.2295	0
		OM-specific Volume (*C5*)	-0.1517	0
		TC-specific Volume (*C6*)	-0.1184	0.001

Time-of-day effects are tested to help compensate for the lack of granularity in the model. For example, where San Ysidro Ready Lane wait times are held constant with respect to volume, could they vary with respect to time? In order to preserve the explanatory power of both the volume and the time of day, the volume coefficients derived in Table 12 are held constant so the general, station specific, and time-of-day constants are estimated separately. The final model is shown in Equation 2 where the constants C_1, C_2, C_3, C_7, C_8, C_9, C_{10}, C_{11} are taken from Table 13, and C_4, C_5, C_6 are taken from Table 12**Error! Reference source not found.**. There are no observations of early AM crossings in the SENTRI lane at Otay Mesa, so there is no estimation of this constant.

Equation 2

$$Wait\ Time_{Lane\ Type} = C_1 + C_2 OM(0,1) + C_3 TC(0,1) + C_7 EA(0,1) + C_8 EV(0,1) + Volume/lane * (C_4 + C_5 OM(0,1) + C_6 TC(0,1) + C_9 OM(0,1)EA(0,1) + C_{10} OM(0,1)EV(0,1) + C_{11} TC(0,1)EV(0,1))$$

TABLE 13: FINAL MODEL CONSTANTS

LANE	R^2	TERM	COEF	P VALUE
Standard	0.696	Generic Const (*C1*)	24.4819	0
		OM-specific Const (*C2*)	38.2236	0
		TC-specific Const (*C3*)	-31.6399	0
		EA Generic Const (*C7*)	-17.017	0.012
		EA OM-specific Const (*C9*)	-29.1815	0.012

		EV Generic Const (C_8)	-33.0233	0
		EV TC-specific Const (C_{11})	21.8711	0.023
		EV OM-specific Const (C_{10})	-11.0587	0.292
Ready	0.643	Generic Const (C_1)	40.0794	0
		OM-specific Const (C_2)	-33.5957	0
		EA Generic Const (C_7)	-1.211E-11	1
		EA OM-specific Const (C_9)	-7.7458	0.265
		EV Generic Const (C_8)	1.584E-12	1
		EV OM-specific Const (C_{10})	-25.8116	0.001
SENTRI	0.020	Generic Const (C_1)	3.5404	0
		OM-specific Const (C_2)	-0.3852	0.469
		EA Generic Const (C_7)	0.1617	0.857
		EV Generic Const (C_8)	0.4187	0.667
		EV OM-specific Const (C_{10})	-1.1282	0.381
Pedestrian	0.0428	Generic Const (C_1)	-3.5615	0.013
		OM-specific Const (C_2)	3.3397	0.074
		TC-specific Const (C_3)	-2.0119	0.402
		EA Generic Const (C_7)	3.4151	0.277
		EA OM-specific Const (C_9)	-4.477	0.285
		EV Generic Const (C_8)	1.9005	0.577
		EV OM-specific Const (C_{10})	-3.8144	0.397

The generic parameters are used for any port of entry while the station specific parameters for Otay Mesa and Tecate are only used for those ports. If modeling a new port of entry, only the generic terms should be used.

Ultimately the project team decided to implement the model without time-of-day effects in order to minimize the number of constant terms in the model. Figure 26 through Figure 29 show the comparison of the observed wait times to the predicted wait times. These fits look reasonable given that the wait time data is often reported in discrete bins (such as 5, 10, or 15 mins in SENTRI lanes) and there is a lot more at play in the border wait times than processing volumes of vehicles (the primary objective is to make sure each border crossing is legal and not to guarantee a speed of crossing). The station specific terms allow the model to capture the behavior at stations with different attributes. The Standard lane at Tecate sees a much higher Volume/Lane due to the low number of lanes at this P.O.E, but the model is able to capture this effect with the station specific terms.

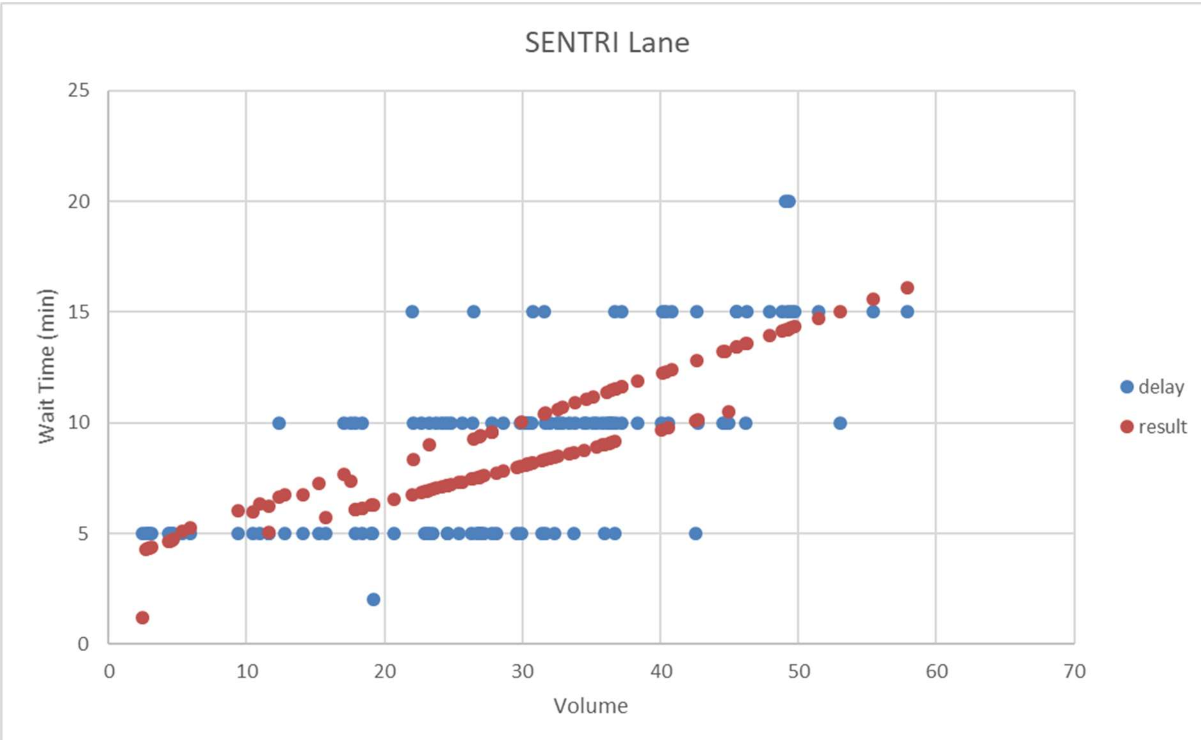


FIGURE 26: SENTRI LANE MODEL RESULTS

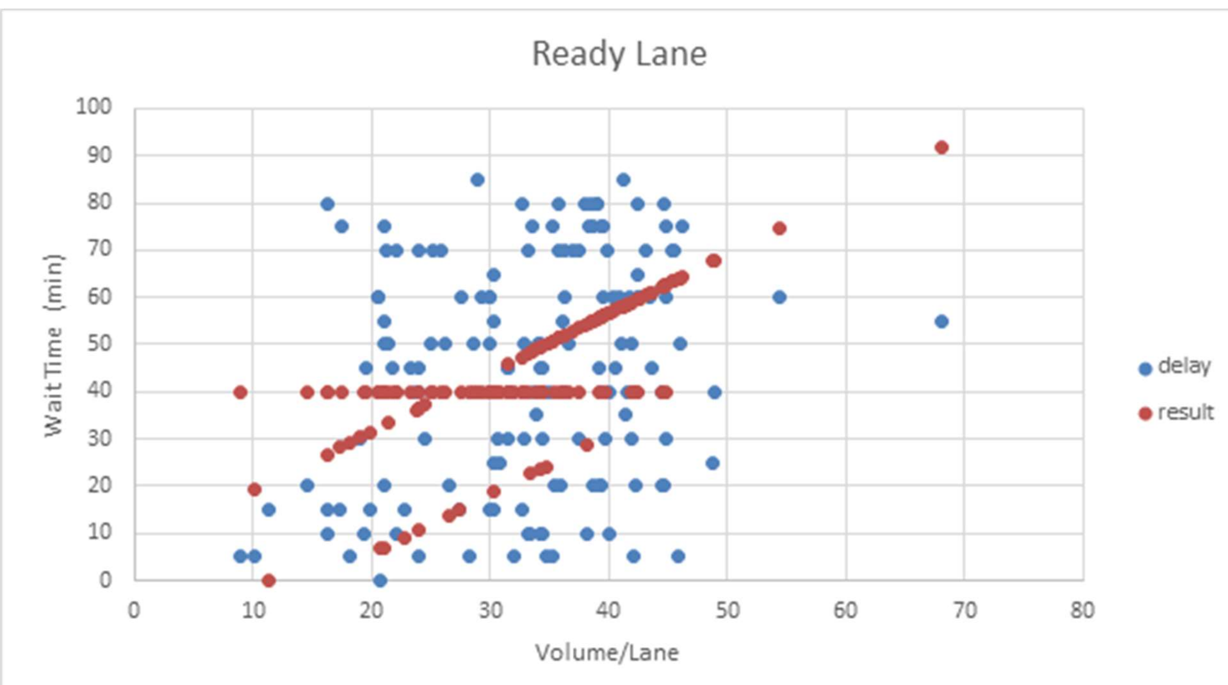


FIGURE 27: READY LANE MODEL RESULTS

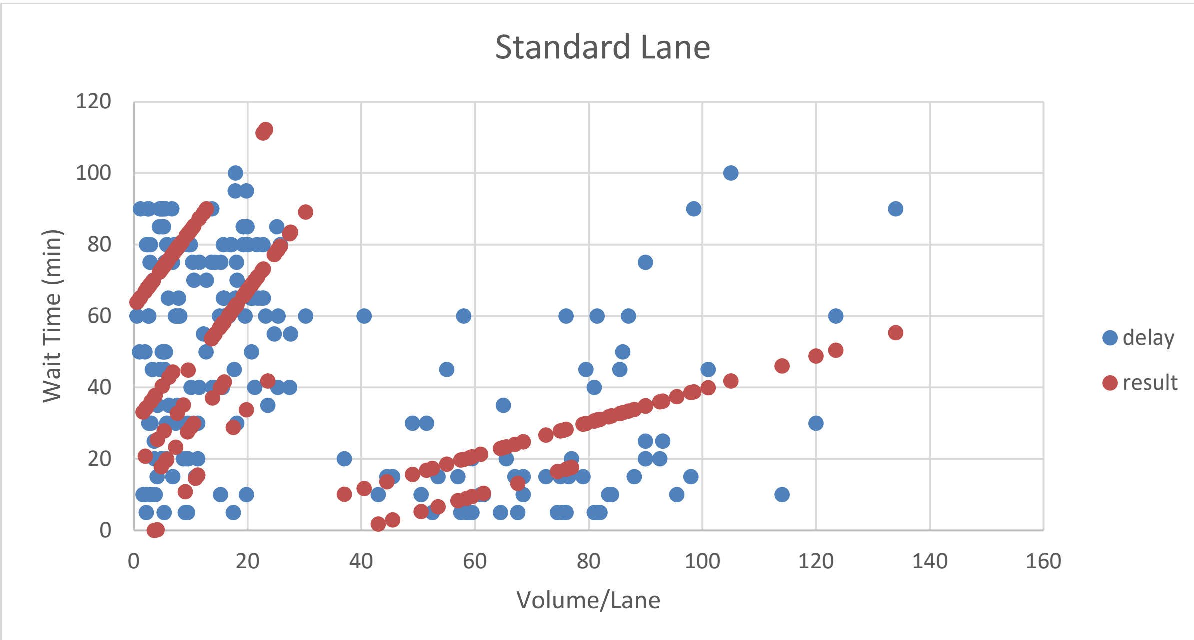


FIGURE 28: STANDARD LANE MODEL RESULTS

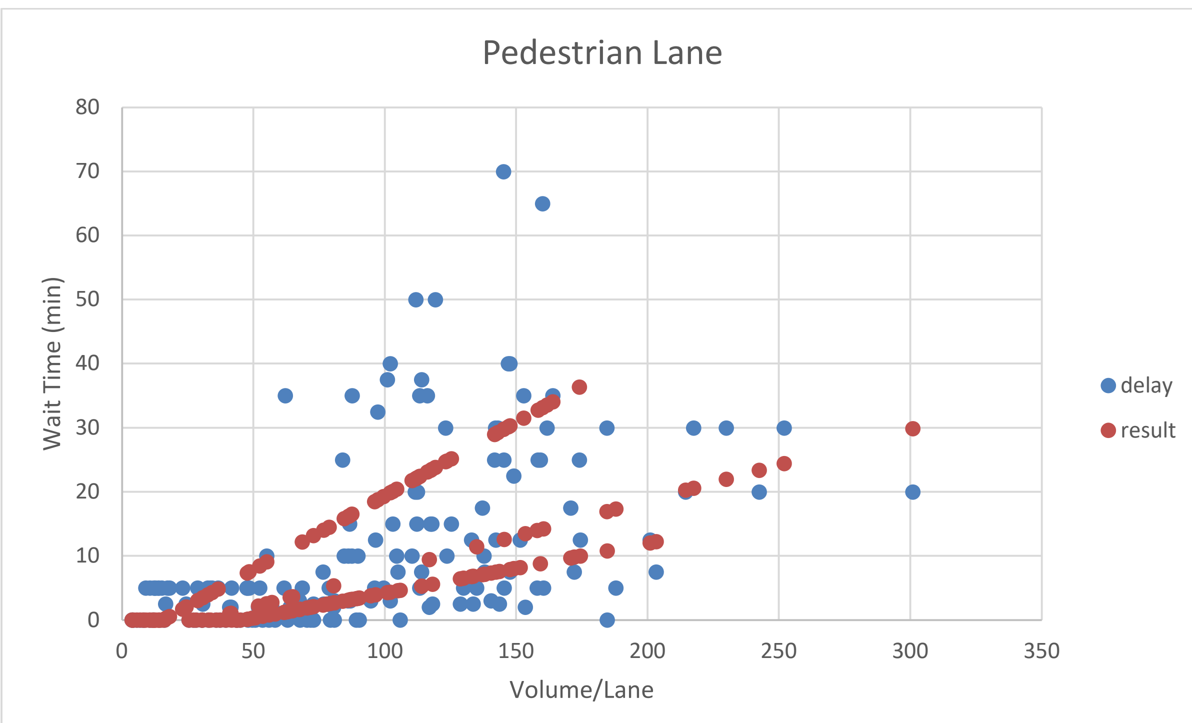


FIGURE 29: PEDESTRIAN LANE MODEL

4.4 TOUR POE, DESTINATION, AND BORDER CROSSING MODE CHOICE

As shown in Figure 23 there are three tour level models in the Cross-Border Travel Model; a joint model of primary destination and border crossing point of entry, followed by a tour time-of-day choice model, and finally a border crossing mode choice model. Initial estimation results of the border crossing mode choice model resulted in illogical coefficients on accessibility variables between the POE and the internal destination MGRA. For this reason, the project team estimated a set of joint models of border crossing POE, internal destination MGRA, and border crossing mode. As will be demonstrated below, estimating the model simultaneously resulted in more logical mode choice accessibilities. This approach to model estimation is similar to the method utilized in the initial model development effort over 10 years ago.

Because the models are implemented sequentially, the construction of the estimation file required some effort. We first modified the utility expression file for the border crossing station and internal destination choice model to write out mode choice logsums by mode (drive-alone, shared 2, shared 3+, and walk) for each sampled border crossing station and internal destination. Running this revised model specification through ActivitySim's estimation mode resulted in an estimation data bundle containing, for each observed cross-border tour, a sample of alternatives consisting of combinations of POEs and internal destinations. The sample includes one MGRA for each Transportation Analysis Zone (TAZ) (chosen at random from within the TAZ according to the ratio of the size term of the MGRA to the total size term across all MGRAs within the TAZ), and every TAZ for each POE. In other words, approximately 15,000 sampled destinations for each cross-border tour. For each of these destinations, the estimation data file includes:

- The mode choice logsum for each POE to MGRA pair for each mode (drive-alone, shared 2, shared 3+, and walk) in both directions (outbound and inbound). In this case the mode choice logsum represents the weighted utility of all modes (including transit) for the tour
- The distance between the POE and the MGRA
- The size of the MGRA (used in sampling but not used in estimation)

To this data is appended information on the observed cross border tour, including:

- The purpose of the tour
- The chosen POE, internal MGRA, and mode
- The pass type held by the border crossing tour (none, Ready, SENTRI)
- The start and end time of the tour

We also append the following data to the estimation file:

- Land-use data (households, employment by type, and enrollment by grade group) for each sampled MGRA
- A population accessibility term for each POE (a destination choice logsum measuring the population of each Colonia in Mexico weighted by distance to the POE)
- The observed wait time to cross at each POE by pass-type and auto versus pedestrian according to the observed tour start time.

A program (*processODchoicedata.awk)* was written to read the various files, merge the data from the estimation data bundle, and sample 99 combinations of POE and internal MGRA for each tour from the approximately 15k combinations of POE and MGRA contained in the ActivitySim output file. To this is appended information for the chosen POE and MGRA combination, resulting in 100 total alternatives for each tour.

