FINAL KERN COG VMIP 2
Model Development Report

Prepared for:
Kern Council of Governments

April 2017

WC15-3242

FEHR ♦ PEERS

Table of Contents

OVERVIEW

The San Joaquin Valley Model Improvement Plan (VMIP 1) began in 2010 and resulted in substantial enhancements to the modeling capabilities of the Metropolitan Planning Organizations (MPOs) within the San Joaquin Valley (SJV). Due to the timing of the original VMIP 1, many data sources pertinent to understanding travel behavior and developing travel forecasting models were not available. As such, older sources were used to supplement data for the base year, making calibration and validation difficult due to the economic downturn relative to the 2001/2003 California Household Travel Survey (CHTS) and 2000 Census which were collected before calibration efforts commenced. VMIP 2 not only takes advantage of the most recent Census and CHTS data and the model structure enhancements developed as part of the VMIP 1, but also new Big Data.

This document provides guidance on the model specifications and data used in developing the components for the San Joaquin Valley Model Improvement Plan, Phase 2 (VMIP 2). The objective of this document is to provide an overview and full technical details of the VMIP 2 models: this includes aspects common to all VMIP 2 models as well as specific calibration and model validation for the Kern Council of Governments (Kern COG) model. Changes between the original VMIP 1 models and the VMIP 2 models receive special emphasis.

In addition to the updated data, VMIP 2 implemented changes to the model structure are based on feedback from the Air Resources Board (ARB) provided during the Regional Transportation Plan (RTP) and Sustainable Communities Strategy (SCS) review process, and MPO staff who applied the models over the last several years. Key enhancements to model sensitivity and usability include:

- Land Use: Simplified residential and employment categories and addition of group quarters population

- Socio-economic: Employee salary and household income relationship for home-work trips

- Inter-regional Travel: Improved control over scenario evaluation of inter-regional assumptions

- Updated Scenario Development: Created single scenario spreadsheets and clear documentation

- Sensitivity to the "Ds:" Used GIS centerline network and included accessibility variables

- Refined Post-Processors: Added flexibility to summary processes including select link assignment

Listed below are recommendations for updating the model, data, or usability beyond VMIP 2.

- Refine trip generation such that person trips and vehicle trips account for under-reporting of travel in the CHTS, and assigned traffic volumes reflect roadway counts.

- Refine economic factors at a more specific geography and calibrate the land use allocation model using the refined data.

- Continue to collect traffic count and transit ridership data, land use development (residential, school, and employees) to perform near-term forecasts.

- Review and update the highway and transit networks for future years, creating a link between the RTP projects and the model.

- Coordinate with other MPOs and update the inter-regional travel components as needed.

- Track demographics, economics, and related Ds variables over time to inform future scenario development.

- Evaluate shifts in future assumptions such as autonomous vehicles, demographics, fuel price, and land use development patterns.

The following sections describe the data collected for model estimation, calibration, and validation.

DATA ACQUISITION, REVIEW, AND SUMMARY

This section describes the data collection, review processes, and provides a summary of the data used in the estimation, calibration, and validation of the VMIP 2 models.

2010 CENSUS/2012 ACS

Updated land use cross-classification tables used 2012 ACS Census data and the finest available geography. Most required data were available at the level of census block group or census tract, but a few multi-dimension tables were only available at the Public Use Microdata Areas (PUMA) level. These cross-classification tables are in a percentage format. Each MPO/County provides the control totals for demographic variables including total population, total numbers of households, and total number of residential units at transportation analysis zone (TAZ) level. The base year for most models is 2008, although some MPOs/Counties have opted to update model base years to 2014 under separate contracts. ACS 2012 cross-classified tables represent demographic characteristic of each TAZ regardless of the model base year. The control total can easily be updated to a new base year after each MPO/County provides recent demographic data at the TAZ level.

2012 CHTS

The original VMIP 1, completed before 2012 CHTS data were available, used the 2001/2003 CHTS for validation of household variables. VMIP 2 used newer data from the 2012 CHTS to re-estimate most model components.

PREPARATION AND CLEANING OF CHTS DATA

The publically available version of the 2012 CHTS required a substantial amount of preparation, including re-weighting, before it was suitable for model development. Details of the data preparation are in Appendix A: Preparation of California Household Travel Survey Data. Data dictionaries for the cleaned and prepared CHTS data, including households, trips, and persons files, are in Appendix B: California Household Travel Survey Data Dictionary.

The following pages describe portions of the CHTS data preparation most relevant to VMIP 2; for full details please see the appendices.

Identification of Trip Purposes

The 2012 CHTS data does not describe trip purposes directly; instead, it contains a "place" file whose attributes include a listing of up to three activities the respondent participated in at that place. A small list of place purposes was distilled from this activity information: HOME, WORK, COLLEGE, K12, SHOP, or OTHER.

Once the purpose for each place has been determined, assigning a purpose to each trip is straightforward.

- If one end of the trip is "HOME" and the other is "WORK," the trip is home-based work ("HBW").

- If one end of the trip is "HOME" and the other is "K12," the trip is home-based K-12 ("HBK").

- If one end of the trip is "HOME" and the other is "COLLEGE," the trip is home-based college ("HBC").

- If one end of the trip is "HOME" and the other is "SHOP," the trip is home-based shop ("HBS").

- If one end of the trip is "HOME" and the other is either "OTHER" or "HOME," the trip is home-based other ("HBO").

- If one end of the trip is "WORK" and the other end is anything but "HOME," the trip is work-based other ("WBO").

- In all other cases, the trip is non-home-based ("NHB").

Identification and Consolidation of Transit Trip Chains

In recording transit trips, the CHTS treats each portion of the transit trip chain as a separate trip. For example, a trip in which the traveler drives to a rail station, takes the train to a second rail station, and then walks to a workplace is listed in the survey as three separate consecutive trips, with three separate modes. This method of record-keeping makes it possible to track the mode of access and egress for a transit trip, but for most travel behavior analyses it is preferable to consider these three trips as a single unit or linked trip. Thus, a necessary step of data preparation is identification and consolidation of chains which make up a single linked transit trip. Details of this process are in Appendix A: Preparation of California Household Travel Survey Data.

Estimation of Survey Weights

Surveys capture the characteristics of an entire population by randomly sampling a small proportion of the population. Often, a perfectly random sample is hard to achieve — some groups are difficult to survey and are under-represented, other groups are over-represented. To balance this bias, estimated sample weights "reshape" the sample. Fehr & Peers estimated household sample weights for the CHTS to balance the survey

sample to match county-level percentages for several variables as reported in the 2012 ACS 5-year estimates. Listed below are variables used as controls for the re-weighting.

- Household size (one to seven or more).

- Household income (nine income categories).

- Number of workers per household (zero to three or more).

- Number of vehicles owned per household (zero to four or more).

- Household residential unit type (three categories).

- Household size (one to five or more) cross-classified by household income (five categories).

- Household size (one to five or more) cross-classified by number of vehicles per household (zero to four or more).

- Household size (one to five or more) cross-classified by number of workers per household (zero to three or more).

Details of the survey weight estimation are in Appendix A: Preparation of California Household Travel Survey Data.

Census Designated Places

Census Designated Places (CDPs) are a useful identification that includes cities as well as unincorporated but named places. The fact that publically-available CHTS data is geo-coded only by census tract made the process of identifying a CDP for each location slightly more complex. Because the boundaries of CDPs do not neatly match census tracts, each census tract may have multiple CDPs associated with them. In cases where multiple CDPs make up a single census tract, the CDP with the largest population in the tract (as identified at the census block level) is used. The CDP is identified as an unincorporated portion of the relevant county if the largest population in the tract is outside all named CDPs.

Place Type

In addition to locating households and trip ends using census tracts, CDPs, and counties, each household location and trip end is assigned a place type category. The place type is based on the number of jobs and the working-age population accessible from the household or trip end. These accessibility metrics are available as part of the EPA Smart Location Database (http://www2.epa.gov/smartgrowth/smart-location-mapping#SLD), and are weighted so nearby jobs and population are more influential than distant jobs and population. The resulting sum of accessible jobs and potential workers are categorized into the following place types.

1. Under 40,000 jobs + workers.

2. 40,000 – 100,000 jobs + workers.

3. 100,000 – 200,000 jobs + workers.

4. 200,000 – 450,000 jobs + workers.

5. Over 450,000 jobs + workers.

"Work" Trips Made by Non-Workers

The CHTS collects both employment data for each participant and trip purpose data for all trips undertaken. However, the survey does not ensure these values are in agreement with one another. There are a small number of persons whose employment status is either not reported (or reported as "retired" or "unemployed") whose trips are categorized as work trips. Because this is not optimal for modeling purposes, any work trips made by a non-employed person is re-categorized; HBW trips are re-assigned as HBO trips, and WBO trips are re-assigned as OBO trips.

ESTIMATION DATASET

The estimation dataset for VMIP 2 consists of a portion of the statewide CHTS data. Only CHTS records which satisfy the following criteria were used.

- For household-level variables, only residents of the eight SJV counties and the six Sacramento Area Council of Governments (SACOG) counties are included. The six SACOG counties had to be included to ensure an adequate sample size.

- Only weekday trips are included.

- Trips are included from the full year of the CHTS, including winter and summer.

- Trips with both trip ends outside the 14-county SJV + SACOG region are excluded.

Table 1 shows the distribution of CHTS households in the estimation counties, the households reported in the ACS, and percentage of samples in the estimation set. Note the table shows the (unweighted) number of households in the estimation set and the full CHTS, while the value in the final column represents the percentage of the overall samples by county.

TABLE 1: GEOGRAPHIC SCALE FOR NEW TAZ VARIABLES

County	Households in Estimation Set	Total households in CHTS	Total households in County (2012 ACS)	Percentage of Estimation Set
Fresno	718	1,115	287,082	14%
Kern	961	1,544	253,178	12%
Kings	199	293	40,767	2%
Madera	205	311	42,063	2%
Merced	297	474	74,496	3%
San Joaquin	468	629	213,632	12%
Stanislaus	383	552	165,999	8%
Tulare	537	799	129,996	6%
Sacramento	567	825	512,496	25%
El Dorado	151	208	67,846	2%
Placer	290	385	131,775	7%
Sutter	130	168	31,635	2%
Yuba	137	205	24,133	1%
Yolo	186	246	70,090	4%
Total	**5,229**	**7,754**	**512,496**	**100%**

CHTS SUMMARIES

Several broad summaries of CHTS data were produced and are suitable both for model development and for general information. Separate summaries were produced for the 14-county estimation region, the eight-county SJV region, the three-county Three County Model region, and each of the eight SJV counties individually. The "simple" and "flat" summaries contain one record per geography, and is suitable for joining to GIS. The "simple" summary contains a smaller number of metrics, while the "flat" summary contains many more details. The "filterable" summary contains many records per geography, and is viewable in Excel. Details and data dictionaries for these summaries are in Appendix C: Simple Summaries of CHTS Data, Appendix D: Flat Summaries of CHTS Data, and Appendix E: Filterable Summaries of CHTS Data.

CHTS SIMPLIFIED DATA

In addition to being useful for model estimation, calibration, and validation, the CHTS data is useful for a wide range of other purposes. To that end, we have provided simplified versions of CHTS data together with instructions for processing that data in Excel. The format is designed to be flexible, easy to use, and able to produce a variety of commonly-requested summaries such as mode shares, trip lengths and origin/destination tables. More information about the simplified data and instructions for using it in Excel is in [Appendix F: Simplified CHTS Data](#).

HOUSING AFFORDABILITY, EMPLOYMENT AND JOBS/HOUSING BALANCE

Demographic and employment data are critical components to any land use, transportation, or integrated land use-transportation modeling effort. An appropriately detailed description of the people who live and work in each geographic zone is essential to understanding their travel behavior and in predicting the region's evolution over time, especially the relationship between the locations of employers paying a given range of wages and the residence locations of workers with similar income levels. There are many sources for this data, necessitating a data merge and verifying its compatibility with other datasets. CoStar led this effort. They used surveyors to call and visit residential, office, and commercial buildings and combined multiple demographic and transportation databases into a single web-accessible dashboard. CoStar continuously updates the data and keeps the historic data so changes in rents, vacancies, and other relevant variables can be evaluated. This data were used to calibrate the bid/rent functions of the land use allocation/disaggregation model, and to assist in the estimation and calibration of trip generation and distribution, allowing additional functionality to better match jobs and household income. The income of households and job salaries are described later in the calibration step.

REFINE MODEL INPUT DATA

TRANSPORTATION ANALYSIS ZONES

The TAZ system for each model is largely unchanged from the original VMIP 1. New TAZ attributes were developed to refine the model's trip distribution, including the matching of jobs to workers by income level and the distribution of trips entering and leaving the model area. In addition, the VMIP 2 models include both accessibility pre-processors and in-model accessibility calculations at the TAZ level, described below.

TAZ ATTRIBUTES

New attributes in the TAZ-level input data are listed below.

- Total acreage of the TAZ (including undeveloped land).

- Percentage of trips produced by the TAZ which enter or leave the model area, by trip purpose.

- Percentage of trips attracted to the TAZ which enter or leave the model area, by trip purpose.

- Percentage of jobs in the TAZ which are high-, medium-, and low-income, by employment category.

Table 2 below describes the geographic scale at which the trips produced/attracted and employment income variables are implemented in the model. The model user can change variables to apply at a different scale if desired, as described in the table.

TABLE 2: GEOGRAPHIC SCALE FOR NEW TAZ VARIABLES

Variables	Description	Scale of current implementation	Scale of potential implementation
HBWH_ix, HBWH_xi, HBWM_ix, HBWM_xi, etc.	Percentages of trips produced & attracted to TAZ, by trip purpose	CDP	TAZ
EMP_EDUH, EMP_EDUM, EMP_EDUL, etc.	Percentages of employment that are high, medium, and low income, by job sector	County	TAZ

The full data dictionary for TAZ-level inputs is in Appendix G: Data dictionary for TAZ data inputs.

ACCESSIBILITY

The VMIP 2 models include two accessibility pre-processors. These are Python scripts, operating on the input TAZ and network shapefiles to produce accessibility metrics.

- Intersections.py produces a count of the number of intersections per TAZ.

- RoadwayMiles.py produces the sum of walkable network miles.

These script outputs, in data base format (DBF), are used during the model input preparation stage to calculate a variety of accessibility metrics at the TAZ level.

A third input file, VMTseed, contains an estimate of the average commuting VMT generated per worker in the TAZ. The starting estimates can be approximate because this estimate is updated throughout the model process.

During the input preparation phase of the model, TAZ-level accessibility metrics and built environment ("D variable") metrics are produced. These metrics are updated as the model runs through its feedback loops. Some of the accessibility metrics are implemented later in the model; others are provided as model outputs. **Table 3** below shows the accessibility metrics used later in the model.