We use ALOGIT to estimate the joint POE, MGRA, and crossing mode choice models for each purpose. We read in the output file from the above program and expand the alternatives four times (one for each mode). We set the chosen alternative according to the chosen mode for the tour. We then estimate a simultaneous model of POE station, internal destination, and border crossing mode. Estimation results are shown in Table 14, which includes the estimated parameter, the t-statistic, and the ratio of the parameter to the wait time variable in order to evaluate the reasonableness of the parameter in terms of minutes of wait time. The variables include:

- Mode choice logsum: The mode choice logsum is the weighted utility of all modes (drive-alone, shared-ride 2, shared-ride 3+, walk, walk-transit) for each border crossing mode between the POE and the internal MGRA.
- Wait time: The observed wait time for the POE, border crossing mode, and pass type according to the observed tour crossing time.
- Distance: The off-peak distance in miles between the POE and the primary destination.
- Alternative-specific constants: A set of constants on POE (San Ysidro is the base alternative) and border crossing mode (drive-alone is the base alternative)
- Population logsum: The distance-weighted population of each Colonia to the POE
- Size terms: Attributes of the internal MGRA including population, enrollment by type, and employment by type according to the following groups:
    - Construction = emp_const_non_bldg_prod + emp_utilities_prod + emp_const_bldg_prod + emp_mfg_prod + emp_whsle_whs + emp_trans
    - Office = emp_const_non_bldg_office + emp_utilities_office + emp_const_bldg_office + emp_mfg_office + emp_personal_svcs_office + emp_personal_svcs_office

- Retail = emp_retail
- Amusement, hotel, restaurant & bar = emp_amusement + emp_hotel + emp_restaurant_bar
- Other = emp_total - (emp_ag + Construction + Office + Retail + Amusement, Hotel, Restaurant)

The mode choice logsum term is positive and larger than models estimated previously. The wait time coefficient is negative and significant, and roughly three times more onerous than the coefficient estimated previously; this suggests that the more refined wait time data used in the current estimation may have improved the estimation results. Distance coefficients are significant and negative for all models except Visiting and Other tours. Note that inclusion of distance terms in the model specification reduce the explanatory power of the logsum coefficient. However, the improvement in goodness of fit when including such terms indicates a better model when including them. Therefore the logsum coefficients are held to the estimated values without a distance term; the reported t-statistic in the table of estimation results reflects this model estimation.

The size of the alternative-specific constants are generally reasonable; all mode constants are negative for all purposes, reflecting disutility for sharing a ride and pedestrian crossing modes to driving alone, all else being equal. Note that it is not possible to estimate an alternative-specific constant for crossing at Tecate for the school purpose due to a lack of observed tours. In application, Tecate should probably be turned off for school tours. The coefficient on the population logsum term at each POE is held to 1.0, since it is essentially a POE size term.

The size terms for each MGRA are generally reasonable as well. For work tours, households (the base category) has the highest coefficient, followed by construction employment. Retail employment and amusement/hotel/restaurant employment have similar size terms. For school tours, College and K-12 enrollment have similar size terms. Retail employment logically has the biggest size term for Shop tours, followed by amusements/hotel/restaurants and other employment. Visiting and other tour purposes are collapsed due to lack of adequate number of observations in each tour purpose alone. For these purposes, households have the highest size term followed by retail employment; office, amusement/hotel/restaurant and other employment have similar, small size term coefficients.

TABLE 14: CROSS-BORDER TOUR POE, INTERNAL DESTINATION, AND BORDER CROSSING MODE ESTIMATION RESULTS

	WORK			SCHOOL			SHOP			VISITING AND OTHER TOURS		
Variables	Coeff.	T-Stat	Ratio to Wait Time	Coeff.	T-Stat	Ratio to Wait Time	Coeff.	T-Stat	Ratio to Wait Time	Coeff.	T-Stat	Ratio to Wait Time
Mode choice logsum	0.349	15.49	-7.0	0.393	8.88	-2.8	0.449	31.18	-4.9	0.346	9.45	-6.9
Wait time	-0.050	-4.31	1.0	-0.139	-3.54	1.0	-0.092	-9.80	1.0	-0.050	-2.58	1.0
Population logsum	1.000	Fixed	-20.2	1.000	Fixed	-7.2	1.000	Fixed	-10.9	1.000	Fixed	-20.1
Distance	-0.029	-2.90	0.6	-0.055	-2.26	0.4	-0.132	-12.27	1.4			
Constants												
Otay Mesa POE	-0.507	-2.25	10.2	-0.575	-1.27	4.1	-0.091	-0.61	1.0	0.638	1.82	-12.8
Tecate POE	0.654	1.18	-13.2	0.000	NA	0.0	-0.521	-0.84	5.7	1.041	1.05	-20.9
Shared 2	-1.925	-6.92	38.8	-0.741	-1.73	5.3	-0.226	-1.81	2.5	-0.369	-0.92	7.4
Shared 3+	-3.630	-7.11	73.2	-2.119	-3.73	15.3	-1.425	-9.98	15.5	-1.580	-3.15	31.7
Pedestrian	-4.402	-6.99	88.8	-5.522	-3.76	39.8	-6.860	-13.69	74.8	-2.742	-3.01	55.1
Exponentiated Size terms												
households	1.000									1.000		
construction	0.499											
office	0.088									0.106		
retail	0.247						1.000			0.309		
amusement, hotel, restaurant	0.258						0.630			0.173		
other	0.148						0.414			0.247		
College enrollment				1.000								
K-12 enrollment				0.934								

Observations	194.000	47.000	504.000	47.000
Initial Likelihood	-818.020	-247.421	-2210.309	-265.573
Final Likelihood	-686.555	-220.620	-1885.470	-222.822

4.5 STOP FREQUENCY MODEL

The stop frequency model assigns to each tour the number of intermediate destinations a person will travel to on each leg of the tour from the origin to tour primary destination and back. The model incorporates the ability for more than one stop in each direction, up to a maximum of 3, for a total of 8 trips per tour (four on each tour leg).

The stop frequency model uses a probability distribution derived from the cross border survey data. The ActivitySim model is limited to a maximum of 3 stops per half tour. Survey tours with more than 3 stops are collapsed to a max of 3 stops. The stop frequency distribution is segmented by tour purpose, tour duration bin (0-4 hours, 4-8 hours, or 8-24 hours), number of outbound stops (0-3), and number of inbound stops (0-3). This distribution is shown in the appendix Table 28.

4.6 STOP LOCATION CHOICE MODEL

The stop location choice model predicts the MGRA of each intermediate stop on each tour. The model is estimated using a sample of 499 alternatives (plus the chosen alternative) selected randomly from the full set of approximately 5000 alternatives output from ActivitySim. The stop destination choice model constructs a utility for each stop based on the following variables:

- The mode choice logsum from the last location to the stop location, and from the stop location to the next known location.
- The distance from the last location to the stop location, and from the stop location to the next known location.
- The size of the stop location, based on the purpose of the stop.

In order to estimate the model, we first modified the utility expression file for the trip (stop) destination choice model to write out mode choice logsums and distance terms for each sampled MGRA. Running this revised model specification through ActivitySim's estimation mode resulted in an estimation data bundle containing, for each stop on each observed cross-border tour, a sample of alternatives consisting of stop locations (MGRAs). The sample includes one MGRA for each TAZ (chosen at random from within the TAZ according to the ratio of the size term of the MGRA to the total size term across all MGRAs within the TAZ). The estimation data bundle includes the mode choice logsum and distance terms described above. To this data is appended information on the observed cross border tour and stop, including:

- The chosen MGRA and mode
- The purpose of the tour and the purpose of the stop

- The number of the trip (stop) and the total number of trips on the tour leg (also known as the tour "half"; the one-way movement from the tour origin to the primary destination, or from the primary destination back to the tour origin.
- Whether the stop is on the outbound leg or the return (inbound) leg of the tour
- The start time and duration of the tour

We also append land-use data (households, employment by type, and enrollment by grade group) for each sampled MGRA

A program (*processTripDestChoicedata.awk)* was written to read the various files, merge the data from the estimation data bundle, and sample 499 combinations of MGRA for each stop from the approximately 5k MGRAs contained in the ActivitySim output file. To this is appended information for the chosen MGRA, resulting in 500 total alternatives for each tour.

We use ALOGIT to estimate the stop destination choice model. We tried various model specifications and selected a preferred model specification based upon reasonableness of estimated parameters, significance, and goodness-of-fit. We estimate two models; one for work tours and one for non-work tours. Estimation results are shown in Table 15, which includes the estimated parameter and the t-statistic.

The variables include:

- Mode choice logsum: The mode choice logsum from the origin to the stop plus the stop to the destination. The coefficient is positive and significant.
- Distance terms: A set of distance terms constraining total distance and reflecting the tendency of stops to be located close to the tour origin. Note that the inclusion of these terms significantly improved model goodness of fit but reduced the explanatory power of the mode choice logsum due to correlation. Therefore, the logsum coefficient is held to the value estimated in a previous specification without inclusion of the distance term, and the model is re-estimated with the distance terms. The reported t-statistic for the mode choice logsum is from the previous specification. The distance terms include the following:
    - Total distance on the journey (origin to stop distance plus stop to destination distance).
    - Origin->stop distance for the first stop on the tour (first stop in outbound direction). The negative coefficient reflects a tendency to locate the stop close to the POE, which is the origin for the tour.
    - Stop->destination distance for the last stop on a tour (last stop inbound direction). The negative coefficient reflects a tendency to locate the stop close to the POE, which is the origin for the tour.

- A size term calculated according to stop purpose based on the size terms shown in Table 14, with a coefficient constrained to 1.0.

We also tried other distance terms and interacting distance with tour duration. The results are illogical or insignificant.

TABLE 15: STOP DESTINATION CHOICE MODEL RESULTS

	WORK TOUR RESULTS		NON-WORK TOUR RESULTS	
Parameter	Coeff	T-Stat	Coeff	T-Stat
Mode Choice Logsum	0.7870	20.80	1.00	38.35
Total distance			-0.0364	-4.87
Origin->Stop distance first stop	-0.2152	-7.07	-0.0337	-1.61
Stop->Dest distance last stop	-0.0897	-3.75	-0.0720	-3.96
Size Term (by stop purpose)	1.000	fixed	1.000	fixed
Observations	256		871	
Initial Likelihood	-1022		-2894	
Final Likelihood	-972		-2845	

5.0 MODEL CALIBRATION

The data described in Chapter 2 of this report is used to calibrate the cross-border special market implementation. This chapter describes the calibration of this model.

The following model components are calibrated as part of this effort:

- Tour origin-destination choice: This model simultaneously predicts the point of entry (POE) into San Diego County, and the destination Master Geographic Reference Area (MGRA) of the tour primary destination in San Diego County. The model assumes that the POE is also used to exit the United States at the end of the tour. Alternative-specific constants are calibrated to better match the distribution of tours by POE and the Pseudo-Metropolitan Statistical Area (PMSA) containing the tour primary destination.
- Tour mode choice: The tour mode choice model predicts the mode used to cross into and out of San Diego County at the POE. Alternative-specific constants are adjusted to better match the observed distribution of tours by crossing mode (drive-alone, shared 2, shared 3+, and walk).
- Trip mode choice: The trip mode choice model predicts the mode for each trip on the tour. Alternative-specific constants are adjusted to better match the observed distribution of trips by trip mode (drive-alone, shared 2, shared 3+, walk, walk-transit, taxi, and Transportation Network Company (TNC)).

For purposes of calibration, we apply the Cross-Border Model using 2016 input land-use and network data but calibrate the model to 2019 survey conditions. Calibration of constants follows the standard process of taking the natural log of the observed share of the segment over the model share of the segment and adding to the existing constant for that segment. If we observe that the model oscillates drastically between calibration iterations, we damp the factor using a multiplier of 0.1 to 0.4 depending on the extent of the oscillation, using professional judgement. Other model components in the Cross-Border Model (tour purpose, tour pass type (none, SENTRI pass, READY pass) ownership, tour scheduling, trip purpose, and trip departure time) utilize observed distributions rather than a statistical model for Monte Carlo simulation. These model components are not calibrated since the observed distributions are generated from observed data. However, we summarize the results of applying these model components below to ensure that they adequately represent the observed distributions.

5.1 TOUR PURPOSE AND PASS TYPE

The absolute number of person tours by residents of Mexico crossing into and out of San Diego county are specified in a user-provided input. These tours are then attributed with a tour purpose (Work, School, Shop, Social/Visit, or Other) and a pass type (none, SENTRI, or READY). These models are applied by simulating from observed probability distributions. These distributions are shown in Figure 30.

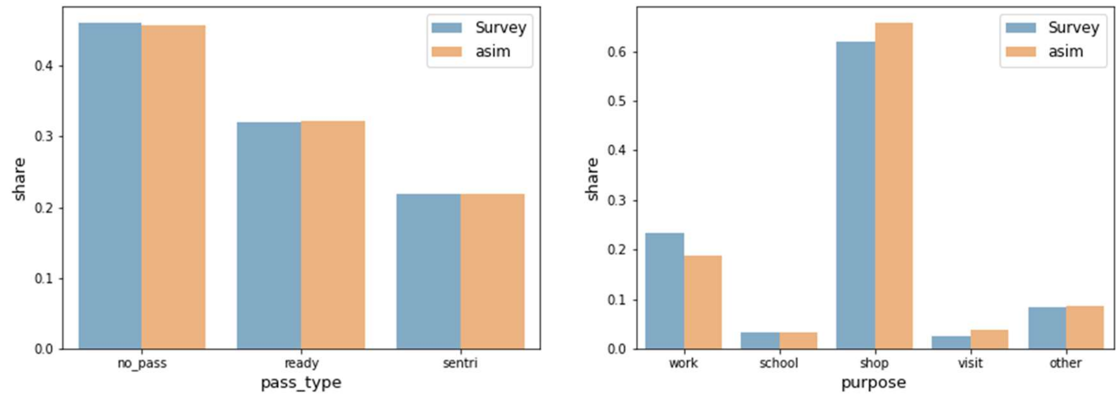


FIGURE 30: TOUR PASS TYPE AND PURPOSE

5.2 TOUR SCHEDULING

The tour scheduling model predicts the half-hour time period that the tour crosses into and out of the United States. There are 48 half-hour period alternatives in the model. This model uses Monte Carlo simulation from an observed probability distribution that is applied based on tour purpose. Tour start and end distributions are directly derived from observed border crossings in the 2019 Survey. The Cross-Border Model schedules tours directly from this distribution which ensures the distributions match very closely. The tour departure half-hour distribution is shown in Figure 31. Most tours cross the border between 7am and 3pm. The tour return distribution is shown in Figure 32. Most tours return after 12pm.

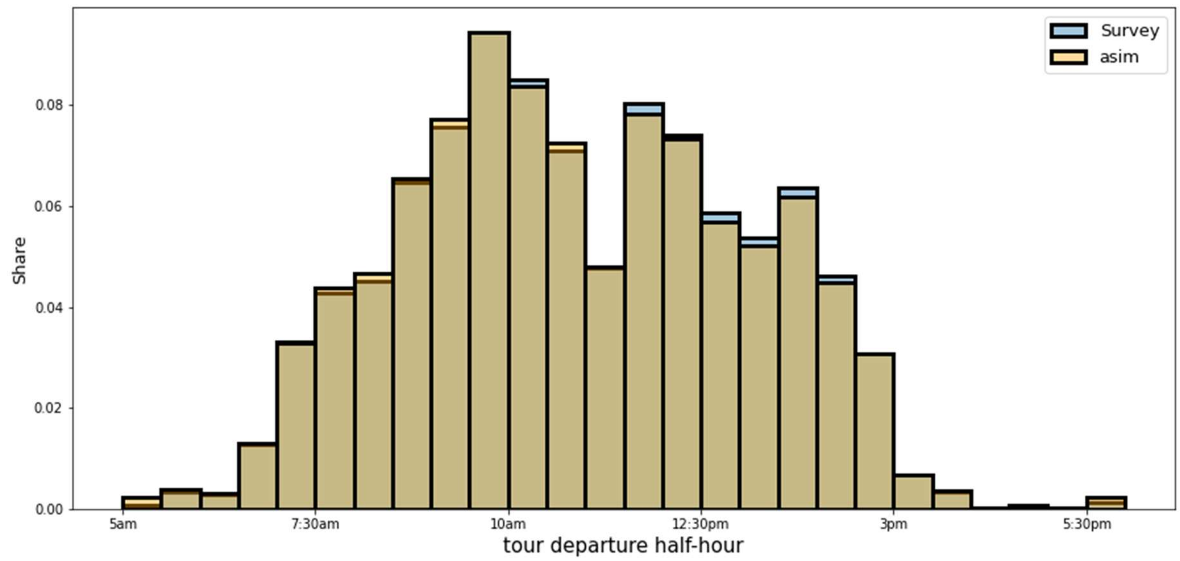


FIGURE 31: TOUR DEPARTURE HALF-HOUR

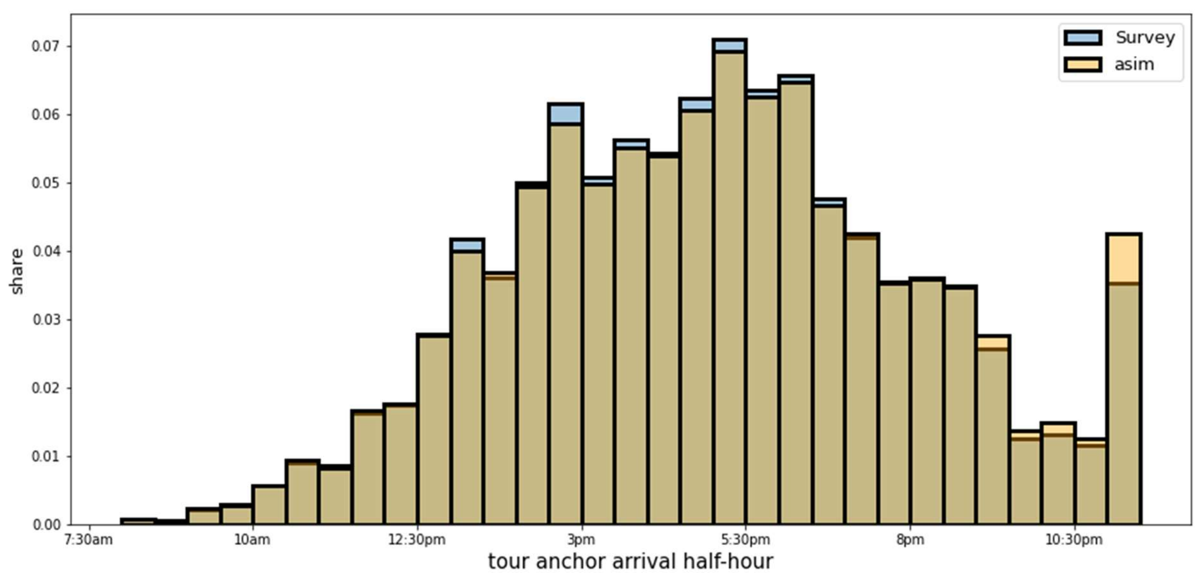


FIGURE 32: TOUR ARRIVAL HALF-HOUR

5.3 TOUR ORIGIN-DESTINATION CHOICE

The tour origin POE and primary destination MGRA choice are simultaneously modeled based on POE attributes including a population logsum term at each POE, the wait time by border crossing mode at each POE, the size term for each tour primary destination MGRA, the tour mode choice logsum between the POE and the tour primary destination, and the distance between the POE and the tour primary destination MGRA. Tour origin-destination choice models are calibrated to better match the shares of tour origins by POE, tour primary destination by pseudo-Metropolitan Statistical Area (PMSA), and the share of tours by tour purpose and POE. These results are described below.

Port of Entry Calibration

In the base-year, cross-border tours cross at one of three existing ports of entry. The largest is San Ysidro, Otay Mesa further to the east, and Tecate, which is the eastern most crossing as shown in Figure 33. Port of entry wait times are modeled based on the total number of lanes at the POE and the total crossing volume. For purposes of calibration, observed crossing wait times are used in the tour origin-destination choice model. Two constants are added for the total crossing volume through Otay Mesa and Tecate to help distribute the tours across the POEs shown in Figure 16. The final crossing shares by POE closely match the observed crossing volumes shown in Figure 34.

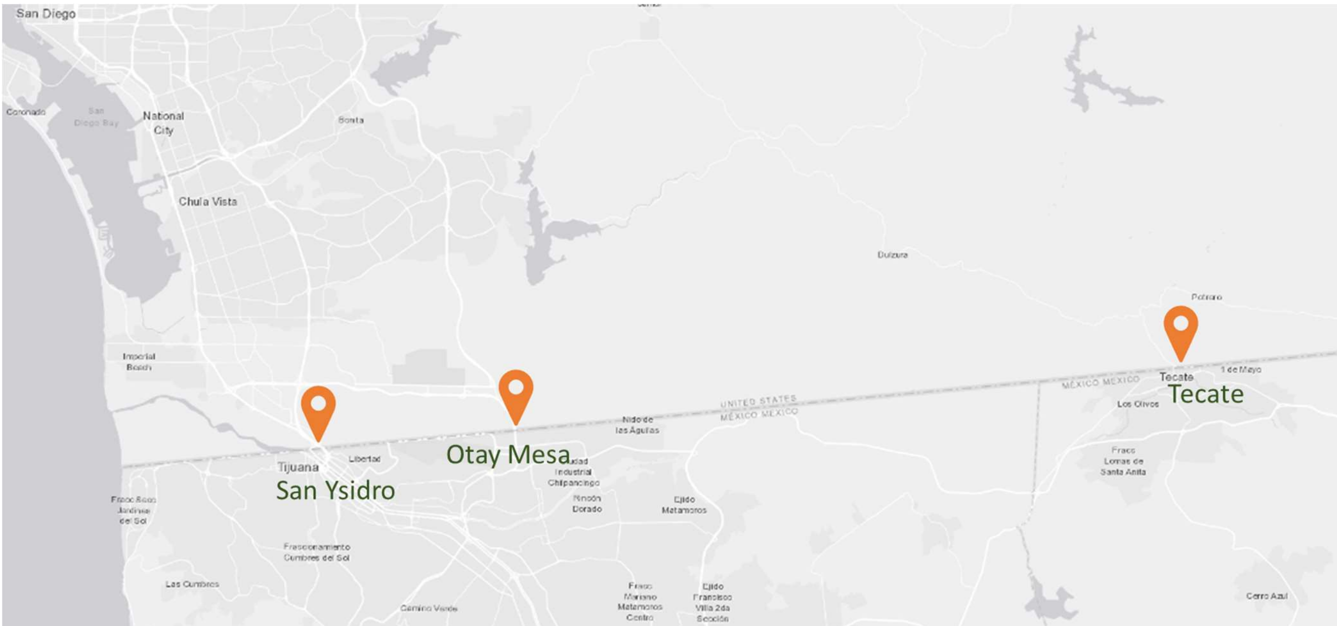


FIGURE 33: SAN DIEGO BORDER PORTS OF ENTRY

TABLE 16: POE CROSSING CONSTANTS

CONSTANT	DESCRIPTION	VALUE
coef_tecate	Total crossing volume through Tecate	3.87
coef_otay	Total crossing volume through Otay Mesa	0.49

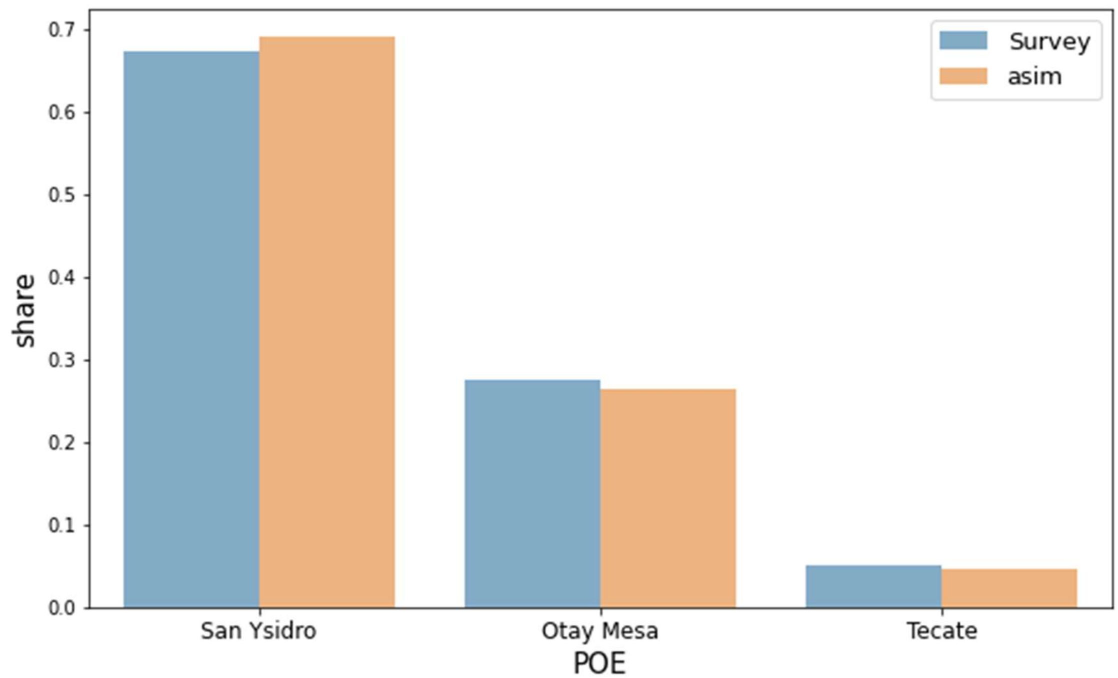


FIGURE 34: POE CHOICE

Destination Pseudo-MSA Calibration

The SANDAG region is divided into PMSA districts as shown in Figure 35. Matching the PMSA destination helps calibrate the average tour distances and better consistency between estimated and observed traffic volumes and transit boardings. Initial model results indicated that the origin-destination choice model generally over-estimated tour destinations to PMSA 2 (San Diego) and under-estimated destinations to PMSA 4 encompassing the area just to the north of the San Ysidro and Otay Mesa POEs. The model also under-estimated tours crossing at Tecate to PMSA 8 in the east county. Constants are added to calibrate tour destinations to PMSAs 2 and 4 at all POEs and destination PMSA 8 at Tecate. The final constants are shown in Table 17. The estimated versus observed share of tour primary destination PMSA for all tours is shown in Figure 36 , and Figure 37 shows the choice by POE. After calibration, the tour destination choice by PMSA closely matches the observed tour destinations and the tour distances fall within 5% of observed values.

.

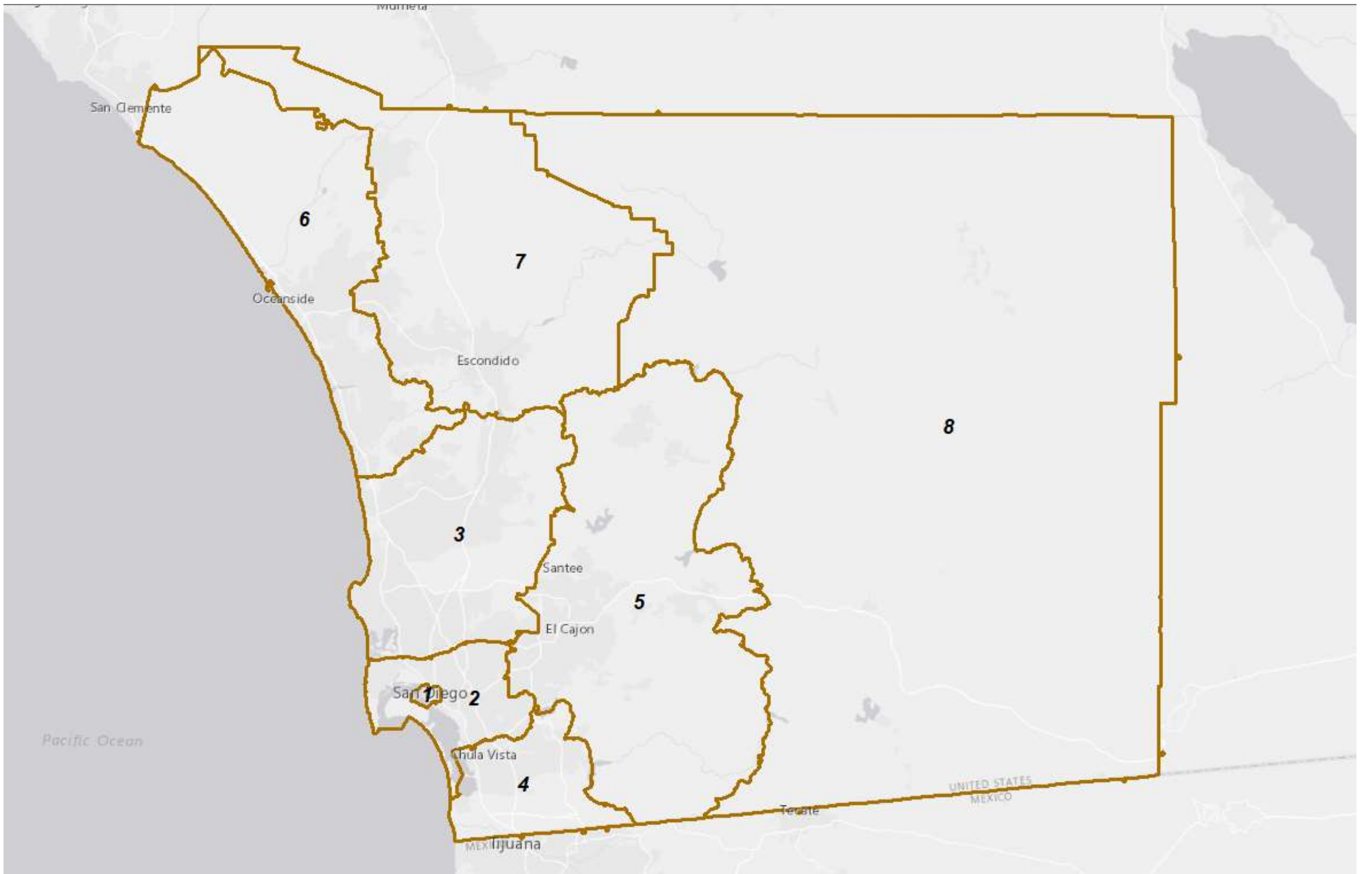


FIGURE 35: SANDAG PSEUDO-MSA BOUNDARIES

TABLE 17:POE TO DESTINATION PSEUDO-MSA CONSTANTS

DESTINATION PMSA CONSTANTS	DESCRIPTION	VALUE
coef_san_ysidro_pmsa2	San Ysidro to PMSA 2	-1.03572
coef_san_ysidro_pmsa4	San Ysidro to PMSA 4	0.747354
coef_otay_mesa_pmsa2	Otay Mesa to PMSA 2	-1.4706
coef_otay_mesa_pmsa4	Otay Mesa to PMSA 4	0.43201
coef_tecate_pmsa2	Tecate to PMSA 2	-1.70846

DESTINATION PMSA CONSTANTS	DESCRIPTION	VALUE
coef_tecate_pmsa4	Tecate to PMSA 4	0.016925
coef_tecate_pmsa8	Tecate to PMSA 8	1.311931

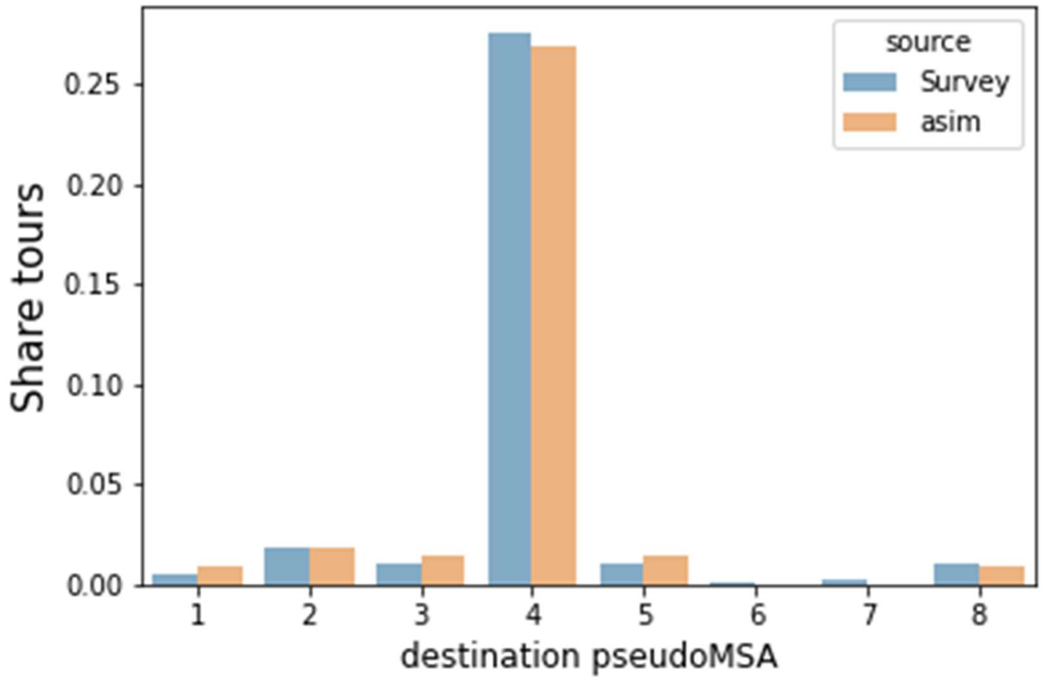


FIGURE 36: DESTINATION PSEUDOMSA CHOICE

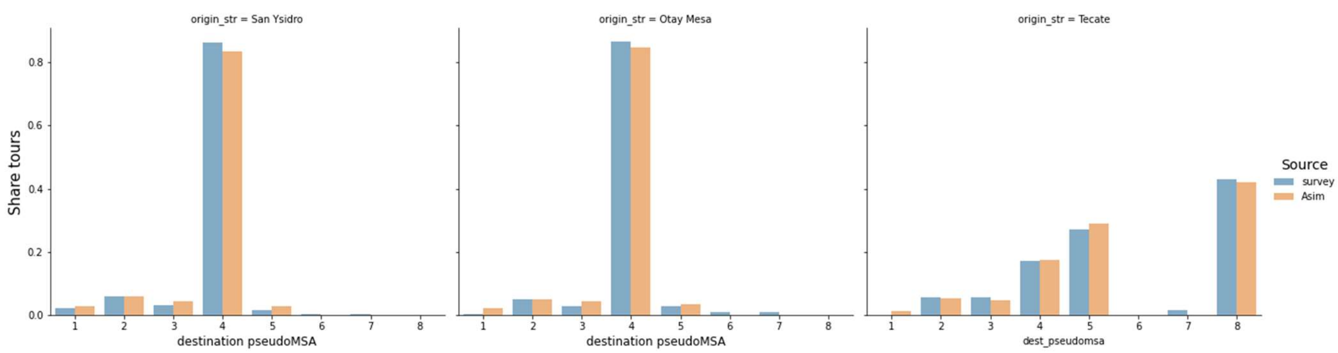


FIGURE 37: DESTINATION PSEUDO MSA CHOICE BY POE

With tour origin and destination calibrated well, the tour distances match the survey tour distances overall within 5.1%. Tour distances by POE are shown in Table 18. San Ysidro has the highest number of crossings and shows a somewhat larger difference in average tour distance compared to the other ports of entry. However, the average estimated difference to the tour primary destination is only one mile longer than the average observed distance. Tour distance by tour mode is shown in Table 19 and each mode is well within 10% of the observed tour distances.

TABLE 18: HALF TOUR DISTANCE BY POE

ORIGIN_STR	ASIM MEAN TOUR DIST	SURVEY MEAN TOUR DIST	DIFF	PCT DIFF
San Ysidro	8.45	7.40	1.05	14.10%
Otay Mesa	11.13	11.49	-0.36	-3.10%
Tecate	20.48	21.37	-0.89	-4.10%
Total	9.72	9.25	0.47	5.10%

TABLE 19: HALF TOUR DISTANCE BY MODE

TOUR_MODE	ASIM MEAN TOUR DIST	SURVEY MEAN TOUR DIST	DIFF	PCT DIFF
DRIVEALONE	11.45	10.83	0.62	5.70%
SHARED2	10.40	9.62	0.77	8.00%
SHARED3	9.91	9.34	0.56	6.00%
WALK	7.92	7.92	-0.001	0.00%
Total	9.72	9.25	0.47	5.10%

POE By Purpose Calibration

Tour purposes are simulated based on the observed purpose shares from the survey. The overall shares are simulated very closely to the observed purposes as shown in Figure 38. Tour purpose shares for Otay Mesa and Tecate are then calibrated to ensure the shares at each POE also matched the survey distribution. Visit and Other tours are estimated simultaneously so they are also calibrated together. The final purpose by POE constants are shown in Table 20. The final distribution of tour purposes by POE is shown in Figure 39.