TABLE 3: ACCESSIBILITY METRICS USED IN VMIP 2 MODELS

Metric	Description	Where used
EMP_30AUT	Jobs within 30 minutes by auto	Place Type calculation
WRK_30AUT	Working-age population within 30 minutes by auto	Place Type Calculation
ATYPE	Place Type categorization of job+worker to five categories. (See Table 4 below).	Trip Generation
LOG_EMPD	Log of employment density (jobs per developed acre)	Auto Ownership, Mode Choice
INTDEN	Intersection density (intersections per square mile)	Auto Ownership, Mode Choice
EMP_30TRN	Jobs within 30 minutes by transit	Auto Ownership, Mode Choice
COMMUTECOST	Average annual commute cost	Auto Ownership

Place type is calculated from the sum of jobs within 30 minutes by auto- and working-age population within 30 minutes by auto, and categorized into the five categories listed in **Table 4** below.

TABLE 4: PLACE TYPES

Place Type Category	Alternate Name	Description
1	POP1	Under 40,000 jobs + working-age population within 30 minutes by auto
2	POP2	Between 40,000 and 100,000 jobs + working-age population within 30 minutes by auto
3	POP3	Between 100,000 and 200,000 jobs + working-age population within 30 minutes by auto
4	POP4	Between 200,000 and 450,000 jobs + working-age population within 30 minutes by auto
5	POP5	Over 450,000 jobs + working-age population within 30 minutes by auto

A full data dictionary of the accessibility metrics calculated in the model is in Appendix H: Accessibility Variables.

LAND USE INPUTS

During the original VMIP 1, Census 2000 land use data were used in combination with the CHTS 2001/03 to estimate and calibrate the trip generation rates. After Census 2000, the Census Bureau not only developed continuous sampling and reporting via the American Community Survey, but they also changed the format, variables, and detail of reported data. In 2012 it was discovered all of the variables used in the MIP models are not available at the same cross-classification detailed level as was reported in 2000. As such, we have updated the residential demographic variables at the same time we re-estimated trip generation equations.

In addition to the availability of data provided by the ACS and Census, updating the land use inputs at the same time trip information is estimated and calibrated allowed the opportunity to expand the capabilities to take advantage of the job salary and household mortgage/expense data. While the Census and ACS provide the information for the base year recalibration, the VMIP 2 models can now also use Cube Land to disaggregate the base year land use to reflect the validation conditions, allowing future forecasts of residential demographics to vary based on land use and transportation system changes.

Although the land use data and Cube Land model were implemented for each model, the application of Cube Land is not required. It can be used to disaggregate land use while keeping the totals by zone nearly identical, test brand new scenarios by allocating the control total for each land use type, or a middle scenario where some areas do not change and others can be allocated based on Cube Land.

Table 5 below describes the land use variables used as model inputs:

<div align="center">

TABLE 5: LAND USE INPUT VARIABLES

</div>

Type	Attribute	Description	Units
Geographic	TAZ	Transportation Analysis Zone ID	
	STATE	State	
	COUNTY	County	
	PUMA	Census Public Use Microdata Area	
	CITY	City	
	TRACT	Census tract ID	
	BLOCK	Census block ID	
	MODEL	Model ID	
	PLACETYPE [1]	Placetype category	
Residential	TOTHH	Total Households	Households
	RU1, RU2, ... RU10 [2]	Households by Residential Unit Type	Households
	RUG1, RUG2, RUG3 [2]	Households by Residential Unit Type Groups	Households
	RUG1SPARE, ... RUG7SPARE	Unused in current model but available for expanding grouping of residential unit types.	
Non-residential[3]	TOTEMP	Total employees	Employees
	EMPEDU	Educational Services (61-63)	Employees
	EMPFOO	Accommodations (721), Food Services (722), Arts, Entertainment and Recreation (71)	Employees
	EMPGOV	Public Administration (92)	Employees
	EMPIND	Utilities (22), Construction (23), Other Services Except Public Administration (81), Wholesale Trade (42), Transportation and Warehousing (48-49)	Employees
	EMPMED	Health Care and Social Assistance (62)	Employees
	EMPOFC	Information (51), Finance and Insurance (52), Real Estate, Rental and Leasing (53), Professional, Scientific, and Technical Services (54), Management of Companies and Enterprises (55), Administrative/Support, Waste Management & Remediation (56)	Employees
	EMPOTH	Mining, Quarrying, Oil and Gas Extraction (21), Manufacturing (31-33)	Employees

TABLE 5: LAND USE INPUT VARIABLES

Type	Attribute	Description	Units
	EMPRET	Retail Trade (44-45)	Employees
	EMPAGR	Agriculture, Forestry, Fishing and Hunting (11)	Employees
	EMPSPARE1, … EMPSPARE8	Unused in current model but available for expanding employment categories	
	POPDORM	Group Quarters population: School (Dormitory, Fraternity, Sorority)	People
	POPASSIST	Group Quarters Population: Medical (Assisted living, retirement home)	People
	POPMILITARY	Group Quarters Population: Military (Military base if not special generator)	People
	POPINST	Group Quarters Population: Institutionalized population (prison, mental health, etc.)	People
	ELEM	Elementary and middle school enrollment	Student Enrollment
	HS	High school enrollment	Student Enrollment
	COLLEGE	College enrollment	Student Enrollment
Scenario	YEAR	Scenario year	
	SCEN	Scenario name	
	MPO	MPO	
	Comments	Scenario comments	

Notes:
     1. See Table 4 for place type categories.
     2. See Table 8 for residential unit type categories.
     3. Non-residential description contains NAICS sector number(s).

The land use inputs above are combined with the Census cross-classification rates to create the SE Detail file, described in **Table 6** below.

TABLE 6: SOCIO-ECONOMIC DETAIL

Type	Attribute	Description	Units
Geographic	TAZ	Transportation Analysis Zone ID	
	STATE	State	
	COUNTY	County	
	PUMA	Census Public Use Microdata Area	

TABLE 6: SOCIO-ECONOMIC DETAIL

Type	Attribute	Description	Units
Residential	CITY	City	
	TRACT	Census tract ID	
	BLOCK	Census block ID	
	MODEL	Model ID	
	PLACETYPE [1]	Placetype category	
	TOTHH	Total Households	Households
	RUG1, RUG2, RUG3 [2]	Households by Residential Unit Type Groups	Households
	RUG1SPARE, ... RUG7SPARE	Unused in current model but available for expanding grouping of residential unit types.	
	RU1_HHPOP, RU3_HHPOP, RU6_HHPOP [2]	Population in households by residential unit type	People
	RUSPARE1, ... RUSPARE7	Unused in current model but available for expanding grouping of residential unit types	
	RU1_HHSIZE1_INC1, RU9_HHSIZE5_INC5 [2,3,4]	Households cross-classified by Residential Unit Type, Household Size, and Household Income	Households
	RU1_AGE1524, ... RU9AGE75 [2,5]	Households cross-classified by Residential Unit Type and Household Age category.	Households
	POP0005, ... , POP75 [6]	Population by age range	People
Non-residential [7]	TOTEMP	Total employees	Employees
	EMPEDU	Educational Services (61-63)	Employees
	EMPFOO	Accommodations (721), Food Services (722), Arts, Entertainment and Recreation (71)	Employees
	EMPGOV	Public Administration (92)	Employees
	EMPIND	Utilities (22), Construction (23), Other Services Except Public Administration (81), Wholesale Trade (42), Transportation and Warehousing (48-49)	Employees
	EMPMED	Health Care and Social Assistance (62)	Employees