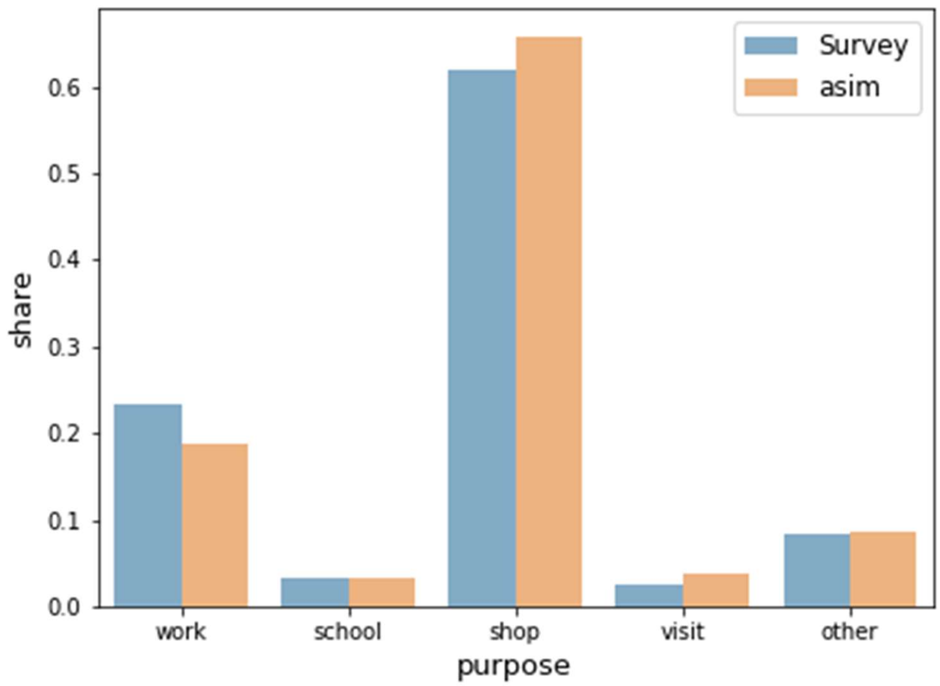


FIGURE 38: TOUR PURPOSE SHARE

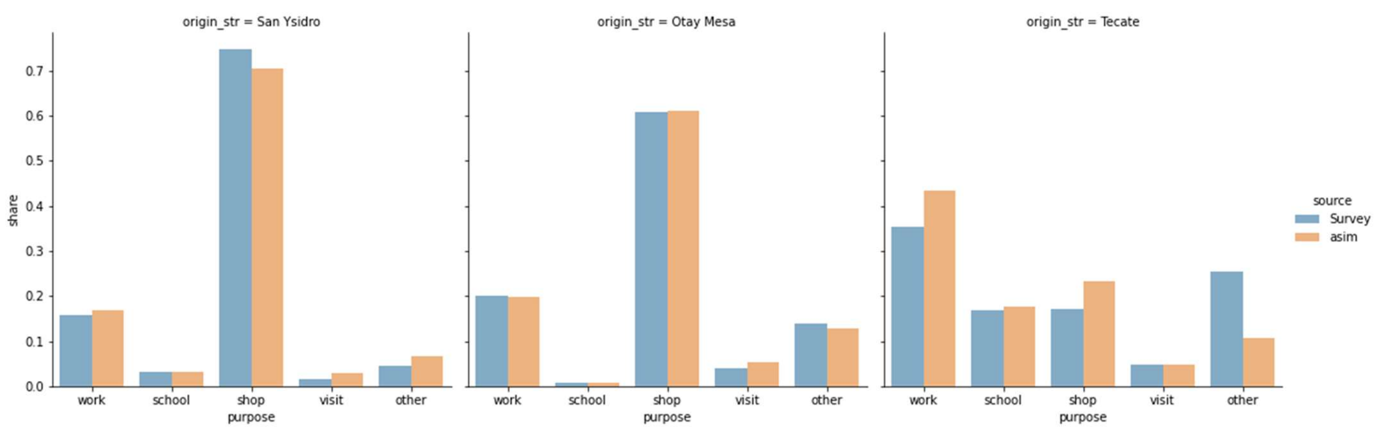


FIGURE 39: PURPOSE SHARES BY POE

TABLE 20: POE PURPOSE SHARE CONSTANTS

CONSTANT	DESCRIPTION	VALUE
coef_otay_mesa_work	Work Tours through Otay Mesa	-0.07558
coef_otay_mesa_school	School Tours through Otay Mesa	-1.63495
coef_otay_mesa_shop	Shop tours through Otay Mesa	-0.19834
coef_otay_mesa_visit_other	Visit and Other tours through Otay Mesa	0.679925
coef_tecate_work	Work tours through Tecate	-0.7198
coef_tecate_school	School tours through Tecate	1.133874
coef_tecate_shop	Shop tours through Tecate	-1.53362
coef_tecate_visit_other	Visit and Other tours through Tecate	4.092507

5.4 TOUR MODE CHOICE

Tour mode choice should be well calibrated to the survey as the model should be sensitive to effects from infrastructure and policy changes at different ports of entry on tour mode. The initial estimated tour mode choice model significantly over-estimated tours crossing by shared 2 and underestimated other modes. Alternative specific constants are adjusted to better match observed distributions of tour crossing mode. Figure 40 shows the overall tour mode distributions after mode choice calibration which match the survey well. Alternative specific constants for tour mode are calibrated by tour purpose. The final calibrated constants are shown in Table 21. The final distribution of tour mode by tour purpose is shown in Figure 41.

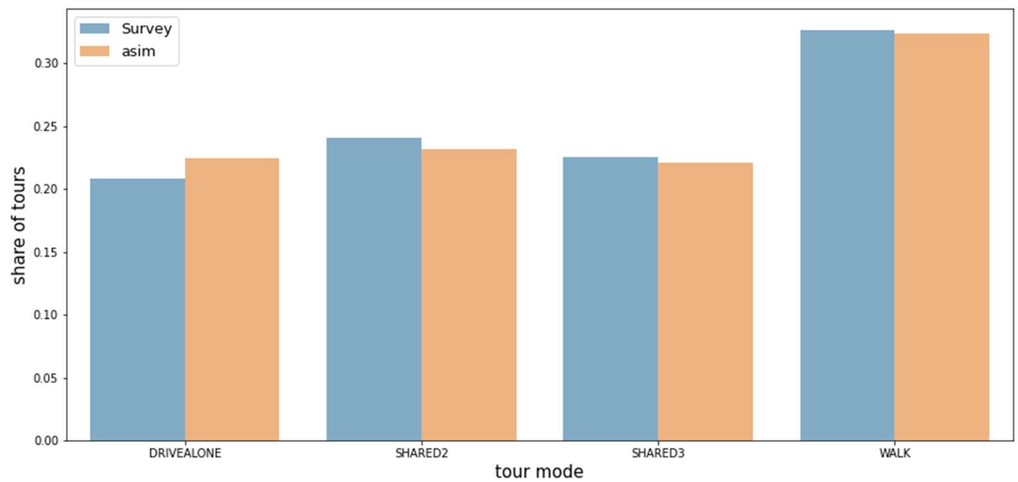


FIGURE 40: TOUR MODE CHOICE

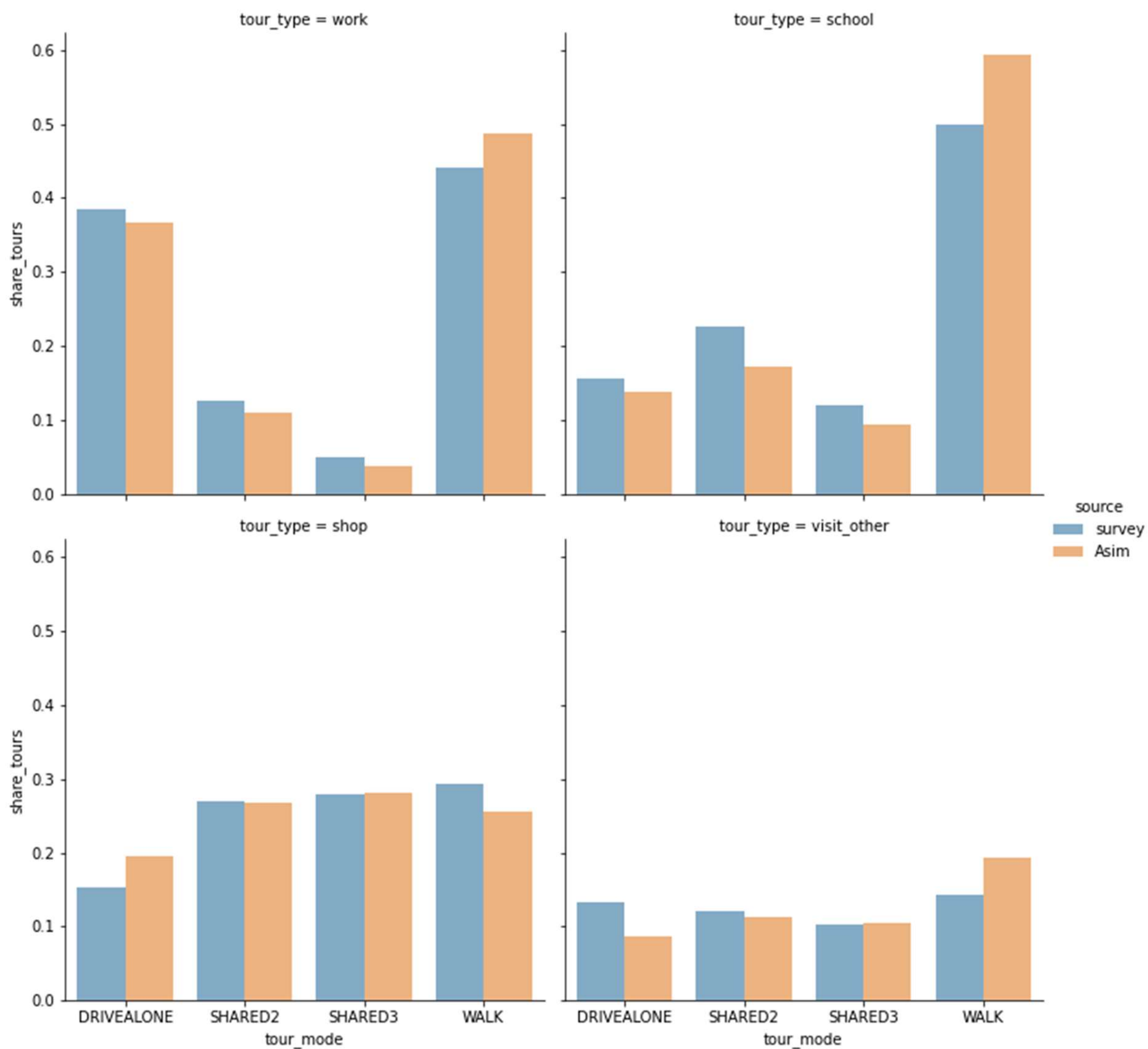


FIGURE 41: TOUR MODE BY TOUR PURPOSE

TABLE 21: TOUR MODE BY PURPOSE CONSTANTS

TOUR MODE BY PURPOSE CONSTANTS	DESCRIPTION	VALUE
asc_SHARED2_work	Shared 2 work tours	-1.56548
asc_SHARED2_school	Shared 2 school tours	-0.47243

TOUR MODE BY PURPOSE CONSTANTS	DESCRIPTION	VALUE
asc_SHARED2_shop	Shared 2 shop tours	-0.53093
asc_SHARED2_visit_other	Shared 2 visit and other tours	-0.75168
asc_SHARED3_work	Shared 3 work tours	-2.83621
asc_SHARED3_school	Shared 3 school tours	-1.37447
asc_SHARED3_shop	Shared 3 shop tours	-0.73553
asc_SHARED3_visit_other	Shared 3 visit and other tours	-1.1503
asc_WALK_work	Walk work tours	-3.7506
asc_WALK_school	Walk school tours	-5.66453
asc_WALK_shop	Walk shop tours	-6.8358
asc_WALK_visit_other	Walk visit and other tours	-3.38339

Tour mode choice is also calibrated by port of entry. The constants for tour mode choice by port of entry are shown in Table 22. The distribution of tour mode by POE is shown in Figure 42.

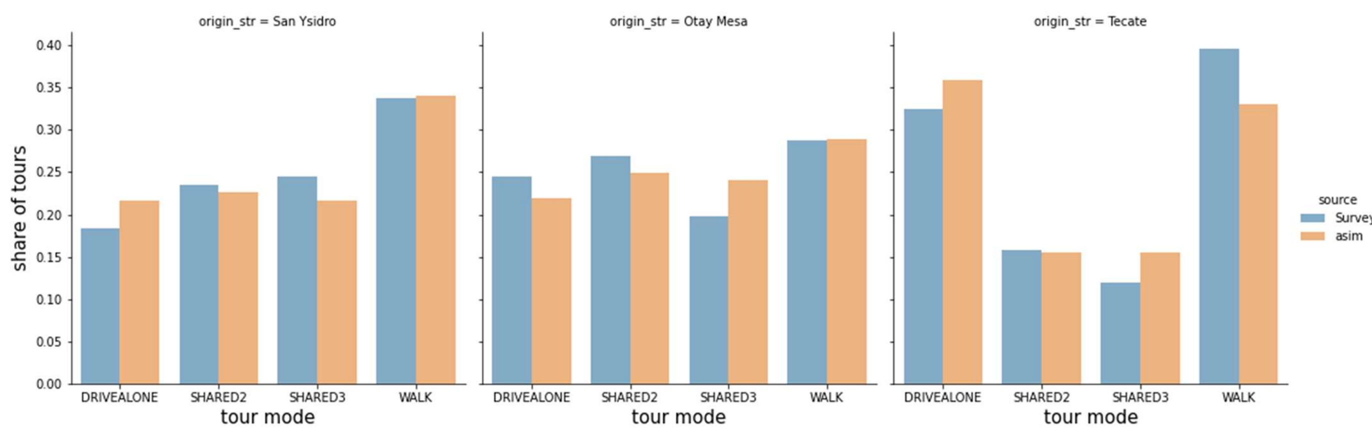


FIGURE 42: TOUR MODE CHOICE BY POE

TABLE 22: TOUR MODE CHOICE BY POE CONSTANTS

TOUR MODE BY POE CONSTANT	DESCRIPTION	VALUE
otay_SHARED2	Shared 2 through Otay Mesa	0.24852
otay_SHARED3	Shared 3 through Otay Mesa	0.359713
otay_WALK	Walk through Otay Mesa	-0.72142
tecate_SHARED2	Shared 2 through Tecate	-0.43687
tecate_SHARED3	Shared 3 through Tecate	-0.1948

tecate_WALK	Walk through Tecate	0.662788

5.5 STOP FREQUENCY

The stop frequency model assigns to each tour the number of intermediate destinations a person will travel to on each leg of the tour from the origin to tour primary destination and back. The model incorporates the ability for more than one stop in each direction, up to a maximum of 3, for a total of 8 trips per tour (four on each tour leg). The stop frequency model uses a probability distribution derived from the cross-border survey data. The ActivitySim model is limited to a maximum of 3 stops per half tour. Survey tours with more than 3 stops are collapsed to a max of 3 stops. The stop frequency distribution is segmented by tour purpose, tour duration bin (0-4 hours, 4-8 hours, or 8-24 hours). This distribution is sampled for assigning trips to tours and this is shown in Figure 43. The number of trips per tour by purpose is shown in Figure 44.

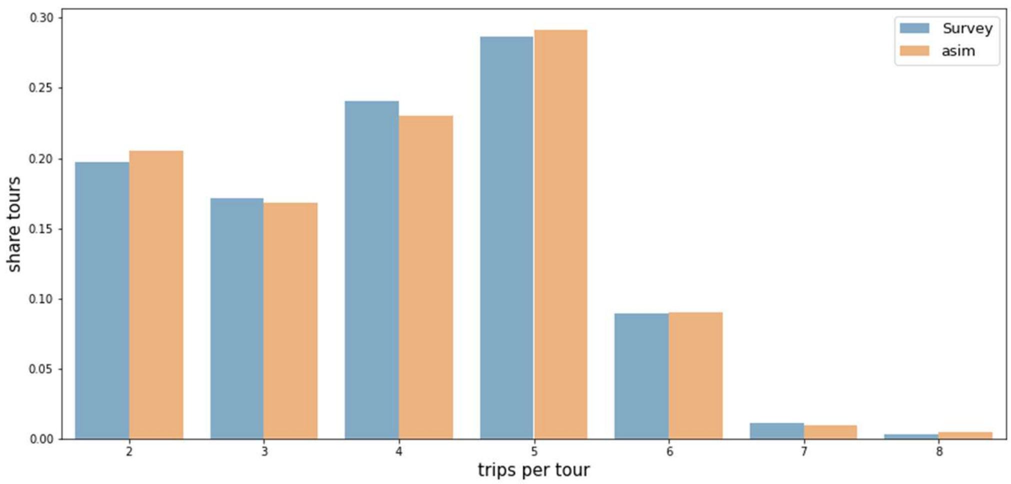


FIGURE 43: TRIPS PER TOUR

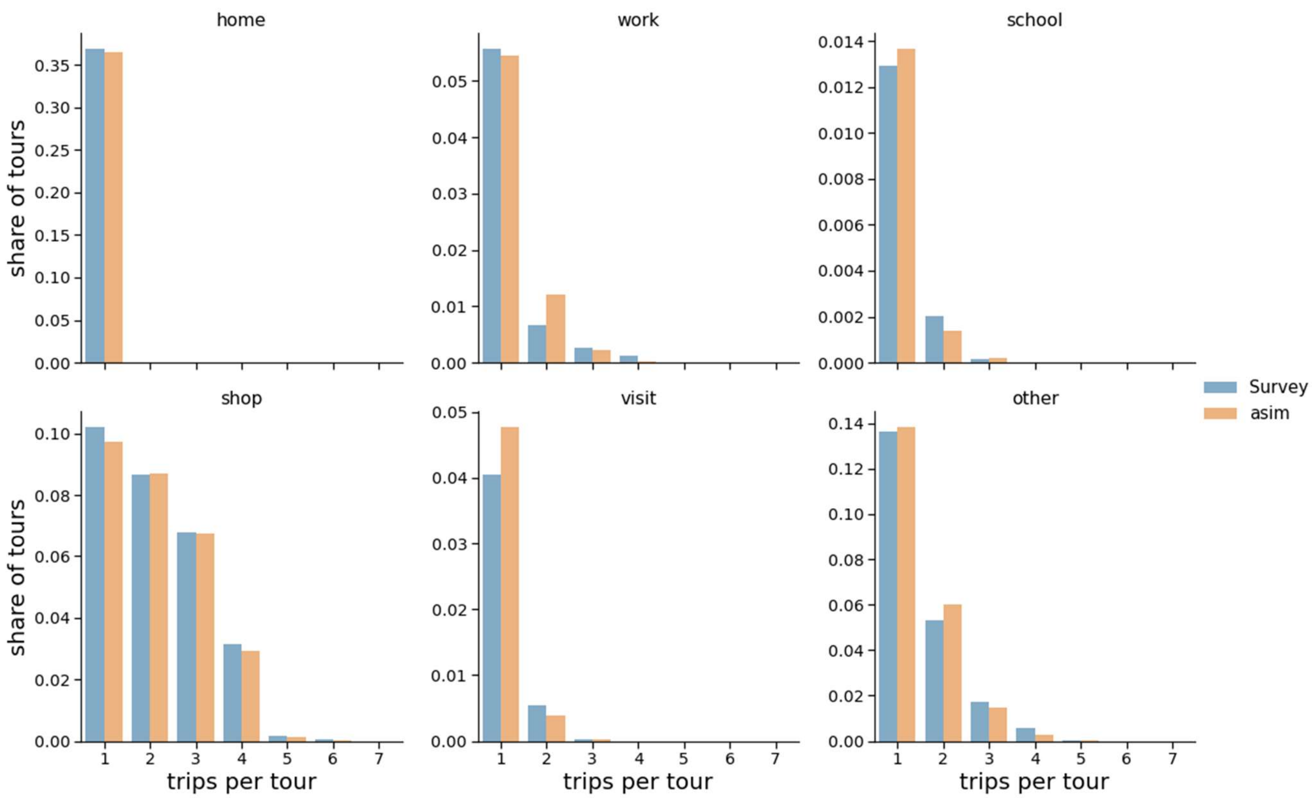


FIGURE 44: TRIPS PER TOUR BY PURPOSE

5.6 TRIP SCHEDULING

Similar to tour scheduling, trip scheduling is sampled from a distribution derived from the 2019 survey data. The outbound stop distribution is segmented by tour purpose, and the upper and lower bounds of remaining half hour periods after the last scheduled trip. The outbound stop duration is then segmented by likelihood the stop departure is in the last inbound/outbound trip period plus 1 to 11 periods. Inbound stop duration is segmented by tour purpose, and the upper and lower bounds of remaining half hour periods after the last scheduled trip. The inbound stop duration is then segmented by likelihood the stop departure is in the tour arrival period plus minus 1 to 7 periods Outbound and inbound trip scheduling results are shown in Figure 45 and Figure 46 respectively. The figures show a reasonably good match between estimated and observed distributions of trips by departure tine (outbound) and arrival time (inbound).

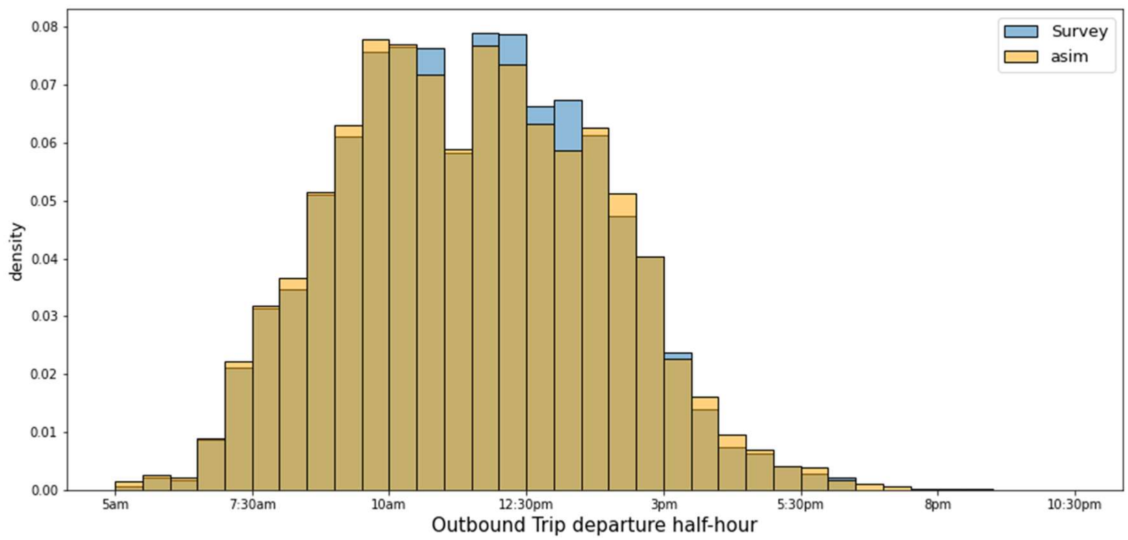


FIGURE 45: OUTBOUND TRIP DEPARTURE

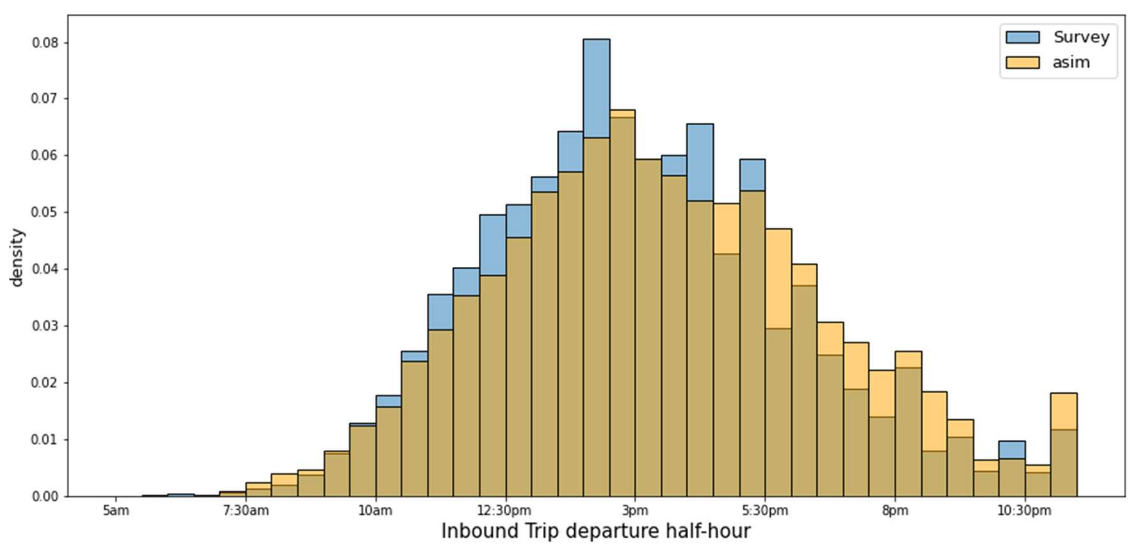


FIGURE 46: INBOUND TRIP DEPARTURE

5.7 TRIP MODE CHOICE

Trip mode choice is calibrated to match the observed distribution of trips by trip mode. The final alternative-specific constants are shown in Table 23, and the distribution is shown in Figure 47. The alternative specific constants for trip mode by tour mode are also calibrated. The calibrated constants are shown in Table 24, and the distribution is shown in Figure 48. Note that unavailable reported trip modes (such as trolley are shown in the table as Other, but only for the survey trips.

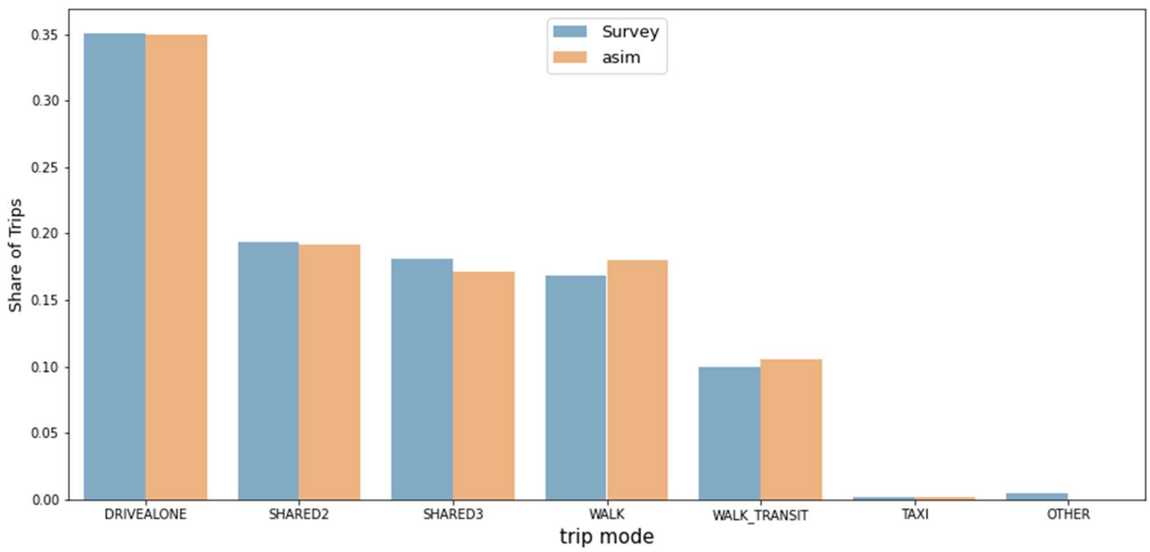


FIGURE 47: TRIP MODE CHOICE

TABLE 23: TRIP MODE CHOICE CONSTANTS

TRIP MODE CONSTANTS	DESCRIPTION	VALUE
coef_s2	Shared 2 trip	0.22603
coef_s3	Shared 3 trip	-0.18617
coef_walk	Walk trip	-1.47438
coef_walk_transi	Walk to transit trip	0.889661

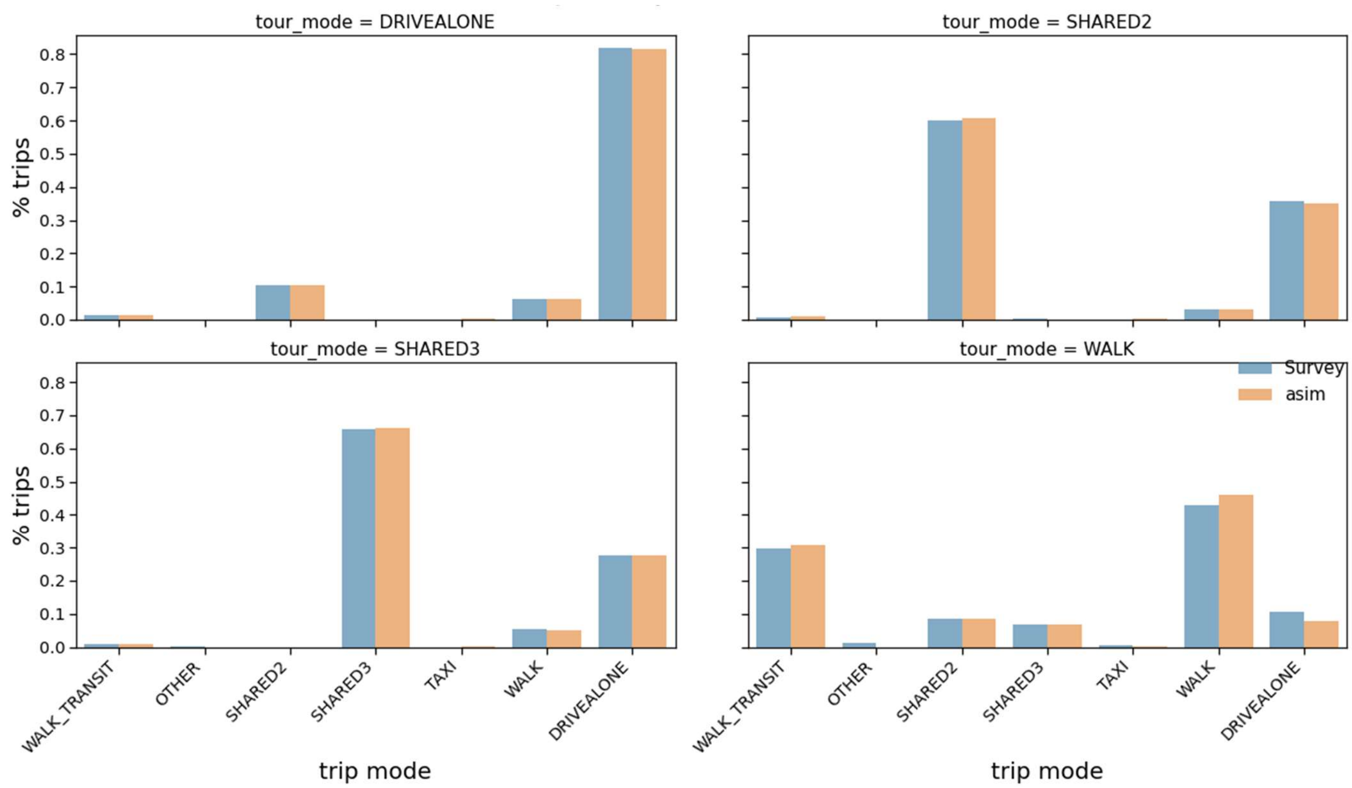


FIGURE 48: TRIP MODE CHOICE BY TOUR MODE

TABLE 24: TRIP MODE ALTERNATIVE SPECIFIC CONSTANTS

TRIP ALTERNATIVE SPECIFIC CONSTANT	DESCRIPTION	VALUE
s2_ASC_tour_da	Shared 2 trip on drive alone tour	-1.21775
s2_ASC_tour_s2	Shared 2 trip on Shared 2 tour	0.754488
s2_ASC_tour_s3	Shared 2 trip on shared 3 tour	1.4795
s2_ASC_tour_walk	Shared 2 trip on walk tour	0.720563
s3_ASC_tour_da	Shared 3 trip on drive alone tour	-5.94073
s3_ASC_tour_s2	Shared 3 trip on shared 2 tour	-5.35244
s3_ASC_tour_s3	Shared 3 trip on shared 3 tour	2.066856
s3_ASC_tour_walk	Shared 3 trip on walk tour	1.429926
walk_ASC_tour_da	Walk trip on drive alone tour	1.956542
walk_ASC_tour_s2	Walk trip on shared 2 tour	1.738203

TRIP ALTERNATIVE SPECIFIC CONSTANT	DESCRIPTION	VALUE
walk_ASC_tour_s3	Walk trip on shared 3 tour	2.611486
walk_ASC_tour_walk	Walk trip on walk tour	5.90987
walk_transit_ASC_tour_da	Walk to transit trip on drive alone tour	-3.47901
walk_transit_ASC_tour_s2	Walk to transit trip on shared 2 tour	-2.15687
walk_transit_ASC_tour_s3	Walk to transit trip on shared 3 tour	-1.6337
walk_transit_ASC_tour_walk	Walk to transit trip on walk tour	3.456636

6.0 SENSITIVITY TESTING

The calibrated model is put through several sensitivity tests to ensure the behavior is reasonable and expected in several directions. Five tests are chosen listed below:

1. Increase the Capacity at Otay Mesa
2. Decrease headways for transit service at San Ysidro
3. Increase employment around Otay Mesa
4. Implement a toll for crossing at Otay Mesa
5. Increase the number of tours crossing

6.2 SENSITIVITY TEST 1

The first test is to increase the capacity at Otay Mesa. The number of vehicle lanes is doubled at Otay Mesa from 13 to 26 lanes. The expected result is that the wait times will decrease and shift tour modes from walk tour to drive modes due to the increased capacity and reduced auto wait times.

The results of this test showed that there is a shift in POE choice from San Ysidro to Otay Mesa shown in Figure 49. The share of tours crossing at Otay Mesa increased by 11% and the share of tours crossing at San Ysidro decreased by 4.3% as shown in Figure 50. This result makes sense, as the capacity at Otay Mesa increases, we see more people choosing to fill that capacity. The wait times have become shorter, and the volumes have stabilized across the POE to account for the new capacity.