TABLE 6: SOCIO-ECONOMIC DETAIL

Type	Attribute	Description	Units
	EMPOFC	Information (51), Finance and Insurance (52), Real Estate, Rental and Leasing (53), Professional, Scientific, and Technical Services (54), Management of Companies and Enterprises (55), Administrative/Support, Waste Management & Remediation (56)	Employees
	EMPOTH	Mining, Quarrying, Oil and Gas Extraction (21), Manufacturing (31-33)	Employees
	EMPRET	Retail Trade (44-45)	Employees
	EMPAGR	Agriculture, Forestry, Fishing and Hunting (11)	Employees
	EMPSPARE1, … EMPSPARE8	Unused in current model but available for expanding employment categories	
	POPDORM	Group Quarters population: School (Dormitory, Fraternity, Sorority)	People
	POPASSIST	Group Quarters Population: Medical (Assisted living, retirement home)	People
	POPMILITARY	Group Quarters Population: Military (Military base if not special generator)	People
	POPINST	Group Quarters Population: Institutionalized population (prison, mental health, etc.)	People
	ELEM	Elementary and middle school enrollment	Student Enrollment
	HS	High school enrollment	Student Enrollment
	COLLEGE	College enrollment	Student Enrollment
Scenario	YEAR	Scenario year	
	SCEN	Scenario name	
	MPO	MPO	
	Comments	Scenario comments	

Notes:
1. See Table 7 for place type categories.
2. See Table 8 for residential unit type categories.
3. See Table 9 for household size categories.
4. See Table 10 for household annual income categories.
5. See Table 11 for household age categories.
6. See Table 12 for population distribution by age range categories.
7. Non-residential description contains NAICS sector number(s).

If desired, preliminary place type descriptions may be included in the land use input. Within the VMIP 2 models, place type is re-calculated as part of the accessibility module described in Accessibility / D Variables.

TABLE 7: PLACE TYPES

Place Type Category	Alternate Name	Description
1	POP1	Under 40,000 jobs + working-age population within 30 minutes by auto
2	POP2	Between 40,000 and 100,000 jobs + working-age population within 30 minutes by auto
3	POP3	Between 100,000 and 200,000 jobs + working-age population within 30 minutes by auto
4	POP4	Between 200,000 and 450,000 jobs + working-age population within 30 minutes by auto
5	POP5	Over 450,000 jobs + working-age population within 30 minutes by auto

TABLE 8: RESIDENTIAL UNIT TYPE

Name	Grouping	Alternate Grouping Name	Description
RU1	RUG1 (SF)	RU1	1, detached
RU2			1, attached
RU3			2 units
RU4			3 to 4 units
RU5	RUG2 (MF)	RU3	5 to 9 units
RU6			10 to 19 units
RU7			20 to 49 units
RU8			50+ units
RU9	RUG3 (Other)	RU9	Mobile home
RU10			Boat, RV, van, etc.

Data sources:

Model input: MPO land use inputs

Estimation: CHTS

Calibration: Census

TABLE 9: HOUSEHOLD SIZE

Category	Description
HHSIZE1	1 person household
HHSIZE2	2 person household
HHSIZE3	3 person household
HHSIZE4	4 person household
HHSIZE5	5 or more person household

Source:

Model Input: MPO land use inputs + census cross-classification percentages

Estimation: CHTS

Calibration: Census

TABLE 10: HOUSEHOLD ANNUAL INCOME

High-med-low grouping	5-category grouping	10-category grouping	Description
LOWINC	INCG1	INC1	Less than $10,000
		INC2	$10,000 to $24,999
	INCG2	INC3	$25,000 to $34,999
		INC4	$35,000 to $49,999
MEDINC	INCG3	INC5	$50,000 to $74,999
	INCG4	INC6	$75,000 to $99,999
HIGHINC	INCG5	INC7	$100,000 to $149,999
		INC8	$150,000 to $199,999
		INC9	$200,00 or more
		INC10	SPARE -- unused

Data sources:

Model Input: MPO land use inputs + census cross-classification percentages

Estimation: CHTS

Calibration: Census

TABLE 11: HOUSEHOLD AGE

Category	Description
Age1524	No household member over age 25 but at least one household member age 15-24.
Age2564	Household has at least one member age 25-64
Age6574	Household has no member age 25-64 but at least one member age 65-74.
Age75	Household has no member age 25-74 but at least one member age 75 or older.

Data sources:

Model Input: MPO land use inputs + census cross-classification percentages

Estimation: CHTS

Calibration: Census

TABLE 12: POPULATION BY AGE RANGE

Category	Description
POP0005	People 0 to 5 years
POP0514	People 5 to 14 years
POP1517	People 15 to 17 years
POP1824	People 18 to 24 years
POP2554	People 25 to 54 years
POP5564	People 55 to 64 years
POP6574	People 65 to 74 years
POP75	People 75 years and over

Source:
Model Input: MPO land use inputs + census cross-classification percentages
Estimation: CHTS
Calibration: Census

Appendix I: Comparison of land use categories shows the residential land use data elements and how the VMIP 2 grouping compares to other data sources including the CHTS, ACS, and VMIP 1 categorization.

NETWORK UPDATE

As part of the VMIP 1, integration of GIS for each of the models took a substantial step forward by utilizing a geodatabase for background data and for storing model outputs. However, the highway and transit networks remained simplistic link and node representations of the actual networks. As part of VMIP 2, the highway network was based on a true shape centerline file in a geodatabase and updated variables to reflect the master network from the RTP/SCS. The transit lines were also updated to match the more detailed highway network and are contained in the geodatabase. The benefits of this are more accurate mapping and distances, easy linkage and comparisons to speed data, and inclusion of local streets for sub-TAZ level analysis. In addition, the GIS network contains many variables to complement those already part of the travel model network, including auto, HOV, transit, truck, bike, and walk accessibility designations. Advanced models such as Activity Based Models (ABMs) and Dynamic Traffic Assignment (DTA) also greatly benefit from the network accuracy and detail.

TABLE 13: STANDARD MASTER HIGHWAY NETWORK VARIABLES

Attribute	Description
Nodes	
X	X-coordinate of node in Nad 83
Y	Y-coordinate of node in Nad 83
N	Node number
TAZ	Traffic Analysis Zone Number
DISTRICT	Super district number used for aggregation
SOI	Sphere of influence used to number TAZs alphabetically
STDID	Study location number used to record turning movements when non-zero
COUNTY	County where node is located
JURISDICTION	Political jurisdiction where node is located
COMMUNITY	Community/district name
Links	
A	A node
B	B node
DISTANCE	Distance in miles
NAME	Local street name
ROUTE	Numerical state route number
TERRAIN	Terrain (F=Flat , R=Rolling, M=Mountain)
JURISDICTION	Political jurisdiction where link is located location
SCREENLINE	Screenline by direction (See Figures 3-1.1 through 3.1.10)
XXXX_PRJID [1]	RTP Project ID number
XXXX_PRJYR [1]	RTP Project Opening Year
XXXX_FACTYP [1]	Facility type by year [2]
XXXX_AREATYP [1]	Area type by year [2]
XXXX_LANES [1]	Number of directional through travel lanes by year [2]
XXXX_AUX [1]	Auxiliary lane (0=no, 1=yes)
XXXX_SPEED [1]	Free-flow speed in miles-per hour by year [3]

TABLE 13: STANDARD MASTER HIGHWAY NETWORK VARIABLES

Attribute	Description
XXXX_CAPCLASS [1]	Capacity class by year (derived from Terrain, Facility type, and Area Type) [2]
XXXX_CAPACITY [1]	Vehicle per hour (calculated based on Lanes and CapClass) [4]
XXXX_USE [1]	Identifies vehicle prohibitions by year [5]
XXXX_TOLL [1]	Code used for cost on toll facilities by year [3]

Notes:
1. XXXX represents BASE (calibration/validation year), IMP1 (status after first improvement), and IMP2 (status after second improvement). In addition to calibration/validation year which varies by MPO, the years required to be covered by improvement are 05, 20, 35, and 40.
2. See Table 14 for details on CapClass by Terrain, Facility Type, and Area Type.
3. See Table 15 for Speed ranges by Terrain, Facility Type, and Area Type.
4. See Table 16 for details on Capacity by Terrain, Facility Type, and Area Type.
5. 0 or 1=facility open to all ("general purpose") ; 2=Carpool 2; 3=Carpool 3+; 4=Combination trucks prohibited; 5=Walk or bike only