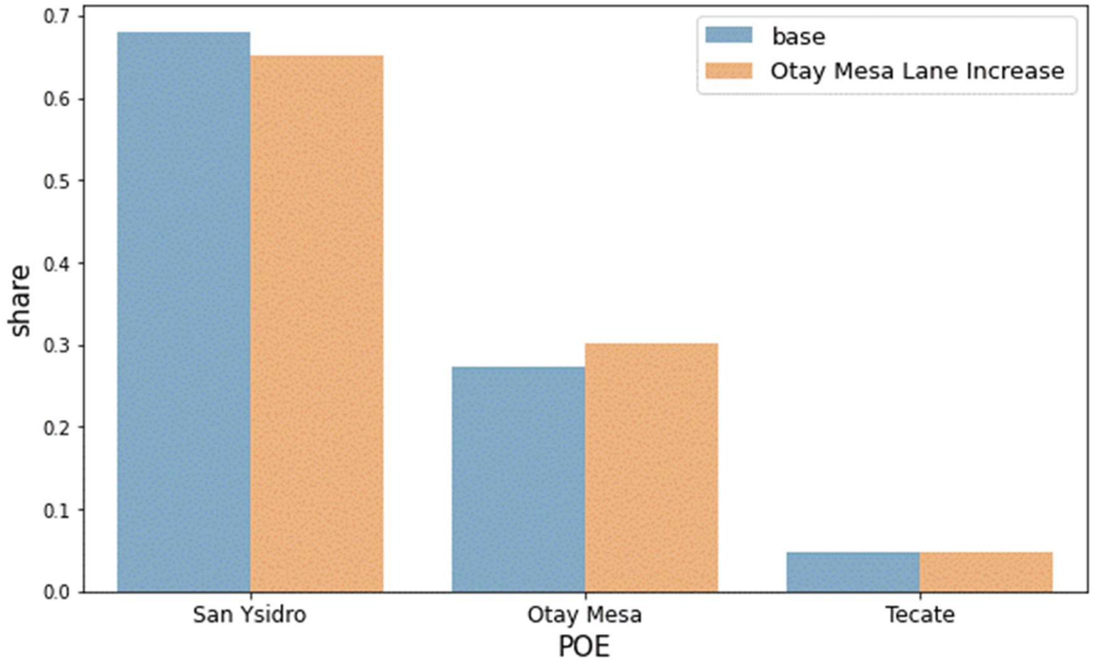


FIGURE 49: SENSITIVITY TEST 1 POE CHOICE

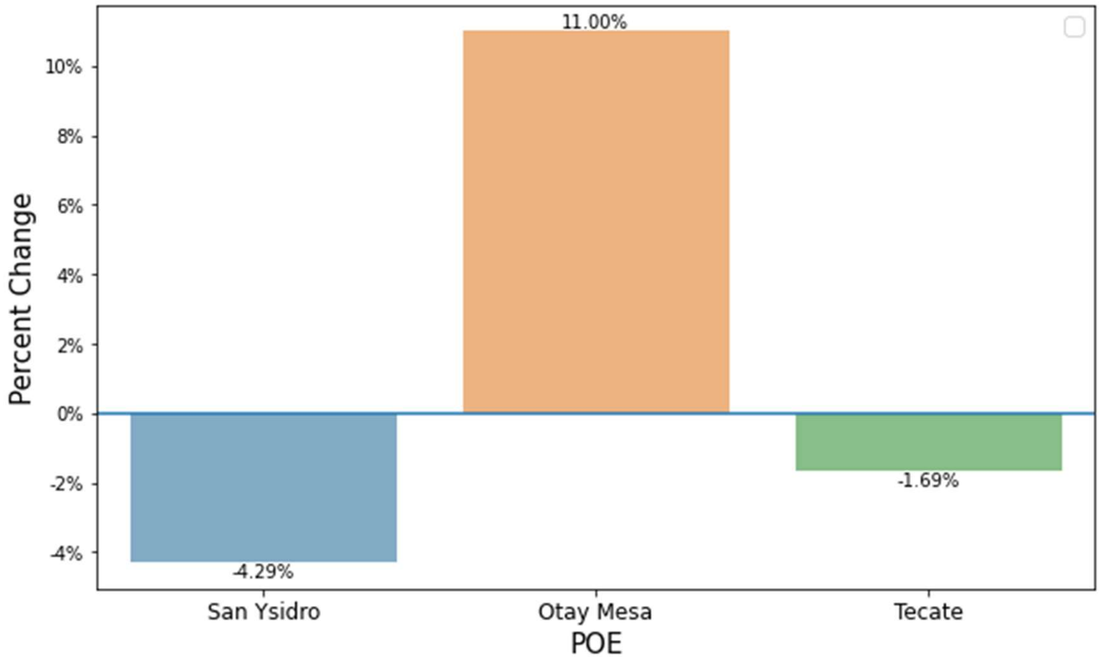


FIGURE 50: SENSITIVITY TEST 1 CHANGE IN POE CHOICE

This test also shows a shift in mode at Otay Mesa shown in Figure 51. As expected, the increase in vehicle lanes at Otay Mesa causes a mode shift from walk to drive tour. The share of walk tours at Otay Mesa decreases by 21% while the share of drive alone, shared 2, and shared 3 tours increases by 7%, 13%, and 15% respectively as shown in Figure 52.

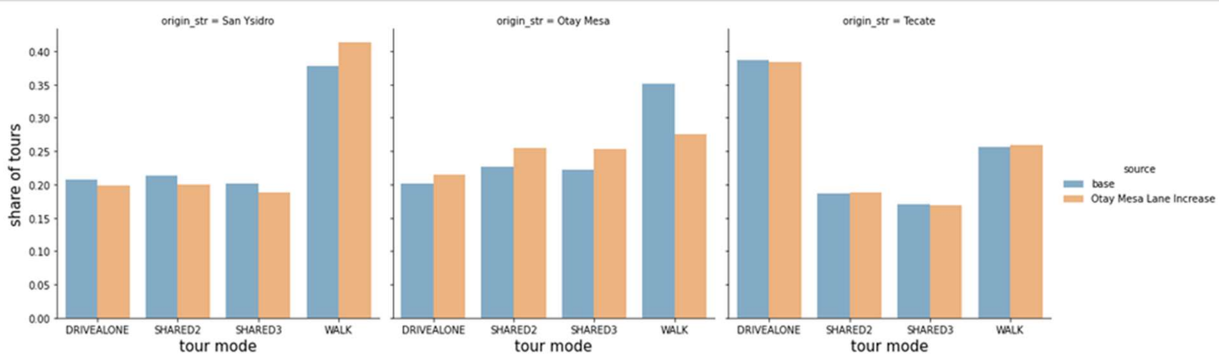


FIGURE 51: SENSITIVITY TEST 1 MODE CHOICE BY POE

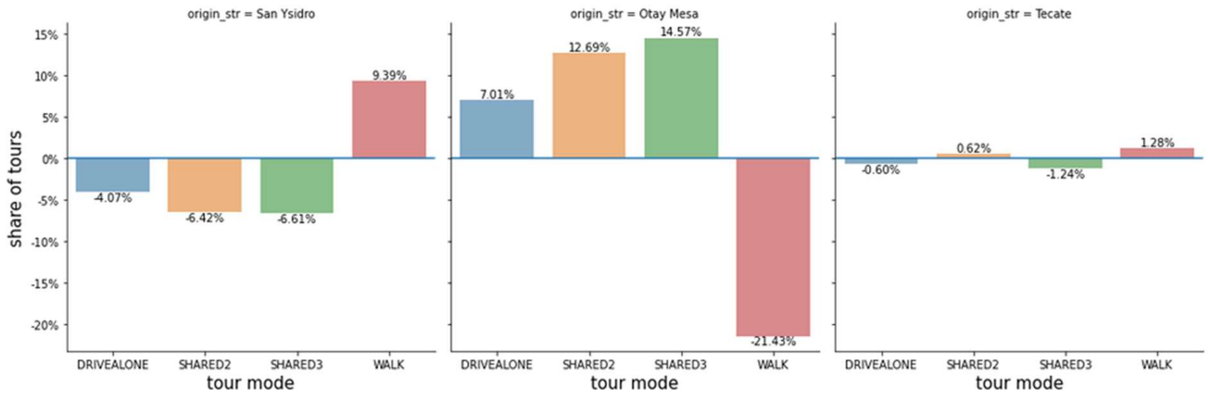


FIGURE 52: SENSITIVITY TEST 1 CHANGE IN MODE CHOICE BY POE

This sensitivity test shows that the model is responsive to POE capacity.

6.3 SENSITIVITY TEST 2

The second sensitivity test is to decrease the headways of transit service near San Ysidro. This is expected to cause a mode shift from auto to walk as the options for transit across the border become better. This is done by manually editing the transit skims such that the first wait of any Transit Access Point (TAP)[11] pair that started or ended at the San Ysidro transit station is halved.

This test showed an overall shift in tour mode choice at San Ysidro shown in Figure 53. The share of walking tours at San Ysidro increased by 2.5%, while auto tours each decreased by about .3-2.9% as shown in Figure 54. This is expected as the transit options have become better at San Ysidro, it's more attractive to walk across the border in a pedestrian lane with a shorter wait time and then take transit.

[11] Transit Access Points, or TAPs, are used in the SANDAG model instead of TAZs. They represent the initial boarding and final alighting stop of a transit trip.

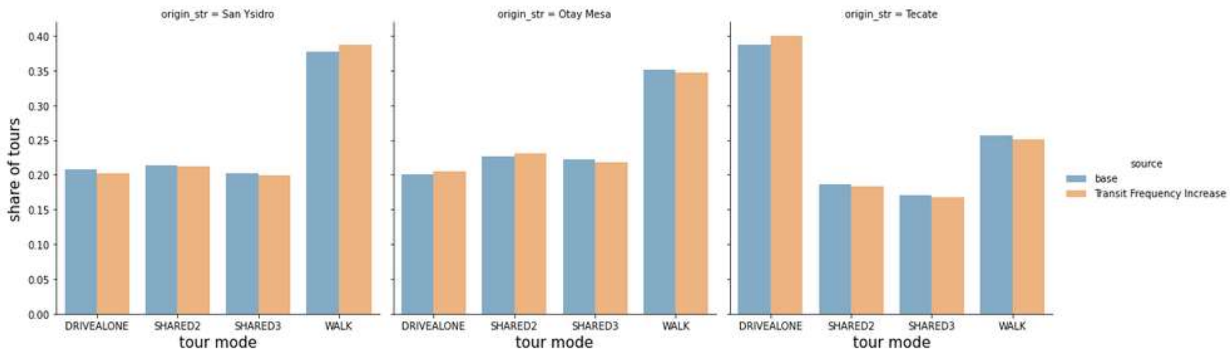


FIGURE 53: SENSITIVITY TEST 2 MODE CHOICE BY POE

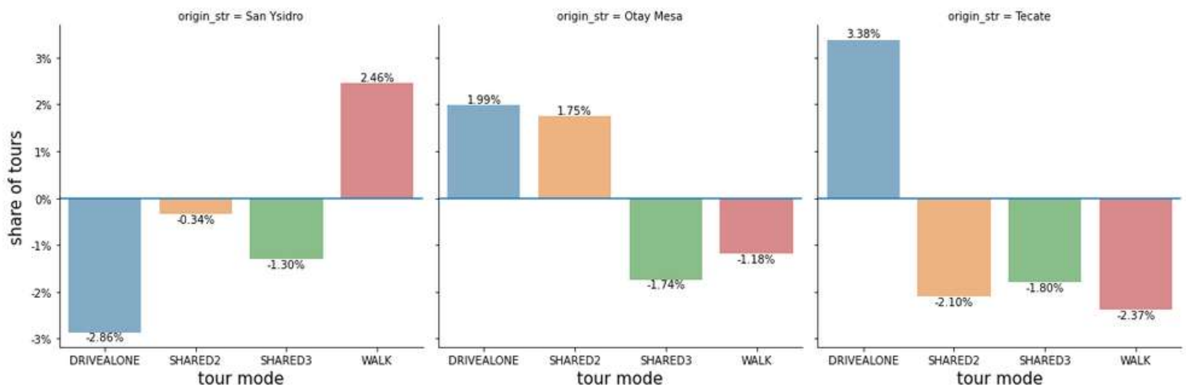


FIGURE 54: SENSITIVITY TEST 2 CHANGE IN MODE CHOICE BY POE

This test also shows a shift in trip mode choice shown in Figure 55. The share of walk to transit trips increased by 3.7% as shown in Figure 56. This is expected, and shows that the walk tours are utilizing the improved transit network.

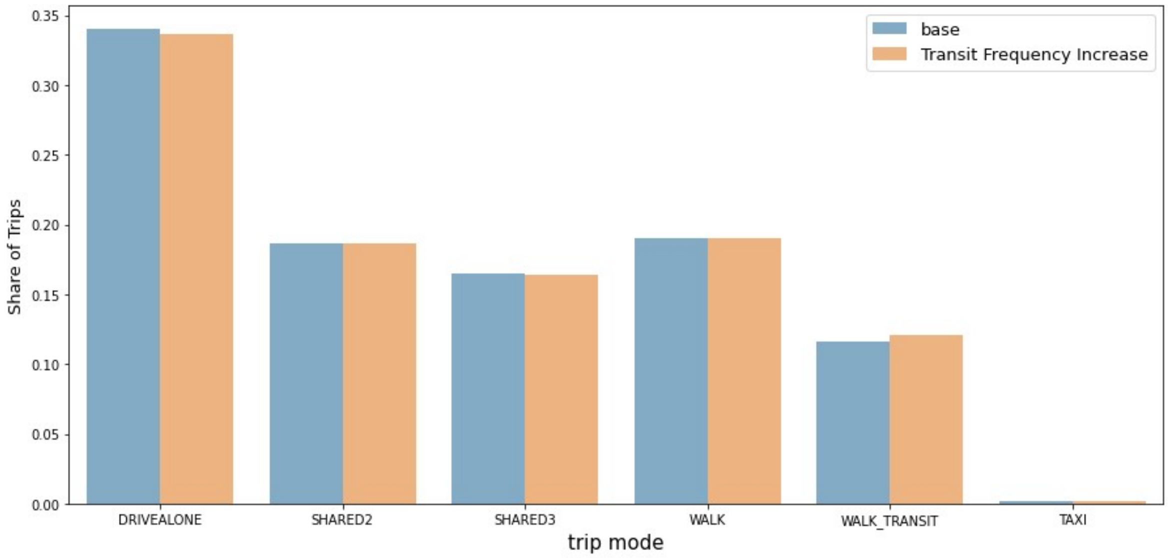


FIGURE 55: SENSITIVITY TEST 2 TRIP MODE CHOICE

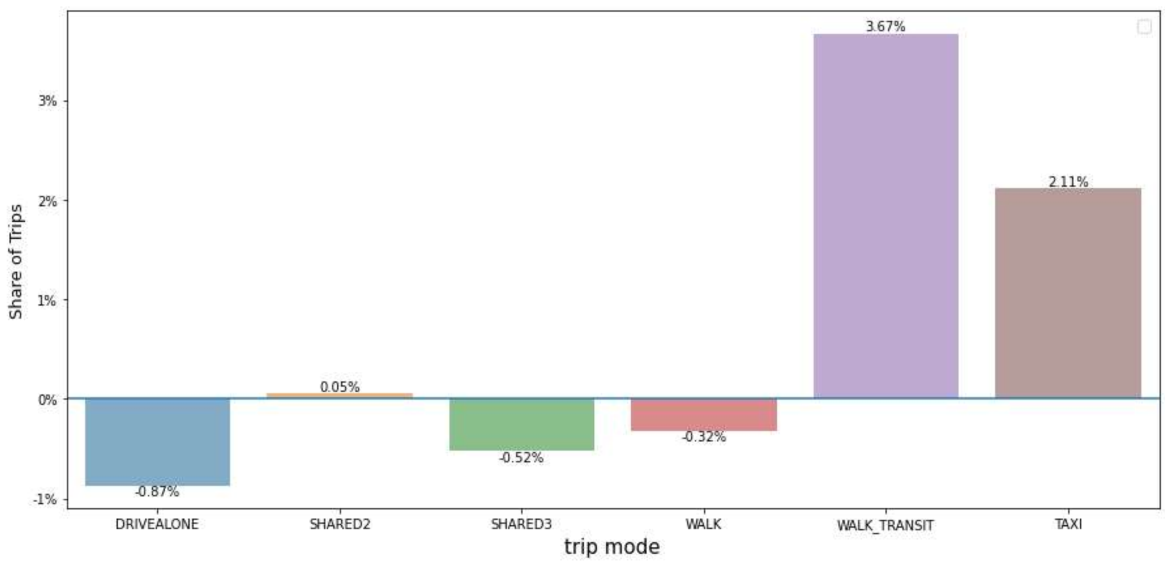


FIGURE 56: SENSITIVITY TEST 2 CHANGE IN TRIP MODE CHOICE

This test shows that the model is sensitive to changes in the transit network and improving the transit accessibility, tours will shift from auto modes to use it.

6.4 SENSITIVITY TEST 3

The third sensitivity test is to increase the employment around Otay Mesa. This should change the primary destination of tours and the POE choice. The employment is increased in 5 MAZs near Otay Mesa, consistent with 2025 SR11 project plans. The total employment is increased by ~4,000 employees in the project MAZs shown in Figure 57. Table 25 shows the breakdown of employment categories for the new employment added for this sensitivity test.

FIGURE 57: SENSITIVITY TEST 3 PROJECT MAZS

TABLE 25: SENSITIVITY TEST 3 EMPLOYMENT CATEGORIES

EMPLOYMENT CATEGORY	EMPLOYEES
Retail	90
Wholesale and Manufacturing	3,335
Africulture	95
Office & Professional Services	466
Total	**3,986**

There is a small shift in POE choice resulting from this increase in employment shown in Figure 58. The share of tours crossing at Otay Mesa increased by ~1% as shown in Figure 59.

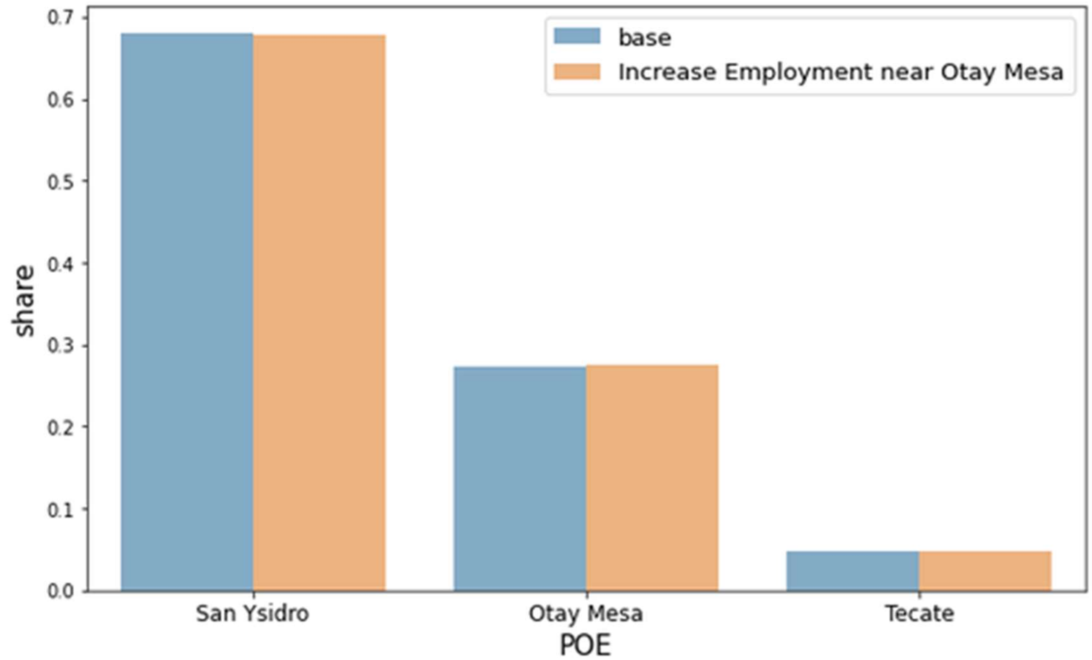


FIGURE 58: SENSITIVITY TEST 3 POE CHOICE

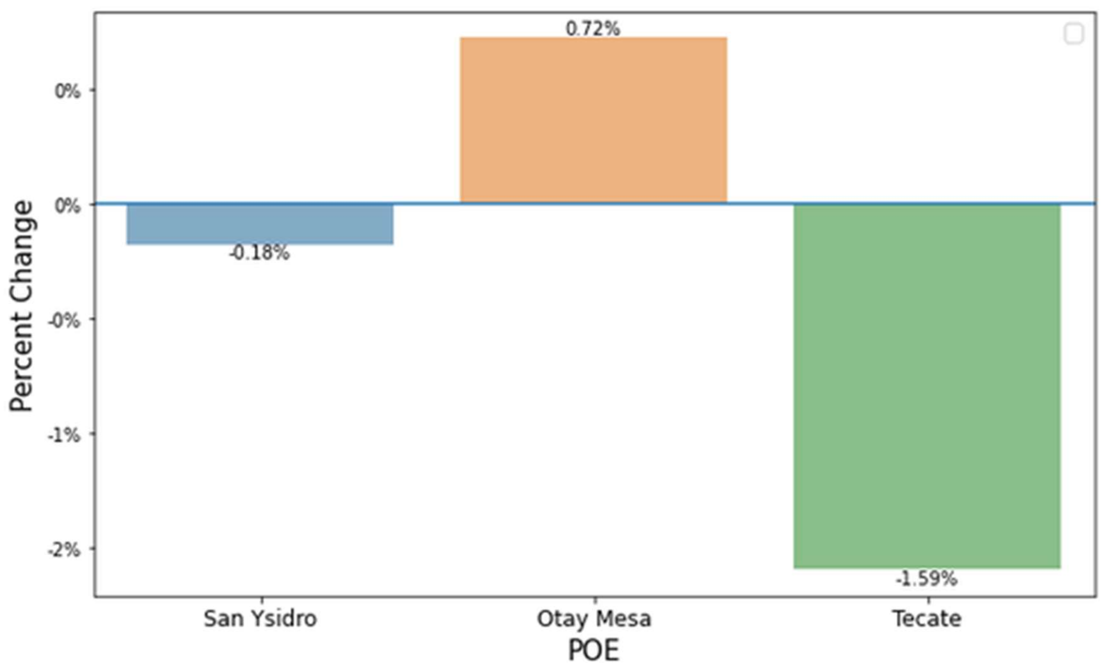


FIGURE 59: SENSITIVITY TEST 3 CHANGE IN POE CHOICE

This test shows that there is an increase in tours and trips to or from the project area. Table 26 shows there are no tours to the project area in the base scenario and the added employment attracted 502 tours to the project. Table 27 shows that the base scenario had only two trips to the project, and the added employment attracted 819 trips of mostly work and shop purposes to the project.

TABLE 26: SENSITIVITY TEST 3 TOURS TO PROJECT AREA

	BASE	BUILD
Tours to Project Area	-	502
Tours Outside the Project Area	113,995	113,493
Total Tours	**113,995**	**113,995**
Percent tours to project area	-	0.44%

TABLE 27: SENSITIVITY TEST 3 TRIPS BY PURPOSE TO PROJECT AREA

	NUMBER OF TRIPS TO/FROM PROJECT	
Trip Purpose	Base	Build
Work	-	363
Shop	-	230
Visit	-	33
Other	2	193
Total	**2**	**819**

This sensitivity test shows that the model is sensitive to changes in land use and that the model estimates that some portion of the new jobs added near the border will be filled by residents of Mexico.

6.5 SENSITIVITY TEST 4

The fourth sensitivity test is to implement a toll at Otay Mesa. This should shift the mode choice at Otay Mesa from auto to walk and possibly decrease overall tours through Otay Mesa. A $5 toll is implemented at all vehicle lanes at Otay Mesa regardless of pass type. Based on the value of time used in the model, a $5 toll is equivalent to a 1 hour wait.

This test showed a large decrease in tours crossing at Otay Mesa shown in Figure 60. The share of tours crossing at Otay Mesa dropped by 35% shown in Figure 61. This is expected as travelers will shift their POE choice to avoid the extra toll.

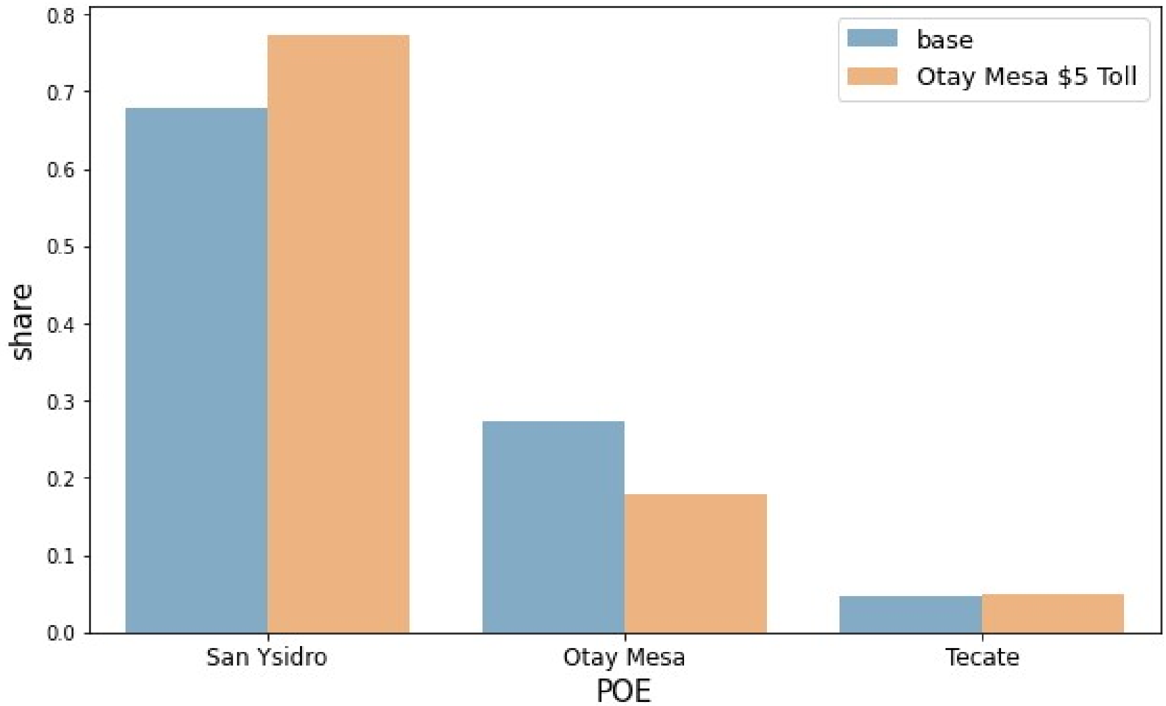


FIGURE 60: SENSITIVITY TEST 4 POE CHOICE

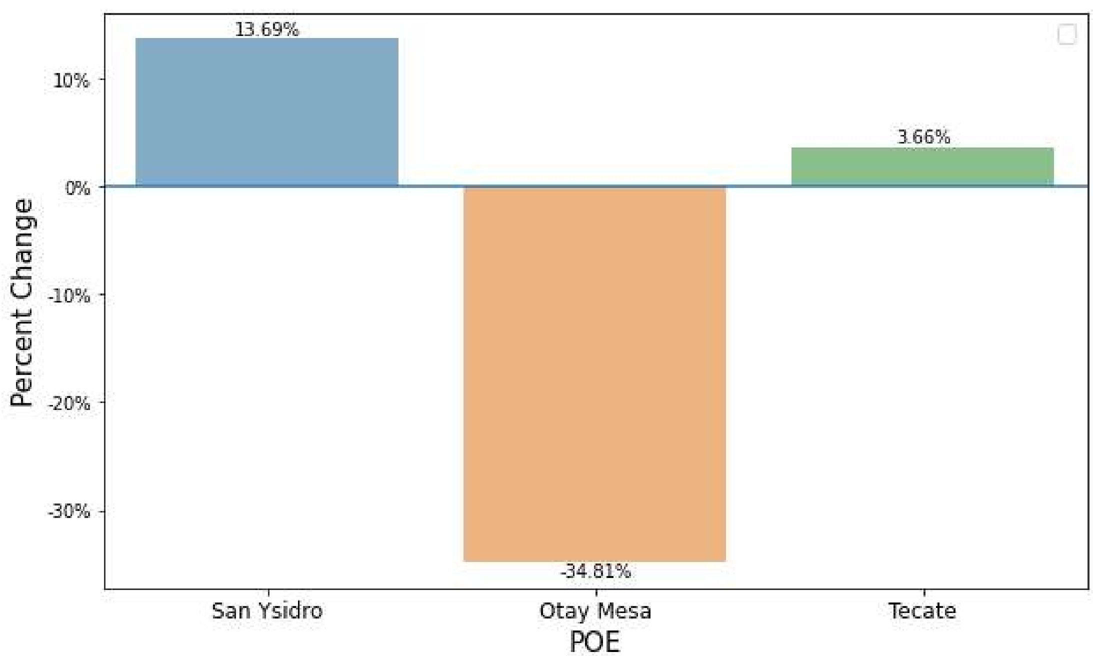


FIGURE 61: SENSITIVITY TEST 4 CHANGE IN POE CHOICE

The tour mode at Otay Mesa showed a large shift from auto to walk as shown in Figure 62. The share of walk tours at Otay Mesa jumped by 115% as shown in Figure 63. This is expected, as travelers will choose the mode which does not incur an extra cost.

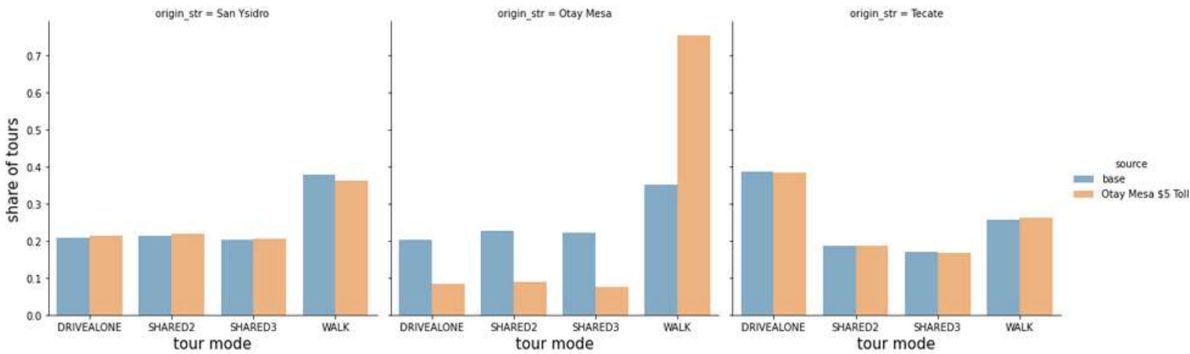


FIGURE 62: SENSITIVITY TEST 4 TOUR MODE BY POE

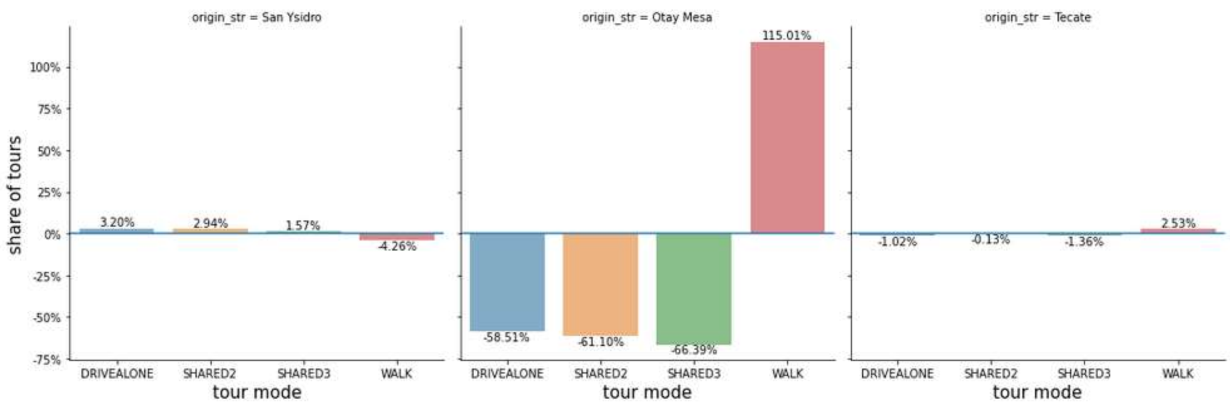


FIGURE 63: SENSITIVITY TEST 4 CHANGE IN TOUR MODE BY POE

This sensitivity test shows the model is highly sensitive to pricing.

6.6 SENSITIVITY TEST 5

The fifth sensitivity test is to increase the total border crossing volumes. This scenario is important for future year scenarios where the number of tours crossing the border is expected to increase. Without any infrastructure changes, increasing the crossing volume is expected to increase the crossing wait times and thus shift mode choice from auto to walk mode. This test is done by doubling the daily crossing volume from 113,995 to 227,990 tours.

The increase in crossing volume caused an overall mode shift from auto to walk as shown in Figure 64. The share of walk tours increased by 16%, while the share of drive alone, shared 2, and shared 3 tours decreased by 8%, 10%, and 11% respectively as shown in Figure 65. This is expected as the wait times in the vehicle lanes increase due to the increase in crossing volumes.

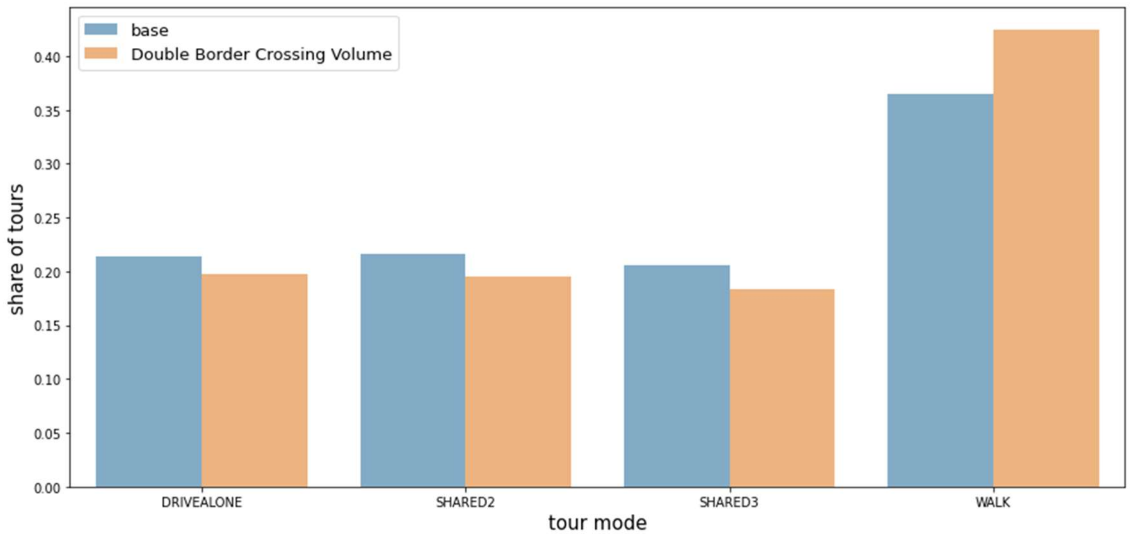


FIGURE 64: SENSITIVITY TEST 5 TOUR MODE CHOICE

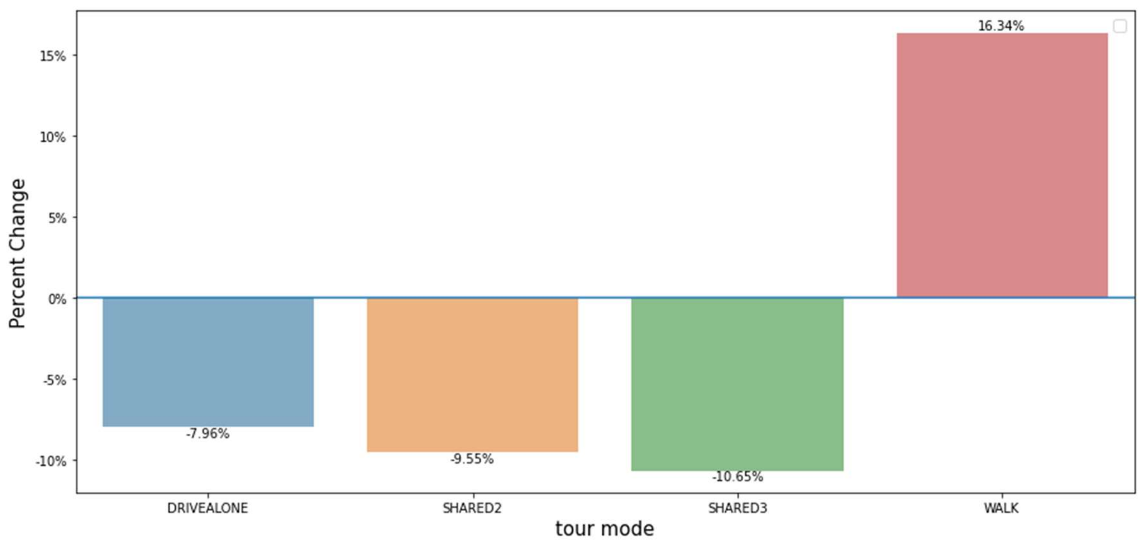


FIGURE 65: SENSITIVITY TEST 5 CHANGE IN TOUR MODE CHOICE

This sensitivity test shows that the model is sensitive to changes in border crossing volumes.

7.0 CONCLUSION

In this report, we have presented the data used, models designed, estimated, and implemented as well as the validations and sensitivity tests conducted in the SANDAG Cross Border Model update. The goal of this project is to update the border crossing and resident travel model to reflect the most recent (2019) cross border survey data, as well as create a new border crossing wait time model that can be used to model changes in number of lanes and total demand. We deployed this model using the ActivitySIm framework. This is the first 'special market' model to be implemented within the ActivitySim framework. Runtime of this new model has been improved over the previous iteration of the Cross-Border Model. The model calibration revealed that the model had a good fit to the observed data. The sensitivity test results also show reasonable results, aligned to what one would expect for the various changes implemented. The model has been handed off to SANDAG staff, who are now running the model in house.

ACKNOWLEDGEMENT

This project is funded by a Caltrans Planning Grant. The project team greatly appreciates the Caltrans funding support. Kate Deutsch-Burgner of Data Perspectives Consulting provided editorial assistance with this report.