TABLE 14: CAPACITY CLASS BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
Flat					
Freeway	1	11	21	31	41
Highway	2	12	22	32	42
Expressway	3	13	23	33	43
Arterial	4	14	24	34	44
Collector	5	15	25	35	45
Local	6	16	26	36	46
Ramp: Freeway-Freeway	7	17	27	37	47
Ramp: Slip	8	18	28	38	48
Ramp: Loop	9	19	29	39	49
Connector: Dist. ≤ 0.25	10	N/A	N/A	N/A	N/A
Connector: Dist. > 0.25	20	N/A	N/A	N/A	N/A

TABLE 14: CAPACITY CLASS BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
Rolling					
Freeway	51	61	71	81	91
Highway	52	62	72	82	92
Expressway	53	63	73	83	93
Arterial	54	64	74	84	94
Collector	55	65	75	85	95
Local	56	66	76	86	96
Ramp: Freeway-Freeway	57	67	77	87	97
Ramp: Slip	58	68	78	88	98
Ramp: Loop	59	69	79	89	99
Connector: Dist. ≤ 0.25	60	N/A	N/A	N/A	N/A
Connector: Dist. > 0.25	70	N/A	N/A	N/A	N/A
Mountain					
Freeway	101	111	121	131	141
Highway	102	112	122	132	142
Expressway	103	113	123	133	143
Arterial	104	114	124	134	144
Collector	105	115	125	135	145
Local	106	116	126	136	146
Ramp: Freeway-Freeway	107	117	127	137	147
Ramp: Slip	108	118	128	138	148
Ramp: Loop	109	119	129	139	149
Connector: Dist. ≤ 0.25	110	N/A	N/A	N/A	N/A
Connector: Dist. > 0.25	120	N/A	N/A	N/A	N/A

TABLE 15: TYPICAL SPEEDS BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
Flat					
Freeway	70	65-70	55-65	55-65	55-65
Highway	40-45	40-45	40-45	40-45	40-45
Expressway	55	45-55	45-55	45-55	40-45
Arterial	40-45	30-45	25-45	30-45	25-45
Collector	50	50	35-40	35-40	35-40
Local	25-40	25-40	25-40	25-40	25-40
Ramp: Freeway-Freeway	50	50	50	50	50
Ramp: Slip	50	50	50	50	50
Ramp: Loop	45	45	45	45	45
Connector: Dist. ≤ 0.25	35	35	35	35	35
Connector: Dist. > 0.25	15	15	15	15	15
Rolling					
Freeway	65-70	65-70	65-70	65-70	65-70
Highway	40-45	40-45	40-45	40-45	40-45
Expressway	50-65	50-65	50-65	50-65	50-65
Arterial	30-45	30-45	30-45	30-45	30-45
Collector	50	50	50	50	50
Local	50	50	50	50	50
Ramp: Freeway-Freeway	50	50	50	50	50
Ramp: Slip	50	50	50	50	50
Ramp: Loop	45	45	45	45	45
Connector: Dist. ≤ 0.25	35	35	35	35	35
Connector: Dist. > 0.25	15	15	15	15	15
Mountain					

TABLE 15: TYPICAL SPEEDS BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
Freeway	65	65	65	65	65
Highway	40-45	40-45	40-45	40-45	40-45
Expressway	40-55	40-55	40-55	40-55	40-55
Arterial	30-45	30-45	30-45	30-45	30-45
Collector	25-40	25-40	25-40	25-40	25-40
Local	25-40	25-40	25-40	25-40	25-40
Ramp: Freeway-Freeway	50	50	50	50	50
Ramp: Slip	45	45	45	45	45
Ramp: Loop	35	35	35	35	35
Connector: Dist. ≤ 0.25	15	15	15	15	15
Connector: Dist. > 0.25	25	25	25	25	25

Note: Speed shown as miles per hour (MPH)

TABLE 16: DEFAULT CAPACITY BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
Flat					
1. Freeway	2,000	2,000	1,800	1,750	1,750
2. Highway	1,800	1,800	1,600	1,500	1,300
3. Expressway	1,100	1,100	1,000	900	800
4. Arterial	900	900	900	800	750
5. Collector	700	700	800	800	700
6. Local	600	600	700	700	600

TABLE 16: DEFAULT CAPACITY BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type		Area Type				
		Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
7.	Ramp: Freeway-Freeway	1,800	1,800	1,800	1,800	1,800
8.	Ramp: Slip	1,500	1,500	1,500	1,500	1,500
9.	Ramp: Loop	1,250	1,250	1,250	1,250	1,250
10.	Connector: Internal	N/A	N/A	N/A	N/A	N/A
Rolling						
20.	Connector: External (except major gateways)	N/A	N/A	N/A	N/A	N/A
21.	Freeway	1,800	1,800	1,620	1,580	1,580
22.	Highway	1,460	1,460	1,300	1,220	1,060
23.	Expressway	890	890	810	730	650
24.	Arterial	730	730	730	650	610
25.	Collector	570	570	650	650	570
26.	Local	550	550	640	640	550
27.	Ramp: Freeway-Freeway	1,800	1,800	1,800	1,800	1,800
28.	Ramp: Slip	1,500	1,500	1,500	1,500	1,500
29.	Ramp: Loop	1,250	1,250	1,250	1,250	1,250
Mountain						
31.	Freeway	1,500	1,500	1,350	1,310	1,310
32.	Highway	790	790	700	660	570
33.	Expressway	480	480	440	390	350
34.	Arterial	390	390	390	350	330
35.	Collector	310	310	350	350	310
36.	Local	330	330	380	380	330

TABLE 16: DEFAULT CAPACITY BY TERRAIN, FACILITY TYPE, AND AREA TYPE

Facility Type	Area Type				
	Rural (R)	Suburban (SU)	Urban (U)	Fringe (F)	Central Business District (CBD)
37. Ramp: Freeway-Freeway	1,800	1,800	1,800	1,800	1,800
38. Ramp: Slip	1,500	1,500	1,500	1,500	1,500
39. Ramp: Loop	1,250	1,250	1,250	1,250	1,250

Note: Capacity shown as vehicles per hour per lane (VPHPL)

ESTIMATION, CALIBRATION, AND VALIDATION

This section covers the model estimation with the enhancements of integrating D variables within the sub-models and a revised inter-regional process to capture the interaction between household income and job salary. Values presented in this section are those estimated based on the entire survey data set, and **Appendix L** contains the resulting calibrated values.

ECONOMIC LAND USE FORECASTING

VMIP 1 developed and implemented an integrated transportation and standard socioeconomic land use forecasting model structure by expanding the pilot project for Kern COG. This system supports the travel demand models by allocating study area forecast control totals of households and jobs by type to zones within the study area based upon bid-rent economic principles. This approach to land use forecasting provides a way of recognizing the important effects that linkages between spatial distributions of housing costs, household incomes, and job industries have on intra- and inter-regional travel. It also provides a way to automate the otherwise tedious and error-prone process of disaggregating land use assumptions developed through scenario visioning exercises into more detailed household and job type stratifications for travel modeling.

TRAVEL MODEL ESTIMATION

VMIP 2 re-estimated the trip generation, auto availability, and mode choice model components using data from the 2012 CHTS. The fairly limited sample size, particularly for transit and non-motorized trips, precluded the estimation of county-specific mode choice model coefficients. Instead, models were estimated using data from all eight San Joaquin Valley counties, together with the six SACOG counties. Each model was then calibrated to fit local conditions using CHTS data for its county/counties alone. Calibration values are in Appendix L: Calibrated Parameters.