8.0 APPENDIX

Stop Frequency Distribution

TABLE 28: STOP FREQUENCY DISTRIBUTION

PURPOSE	DURATIONLO	DURATIONHI	OUTBOUND	INBOUND	PERCENT
work	0	4	0	0	0.354003353
work	0	4	0	1	0.118589503
work	0	4	0	2	0.238526452
work	0	4	0	3	0.095724807
work	0	4	1	0	0.032904022
work	0	4	1	1	0.015264207
work	0	4	1	2	0.041058151
work	0	4	1	3	0
work	0	4	2	0	0.046983846
work	0	4	2	1	0.018775127
work	0	4	2	2	0
work	0	4	2	3	0
work	0	4	3	0	0.030173
work	0	4	3	1	0.007997532
work	0	4	3	2	0
work	0	4	3	3	0
work	4	8	0	0	0.358953731
work	4	8	0	1	0.098461968
work	4	8	0	2	0.064925378
work	4	8	0	3	0.141517964
work	4	8	1	0	0.073681676
work	4	8	1	1	0.03081633
work	4	8	1	2	0.011623154
work	4	8	1	3	0.05092802
work	4	8	2	0	0.075629879
work	4	8	2	1	0.017721792
work	4	8	2	2	0.029161393
work	4	8	2	3	0.005979984
work	4	8	3	0	0.005811577
work	4	8	3	1	0.028975577

work	4	8	3	2	0.005811577
work	4	8	3	3	0
work	8	24	0	0	0.281400799
work	8	24	0	1	0.069666637
work	8	24	0	2	0.070407877
work	8	24	0	3	0.08130446
work	8	24	1	0	0.095132038
work	8	24	1	1	0.088477323
work	8	24	1	2	0.017705966
work	8	24	1	3	0.038071052
work	8	24	2	0	0.080581126
work	8	24	2	1	0.031830285
work	8	24	2	2	0.027653863
work	8	24	2	3	0.003218971
work	8	24	3	0	0.062021881
work	8	24	3	1	0.033906556
work	8	24	3	2	0.008335158
work	8	24	3	3	0.010286008
school	0	4	0	0	0.409105813
school	0	4	0	1	0.327913769
school	0	4	0	2	0
school	0	4	0	3	0
school	0	4	1	0	0.153482472
school	0	4	1	1	0.109497945
school	0	4	1	2	0
school	0	4	1	3	0
school	0	4	2	0	0
school	0	4	2	1	0
school	0	4	2	2	0
school	0	4	2	3	0
school	0	4	3	0	0
school	0	4	3	1	0
school	0	4	3	2	0
school	0	4	3	3	0
school	4	8	0	0	0.128094015
school	4	8	0	1	0.195366472
school	4	8	0	2	0.11400614

school	4	8	0	3	0.119162558
school	4	8	1	0	0.027425159
school	4	8	1	1	0.059895613
school	4	8	1	2	0.020347213
school	4	8	1	3	0.059895613
school	4	8	2	0	0.028367844
school	4	8	2	1	0
school	4	8	2	2	0.037880645
school	4	8	2	3	0
school	4	8	3	0	0.152950922
school	4	8	3	1	0.056607806
school	4	8	3	2	0
school	4	8	3	3	0
school	8	24	0	0	0.131899769
school	8	24	0	1	0.04725762
school	8	24	0	2	0.078568681
school	8	24	0	3	0.468493428
school	8	24	1	0	0
school	8	24	1	1	0
school	8	24	1	2	0
school	8	24	1	3	0.040050879
school	8	24	2	0	0
school	8	24	2	1	0.040050879
school	8	24	2	2	0
school	8	24	2	3	0
school	8	24	3	0	0.069563751
school	8	24	3	1	0.124114993
school	8	24	3	2	0
school	8	24	3	3	0
shop	0	4	0	0	0.205091194
shop	0	4	0	1	0.15822808
shop	0	4	0	2	0.20267432
shop	0	4	0	3	0.105460365
shop	0	4	1	0	0.071542129
shop	0	4	1	1	0.093503144
shop	0	4	1	2	0.045642428
shop	0	4	1	3	0.012337861

shop	0	4	2	0	0.04571784
shop	0	4	2	1	0.024145894
shop	0	4	2	2	0.007636373
shop	0	4	2	3	0.00182955
shop	0	4	3	0	0.018790659
shop	0	4	3	1	0.004852946
shop	0	4	3	2	0.002547217
shop	0	4	3	3	0
shop	4	8	0	0	0.107405621
shop	4	8	0	1	0.11626285
shop	4	8	0	2	0.134552278
shop	4	8	0	3	0.228482531
shop	4	8	1	0	0.032274088
shop	4	8	1	1	0.059274338
shop	4	8	1	2	0.085288909
shop	4	8	1	3	0.081589275
shop	4	8	2	0	0.04994866
shop	4	8	2	1	0.037522738
shop	4	8	2	2	0.017979199
shop	4	8	2	3	0.012131798
shop	4	8	3	0	0.026404749
shop	4	8	3	1	0.005362097
shop	4	8	3	2	0.002588036
shop	4	8	3	3	0.002932832
shop	8	24	0	0	0.061242337
shop	8	24	0	1	0.044418589
shop	8	24	0	2	0.097112444
shop	8	24	0	3	0.326771739
shop	8	24	1	0	0.037085657
shop	8	24	1	1	0.055080115
shop	8	24	1	2	0.058306791
shop	8	24	1	3	0.142525834
shop	8	24	2	0	0.018116341
shop	8	24	2	1	0.026241192
shop	8	24	2	2	0.044618324
shop	8	24	2	3	0.011496019
shop	8	24	3	0	0.037564441

shop	8	24	3	1	0.014951856
shop	8	24	3	2	0.006311058
shop	8	24	3	3	0.018157264
visit	0	4	0	0	0.639330063
visit	0	4	0	1	0.268930645
visit	0	4	0	2	0
visit	0	4	0	3	0
visit	0	4	1	0	0
visit	0	4	1	1	0
visit	0	4	1	2	0
visit	0	4	1	3	0
visit	0	4	2	0	0.091739292
visit	0	4	2	1	0
visit	0	4	2	2	0
visit	0	4	2	3	0
visit	0	4	3	0	0
visit	0	4	3	1	0
visit	0	4	3	2	0
visit	0	4	3	3	0
visit	4	8	0	0	0.179504382
visit	4	8	0	1	0.093077962
visit	4	8	0	2	0.028053068
visit	4	8	0	3	0.146246988
visit	4	8	1	0	0.138119976
visit	4	8	1	1	0.095631673
visit	4	8	1	2	0.011453481
visit	4	8	1	3	0
visit	4	8	2	0	0
visit	4	8	2	1	0.129133628
visit	4	8	2	2	0
visit	4	8	2	3	0
visit	4	8	3	0	0.178778843
visit	4	8	3	1	0
visit	4	8	3	2	0
visit	4	8	3	3	0
visit	8	24	0	0	0.586261862
visit	8	24	0	1	0.047726627

visit	8	24	0	2	0.098603567
visit	8	24	0	3	0.026071158
visit	8	24	1	0	0.067065071
visit	8	24	1	1	0.081878389
visit	8	24	1	2	0
visit	8	24	1	3	0
visit	8	24	2	0	0
visit	8	24	2	1	0.020699081
visit	8	24	2	2	0
visit	8	24	2	3	0
visit	8	24	3	0	0.042694275
visit	8	24	3	1	0
visit	8	24	3	2	0.02899997
visit	8	24	3	3	0
other	0	4	0	0	0.562285413
other	0	4	0	1	0.197379987
other	0	4	0	2	0.061889215
other	0	4	0	3	0.029500835
other	0	4	1	0	0.100103837
other	0	4	1	1	0.00830746
other	0	4	1	2	0.004980763
other	0	4	1	3	0.010808858
other	0	4	2	0	0.024743633
other	0	4	2	1	0
other	0	4	2	2	0
other	0	4	2	3	0
other	0	4	3	0	0
other	0	4	3	1	0
other	0	4	3	2	0
other	0	4	3	3	0
other	4	8	0	0	0.439351238
other	4	8	0	1	0.23945896
other	4	8	0	2	0.050133082
other	4	8	0	3	0.148180425
other	4	8	1	0	0.009824806
other	4	8	1	1	0.064441866
other	4	8	1	2	0

other	4	8	1	3	0.02919389
other	4	8	2	0	0.013502739
other	4	8	2	1	0
other	4	8	2	2	0
other	4	8	2	3	0
other	4	8	3	0	0.005912995
other	4	8	3	1	0
other	4	8	3	2	0
other	4	8	3	3	0
other	8	24	0	0	0.646659489
other	8	24	0	1	0.057291926
other	8	24	0	2	0.056909015
other	8	24	0	3	0
other	8	24	1	0	0.017464522
other	8	24	1	1	0.022782695
other	8	24	1	2	0.057484387
other	8	24	1	3	0
other	8	24	2	0	0.066138738
other	8	24	2	1	0.058036759
other	8	24	2	2	0
other	8	24	2	3	0
other	8	24	3	0	0.017232469
other	8	24	3	1	0
other	8	24	3	2	0
other	8	24	3	3	0

Stop Purpose Distribution

TABLE 29: STOP PURPOSE DISTRIBUTION

TOURPURP	INBOUND	STOPNUM	MULTIPLE	STOPPURP WORK	STOPPURP SCHOOL	STOPPURP SHOP	STOPPURP VISIT	STOPPURP OTHER
work	0	1	0	0.05344	0.076938	0.254353	0.021591	0.593678
		1	1	0.038117	0.062798	0.392458	0.055513	0.451113
		2	1	0.059612	0.025427	0.300058	0.090512	0.524391
		3	1	0.117583	0.032863	0.371829	0.04861	0.429115
	1	1	0	0.129682	0	0.31624	0.082712	0.471366
		1	1	0.228763	0.022011	0.29491	0.0442	0.410116

		2	1	0.256711	0.013777	0.345805	0.029283	0.354424
		3	1	0.188082	0.021373	0.334748	0.120509	0.335288
school	0	1	0	0	0	0.136145	0.149024	0.714831
		1	1	0	0	0.394676	0.036447	0.568877
		2	1	0	0	0.642248	0	0.357752
		3	1	0	0	0.488898	0	0.511102
	1	1	0	0	0	0.571066	0.124658	0.304276
		1	1	0	0.109461	0.408143	0.257706	0.22469
		2	1	0	0.149145	0.508879	0.117253	0.224723
		3	1	0	0.16961	0.285576	0.058764	0.48605
shop	0	1	0	0	0	0.069724	0.11923	0.811046
		1	1	0	0	0.22836	0.059967	0.711673
		2	1	0	0	0.292828	0.085303	0.621869
		3	1	0	0	0.43907	0.110565	0.450364
	1	1	0	0	0	0.702843	0.043899	0.253257
		1	1	0	0	0.71921	0.041408	0.239382
		2	1	0	0	0.686989	0.042206	0.270805
		3	1	0	0	0.718043	0.04994	0.232017
visit	0	1	0	0	0	0.100671	0.056709	0.84262
		1	1	0	0	0.214595	0	0.785405
		2	1	0	0	0.363066	0.108972	0.527962
		3	1	0	0	0.332055	0	0.667945
	1	1	0	0	0	0.34687	0.192934	0.460196
		1	1	0	0	0.119161	0.397293	0.483547
		2	1	0	0	0.119161	0.261393	0.619446
		3	1	0	0	0	0	1
other	0	1	0	0	0	0.387091	0.034835	0.578074
		1	1	0	0	0.59424	0.204095	0.201665
		2	1	0	0	0.326112	0	0.673888
		3	1	0	0	0.496656	0	0.503344
	1	1	0	0	0	0.333342	0.110206	0.556452
		1	1	0	0	0.171466	0	0.828534
		2	1	0	0	0.213663	0.044675	0.741661
		3	1	0	0	0.086657	0.043329	0.870014

Outbound Stop Duration

TABLE 30: OUTBOUND STOP DURATION

REMAININGLOW	REMAININGHIGH	STOP	0	1	2	3	4	5	6	7	8	9	10	11
0	0	1	1.000	-	-	-	-	-	-	-	-	-	-	-
0	0	2	1.000	-	-	-	-	-	-	-	-	-	-	-
0	0	3	1.000	-	-	-	-	-	-	-	-	-	-	-
1	1	1	-	-	-	-	-	-	-	-	-	-	-	-
1	1	2	1.000	-	-	-	-	-	-	-	-	-	-	-
1	1	3	-	-	-	-	-	-	-	-	-	-	-	-
2	2	1	0.143	0.857	-	-	-	-	-	-	-	-	-	-
2	2	2	0.445	0.555	-	-	-	-	-	-	-	-	-	-
2	2	3	-	-	-	-	-	-	-	-	-	-	-	-
3	3	1	0.044	0.696	0.259	-	-	-	-	-	-	-	-	-
3	3	2	0.575	0.347	0.078	-	-	-	-	-	-	-	-	-
3	3	3	0.732	0.156	0.112	-	-	-	-	-	-	-	-	-
4	4	1	0.083	0.384	0.452	0.081	-	-	-	-	-	-	-	-
4	4	2	0.147	0.817	0.037	-	-	-	-	-	-	-	-	-
4	4	3	0.060	0.588	0.352	-	-	-	-	-	-	-	-	-
5	5	1	0.183	0.435	0.279	0.103	-	-	-	-	-	-	-	-
5	5	2	0.048	0.528	0.424	-	-	-	-	-	-	-	-	-
5	5	3	-	0.725	0.147	0.128	-	-	-	-	-	-	-	-
6	6	1	-	0.407	0.447	0.095	0.051	-	-	-	-	-	-	-
6	6	2	0.112	0.472	0.301	0.091	0.023	-	-	-	-	-	-	-
6	6	3	-	0.680	0.320	-	-	-	-	-	-	-	-	-
7	7	1	0.038	0.285	0.271	0.214	0.148	0.022	-	0.023	-	-	-	-
7	7	2	0.085	0.305	0.187	0.423	-	-	-	-	-	-	-	-
7	7	3	0.206	0.321	0.092	0.381	-	-	-	-	-	-	-	-
8	8	1	-	0.240	0.464	0.142	0.144	0.009	-	-	-	-	-	-
8	8	2	0.052	0.337	0.277	0.306	0.027	-	-	-	-	-	-	-

8	8	3	-	0.503	0.383	0.114	-	-	-	-	-	-	-	-
9	9	1	0.029	0.198	0.324	0.265	0.124	0.060	-	-	-	-	-	-
9	9	2	-	0.240	0.527	0.121	0.058	0.054	-	-	-	-	-	-
9	9	3	0.151	0.103	0.305	0.357	0.083	-	-	-	-	-	-	-
10	10	1	-	0.120	0.466	0.113	0.192	0.019	0.041	0.050	-	-	-	-
10	10	2	0.130	0.185	0.397	-	0.288	-	-	-	-	-	-	-
10	10	3	0.088	0.274	0.285	-	0.178	0.175	-	-	-	-	-	-
11	39	1	0.027	0.232	0.242	0.189	0.116	0.074	0.033	0.055	0.013	0.010	0.008	0.002
11	39	2	0.070	0.309	0.302	0.134	0.095	0.008	0.012	0.042	0.025	-	-	0.004
11	39	3	0.052	0.382	0.293	0.186	0.056	0.012	0.019	-	-	-	-	-

Inbound Stop Duration

TABLE 31: INBOUND STOP DURATION

REMAININGLOW	REMAININGHIGH	STOP	0	-1	-2	-3	-4	-5	-6	-7
0	0	1	1.000	-	-	-	-	-	-	-
0	0	2	1.000	-	-	-	-	-	-	-
0	0	3	1.000	-	-	-	-	-	-	-
1	1	1	0.406	0.594	-	-	-	-	-	-
1	1	2	0.362	0.638	-	-	-	-	-	-
1	1	3	0.422	0.578	-	-	-	-	-	-
2	2	1	-	0.748	0.252	-	-	-	-	-
2	2	2	0.110	0.701	0.189	-	-	-	-	-
2	2	3	0.116	0.717	0.167	-	-	-	-	-
3	3	1	0.006	0.320	0.640	0.033	-	-	-	-
3	3	2	0.049	0.353	0.501	0.098	-	-	-	-
3	3	3	0.004	0.244	0.591	0.161	-	-	-	-
4	4	1	-	0.128	0.409	0.428	0.035	-	-	-
4	4	2	0.016	0.087	0.401	0.407	0.088	-	-	-

4	4	3	-	0.095	0.380	0.453	0.073	-	-	-	
5	5	1	0.008	0.049	0.212	0.420	0.258	0.052	-	-	
5	5	2	-	0.057	0.229	0.352	0.321	0.042	-	-	
5	5	3	-	0.068	0.102	0.476	0.315	0.039	-	-	
6	6	1	0.019	0.024	0.080	0.287	0.303	0.233	0.054	-	
6	6	2	-	0.053	0.127	0.136	0.374	0.264	0.046	-	
6	6	3	-	0.045	0.156	0.193	0.298	0.249	0.059	-	
7	7	1	-	0.033	0.019	0.183	0.264	0.211	0.262	0.028	
7	7	2	-	0.014	0.074	0.105	0.262	0.343	0.192	0.009	
7	7	3	-	-	0.080	0.124	0.222	0.334	0.194	0.046	
8	8	1	-	0.017	0.068	0.124	0.139	0.185	0.216	0.251	
8	8	2	-	0.003	-	0.034	0.068	0.321	0.253	0.321	
8	8	3	0.020	0.030	0.042	0.069	0.068	0.241	0.334	0.197	
9	9	1	0.006	0.015	0.010	0.056	0.084	0.155	0.222	0.452	
9	9	2	-	0.007	0.019	0.013	0.044	0.151	0.100	0.665	
9	9	3	-	-	-	0.029	0.135	0.042	0.240	0.554	
10	10	1	-	0.011	-	0.063	0.056	0.107	0.150	0.614	
10	10	2	-	-	0.020	0.036	0.011	0.074	0.142	0.717	
10	10	3	-	0.051	-	0.035	0.061	0.007	0.316	0.529	
11	39	1	0.006	0.040	0.045	0.041	0.036	0.021	0.046	0.764	
11	39	2	-	0.018	-	0.007	0.021	0.068	0.067	0.818	
11	39	3	-	0.022	0.022	0.029	0.159	0.030	0.048	0.689	

the science of insight