The table below shows the re-estimated model components for VMIP 2, including a description of the model structure and a list of variables used. Detailed descriptions of each model component and its estimation are in the following sections.

TABLE 17: RE-ESTIMATED MODEL COMPONENTS

	Vehicle Availability	Trip Generation	Mode Choice
Model Structure	Disaggregate: multinomial logit	Aggregate: 4-dimensional cross-class models or regression models. Stratified by productions vs attractions and trip purpose.	Disaggregate: multinomial logit. Stratified by trip purpose and vehicle availability + household size.
Household Size	HH1, HH2, HH3, HH4, HH5	HH1, HH2, HH3, HH4, HH5	
Household Income	INCG1, INCG2, INCG3, INCG4, INCG5	INCG1, INCG2, INCG3, INCG4, INCG5	
Housing Type	RUG1, RUG3, RUG6	RUG1, RUG3, RUG6	
Accessibility / D variables	Intersection density, transit accessibility to jobs, employment density	Place Types based on auto accessibility to jobs and workers: pop1, pop2, pop3, pop4, pop5	Intersection density, transit accessibility to jobs, employment density
Age of population		POP0005, POP0514, POP1517, POP1824, POP2554, POP5564, POP75	
Employment		EMPEDU, EMPFOO, EMPGOV, EMPIND, EMPMED, EMPOFC, EMPRET, EMPOTH, EMPAGR	
School Enrollment		ELEM, HS, COLLEGE	
In-vehicle travel time			Applies to all modes. Transit amenities, if any, can be discounted here.
Out of vehicle time			Access/egress/transfer walk and waiting time for transit, parking time for drive-to-transit, and passenger pickup for shared ride.
Cost	Commute cost proportion of household income		Transit fare, plus toll and parking costs as well as auto operating costs for drive modes.

ACCESSIBILITY / D VARIABLES

All three of the re-estimated model components make use of built environment ("D variables"), particularly the inclusion of accessibility. The table below describes the variables used.

TABLE 18: ACCESSIBILITY METRICS USED IN VMIP 2 MODELS

Metric	Description	Where used
EMP_30AUT	Jobs within 30 minutes by auto	Place Type calculation
WRK_30AUT	Working-age population within 30 minutes by auto	Place Type Calculation
ATYPE	Place Type categorization of job+worker to five categories. (See table 19 below).	Trip Generation
LOG_EMPD	Log of employment density (jobs per developed acre)	Auto Ownership, Mode Choice
INTDEN	Intersection density (intersections per square mile)	Auto Ownership, Mode Choice
EMP_30TRN	Jobs within 30 minutes by transit	Auto Ownership, Mode Choice
COMMUTECOST	Average annual commute cost	Auto Ownership

Place type is calculated from the sum of jobs within 30 minutes by auto and working-age population within 30 minutes by auto, and categorized into the five categories listed below.

TABLE 19: PLACE TYPES

Place Type Category	Alternate Name	Description
1	POP1	Under 40,000 jobs + working-age population within 30 minutes by auto
2	POP2	Between 40,000 and 100,000 jobs + working-age population within 30 minutes by auto
3	POP3	Between 100,000 and 200,000 jobs + working-age population within 30 minutes by auto
4	POP4	Between 200,000 and 450,000 jobs + working-age population within 30 minutes by auto
5	POP5	Over 450,000 jobs + working-age population within 30 minutes by auto

A full data dictionary of the accessibility metrics calculated in the model is in Appendix H: Accessibility Variables.

VEHICLE AVAILABILITY AND TRIP GENERATION

The original VMIP 1 resulted in all models generating person trips by vehicle availability from a very consistent set of land uses. Household trips were generated for eight different purposes, and truck trips were generated for light, medium, and heavy trucks. With the new CHTS data we have re-estimated the vehicle availability and trip generation rates. In addition to the cross-classifications currently used in the models we have added place classifications that relate jobs/housing, income and long distance commuting, and other factors that were not available in previous data sets. To better link jobs and housing, the HBW trip purpose was split into three purposes corresponding to high, medium, and low income households and jobs.

Auto Operating Cost

Auto operating costs were determined using the methodology outlined in the memo prepared by MTC, SCAG, SACOG, and SANDAG in October 2014 titled *Automobile Operating Cost for the Second Round of Sustainable Communities Strategies*. The method uses county specific base year fuel prices, fleet mix and fuel efficiency from EMFAC, and a consistent growth factor for fuel and non-fuel maintenance and operating costs. See Appendix K: Memo on Auto Operating Cost for the full memo and methodology. The resulting values for years ranging from 2005 to 20420 for each MPO is in Appendix L: Calibrated Parameters.

Vehicle Availability

The vehicle availability model is a disaggregate multinomial logit model which predicts the probability of a household owning 0, 1, 2, or 3, or 4+ vehicles based on the following variables:

TABLE 20: VARIABLES IN VMIP 2 VEHICLE AVAILABILITY MODEL

Category	Variable	Description
Cost Variable	Commute Cost Ratio	Average annual commute cost divided by household income
Accessibility Variables	Intersection Density	Intersections per square mile
	Transit Accessibility	Jobs within 30 minutes via transit
	Employment Density	Log of (jobs per developed acre)
Household Demographic Variables	Household Size	See size categories in Table 9
	Household Income	See income categories in Table 10
	Household Residential Unit Type	See residential unit type groups in Table 11

The commute cost ratio variable is an estimate of the proportion of a household's income required to own vehicles. It is derived from a county-level estimate of per-mile auto ownership costs, tract-level estimates of commuting VMT derived from the EPA's Smart Location Calculator, an annualization factor of 250 working days per year, and the household income. The variable is applied on a per-vehicle basis, so that owning no vehicles incurs no cost, owning two vehicles incurs twice the cost of owning one vehicle, and so on.

The table below provides the coefficients of the auto ownership model. In its draft form the model was estimated without alternative-specific constants. These constants were set for each model individually during model calibration.

TABLE 21: VMIP 2 AUTO OWNERSHIP MODEL COEFFICIENTS

	0 Vehicles	1 Vehicle	2 Vehicles	3 Vehicles	4+ Vehicles
Alternative-Specific Constant					
CommuteCostRatio	7.51	3.95	0.00	0.00	0.00
PedOrIntDens	0.009	0	0	-0.004	-0.004
TransitAccessibility (x1000)	0.009	0.010	0	-0.051	-0.112
LogEmpDensity	0.39	0.24	0	0.00	-0.19
RUGroup=RU1	0	0	0	0	0
RUGroup=RU3	1.27	0.53	0	-1.53	-1.53
RUGroup=RU6	0.27	-.27	0	0	0
HH_size=1	-1.16	1.5	0	-3.15	-4.94
HH_size=2	-3.03	-0.42	0	-2.26	-4.19
HH_size=3	-3.37	-0.24	0	-1.34	-3.40
HH_size=4	-4.02	-0.66	0	-1.61	-3.13
HH_size=5+	-3.50	-0.89	0	-1.32	-2.44
HH_inc=IncG1	0	0	0	0	0
HH_inc=IncG2	-1.33	-0.28	0	0.86	0.98
HH_inc=IncG3	-3.87	-0.93	0	1.2	2.35
HH_inc=IncG4	-2.98	-1.55	0	1.55	2.35
HH_inc=IncG5	-4.23	-1.96	0	1.44	2.87

Note the model uses owning two vehicles as its base, and calculates the relative probability of owning fewer or greater vehicles; thus the model coefficients describe relative probabilities as in the example below:

$$\ln\left(\frac{Prob(0\ vehicles)}{Prob(2\ vehicles)}\right) = 7.51(CommuteCostRatio) + 0.0093(PedOrIntDensity) + \ \dots$$

The coefficients for this model are generally intuitive in direction and scale.

- Higher commuting cost increases the probability of owning 0 or 1 vehicles, and decreases the probability of owning 3 or 4 vehicles, as compared to the baseline of 2 vehicles.

- Higher scores for the three accessibility variables, indicating generally better accessibility by non-auto modes, increase the probability of owning 0 vehicles (and sometimes also 1 vehicle) relative to owning 2; and decrease the probability of owning 3 or 4.

- Household income is the demographic variable which has the largest influence in auto ownership. Generally as incomes go up, probabilities of owning 0 or 1 vehicles go down, and probabilities of owning 3 or 4 vehicles go up.

- Household size behaves in the expected way, with probability of owning 0 or 1 vehicles going down as household size increases and probability of owning 3 or 4 vehicles going up.

- Multi-family unit types are more likely to own 0 or 1 vehicles, and less likely to own 3 or 4 vehicles, than single family. There weren't enough records in the RUG6 "other" category (RV, mobile home, etc.) to distinguish them from single family, and they were generally more similar to single family than multi-family uses, so they share the same coefficients as single family.

An important consideration for future model development is that car sharing and transportation network companies (i.e., UBER, LYFT, etc.) are changing auto availability dynamics and potentially long-term auto ownership. As more data becomes available it may be appropriate to modify the auto ownership model to recognize these changes and focus more on auto availability across multiple sub modes and costs per mile.

Trip Generation

The VMIP 2 models generate person-trips from a consistent set of land uses, using cross-classified residential data, for a number of purposes including non-home-based purposes, K-12 and college trip purposes, and generate small, medium, and heavy truck trips. We have re-estimated trip generation rates, excluding truck rates, with the new CHTS data. The most significant changes in trip generation as compared to original VMIP 1 are listed below.

- Trip generation considers accessibility using the place type variable described in

- Accessibility / D Variables.

- Non-home based trip generation is based on the new categorization of employment.

- HBW trips are expanded into three new categories: HBW-High, HBW-Medium, and HBW-Low. These categories are based on household income on the production side and proportions of worker incomes for each employment category on the attraction side.

- Trips are classified as internal to internal (II), internal to external (IX), or external to internal (XI) based on percentages calculated from CHTS data. These percentages are calculated by trip purpose and by CDP.

Home-Based Productions: Cross-Classification Models

Three of the home-based trip productions (HBW, HBS, and HBO) were estimated using cross-classification models. These models are applied to socio-economic-demographic (SED) data which has been cross-classified by four variables: household size, household income, residential unit type, and place type (as described in section [Accessibility / D Variables](#).

Estimation of trip rates using cross-classification models must ensure all cross-classification groups have large enough sample sizes to produce sufficient variability to obtain a stable average trip rate. Because not all cross-classifications of the variables above do in fact have a large enough sample size, some cross-classifications were estimated in aggregate, resulting in identical trip rates being estimated for some cross-classification combinations.

Variables were added to the cross-classification model sequentially, and with each added variable existing groups were only subdivided if there was sufficient sample size (generally at least 40 households) to support a split. The order in which variables were added to the cross-classification models was as follows.

- Household size

- Household income

- Place Type

- Residential unit type

Although the model is coded to allow for five income categories and five place types, the data available did not allow for distinctions to be determined this finely either because of a lack of sufficient amount of data, or differences which weren't statistically significant, or both. In effect, this means the estimated trip rates differ only among three income categories: low (under $50,000), medium ($50,000 - $100,000), and high (over $100,000); and only between two groups of place types: types 1 and 2 (with fewer than 100,000 workers+jobs within a 30-minute auto trip); and types 3, 4, and 5 (with more than 100,000 workers+jobs

within a 30-minute auto trip). In addition, only a few combinations of household size, household income, and place type yielded different trip rates by residential unit type.

The tables below provide the resulting person-trip production rates:

**TABLE 22: HBW HOUSEHOLD PERSON TRIP PRODUCTION RATES
(DAILY TRIPS PER HOUSEHOLD)**

	1-person HH	**2-person HH**	**3-person HH**	**4-person HH**	**5+-person HH**
Low Income; *Place Types 1 and 2*	0.42 (SF) 0.24 (MF)	0.62 (SF) 0.45 (MF)	0.87	1.28	1.50
Low Income; *Place Types 3, 4, 5*	0.55 (SF) 0.43 (MF)	0.80 (SF) 0.92 (MF)	1.35	1.27	1.49
Medium Income; *Place Types 1 and 2*	0.79	1.13	1.57	1.72	2.40
Medium Income; *Place Types 3, 4, 5*	0.68	1.17	1.62	1.47	2.25
High Income; *Place Types 1 and 2*	0.61	1.42	1.63	1.75	1.84
High Income; *Place Types 3, 4, 5*	0.61	1.26	2.04	1.62	1.84

**TABLE 23: HBS HOUSEHOLD PERSON TRIP PRODUCTION RATES
(DAILY TRIPS PER HOUSEHOLD)**

	1-person HH	2-person HH	3-person HH	4-person HH	5+-person HH
Low Income; Place Types 1 and 2	0.32 (SF) 0.46 (MF)	0.95 (SF) 0.93 (MF)	1.32	1.57	1.75
Low Income; Place Types 3, 4, 5	0.34 (SF) 0.50 (MF)	0.63 (SF) 0.71 (MF)	0.77	1.26	1.67
Medium Income; Place Types 1 and 2	0.36	0.55	0.49	0.62	1.37
Medium Income; Place Types 3, 4, 5	0.45	0.70	1.11	0.81	1.39
High Income; Place Types 1 and 2	0.25	0.56	0.50	0.34	1.01
High Income; Place Types 3, 4, 5	0.25	0.78	1.03	1.14	1.01

**TABLE 24: HBO HOUSEHOLD PERSON TRIP PRODUCTION RATES
(DAILY TRIPS PER HOUSEHOLD)**

	1-person HH	2-person HH	3-person HH	4-person HH	5+-person HH
Low Income; Place Types 1 and 2	1.68 (SF) 0.92 (MF)	2.50	3.57	5.02	7.61
Low Income; Place Types 3, 4, 5	1.35 (SF) 1.14 (MF)	2.69 (SF) 2.59 (MF)	3.83	7.13	9.94
Medium Income; Place Types 1 and 2	1.44	2.17	3.09	5.59	9.06
Medium Income; Place Types 3, 4, 5	1.57	2.92	4.30	6.84	11.10
High Income; Place Types 1 and 2	1.73	1.94	4.94	6.45	8.51
High Income; Place Types 3, 4, 5	1.73	2.69	4.04	7.50	8.51

The graphs below show the cross-classified trip production rates.
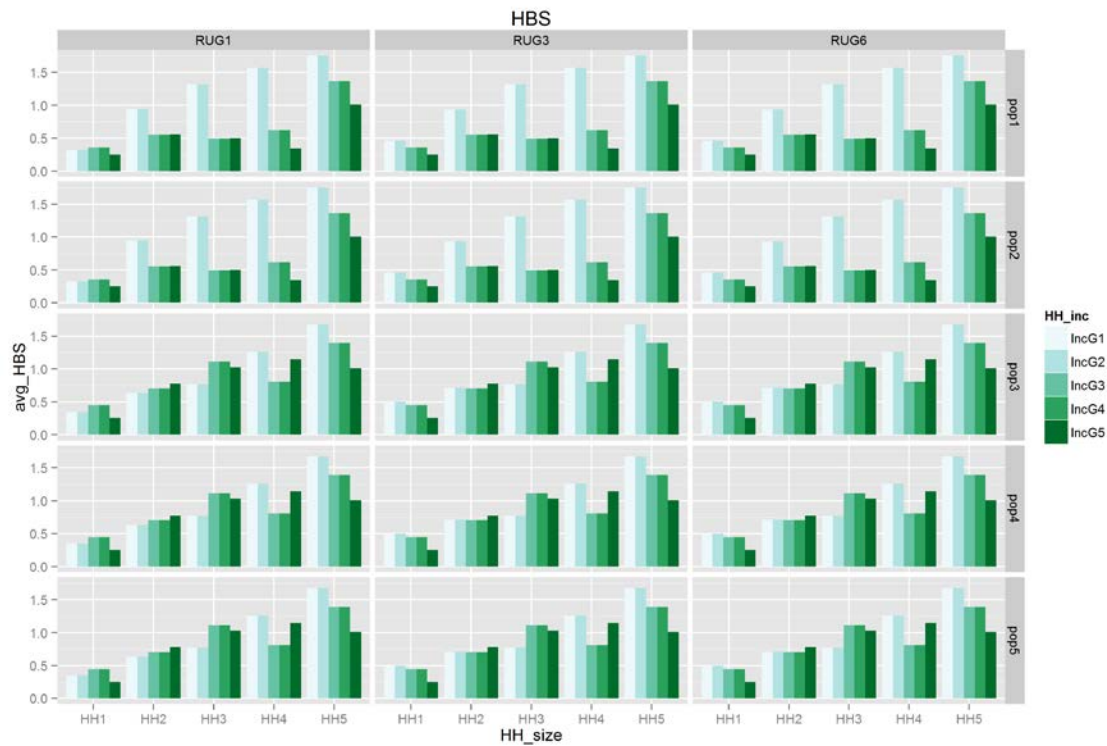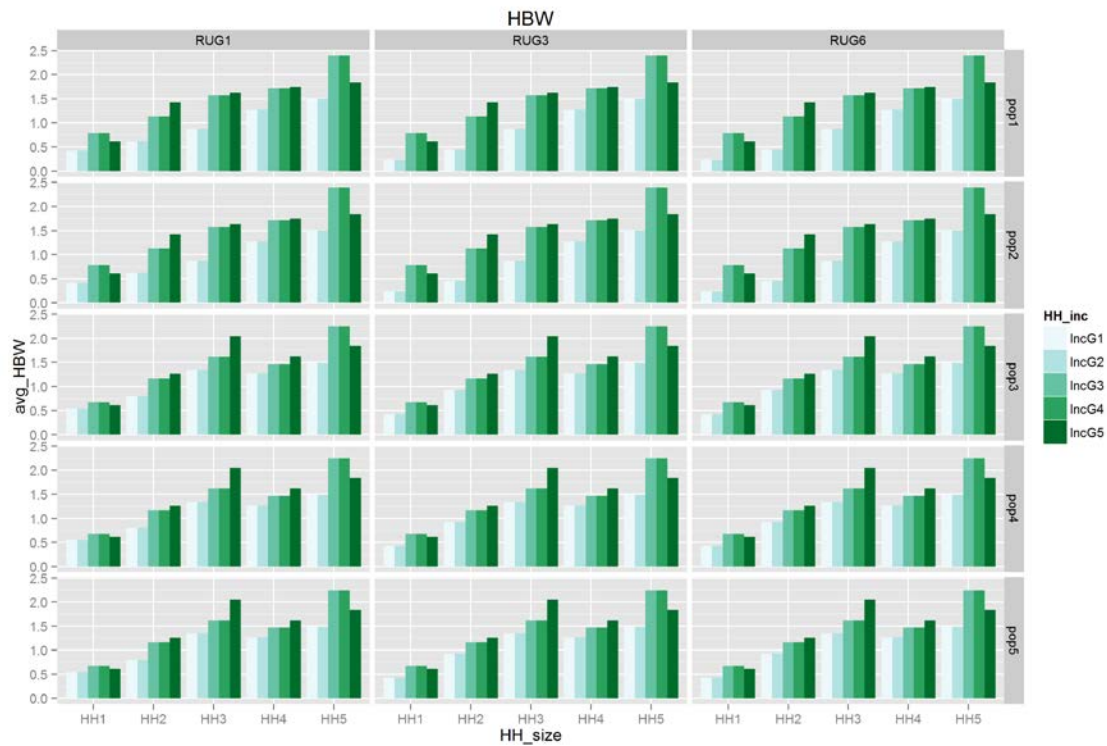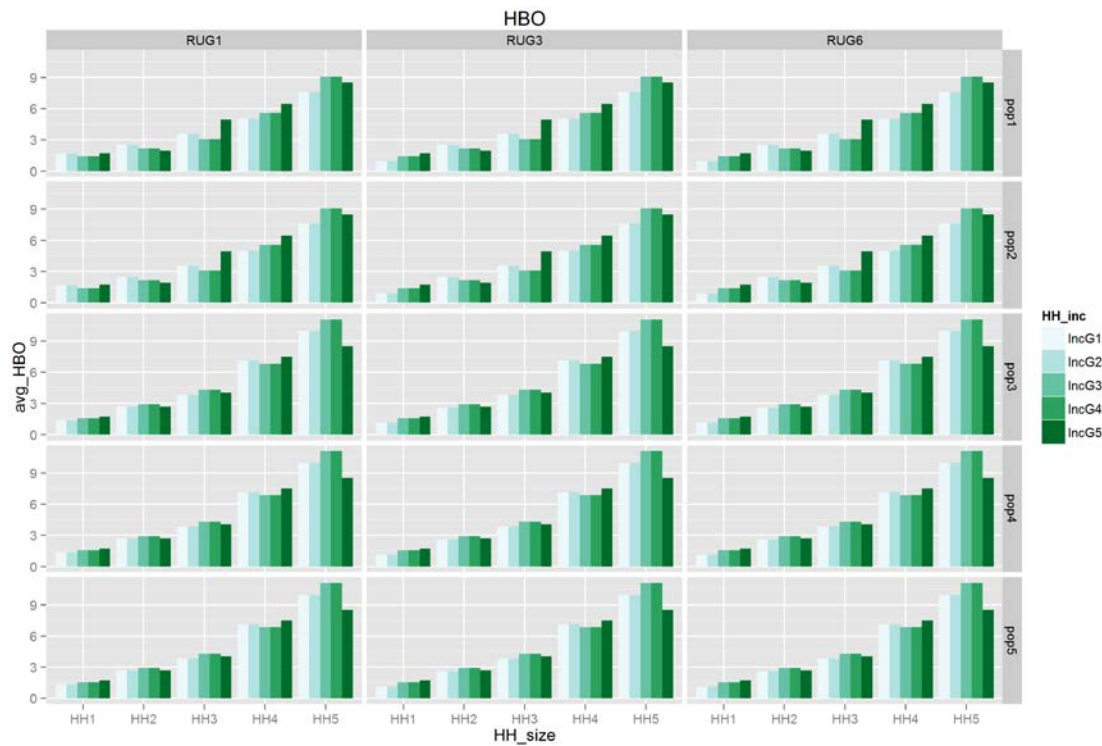
Home-Based Productions: School Purposes

The remaining home-based trip productions, HBK and HBC, were estimated using regression models. The units of analysis for these models were households, and the explanatory variables were the numbers of household members in various age categories: Age 0-4, Age 5-14, Age 15-17, Age 18-24, and Age 25-54.

Two separate models were estimated for each trip purpose, one for households in place types 1 and 2 (with fewer than 100,000 workers+jobs within a 30-minute auto trip); and one for households in place types 3, 4, and 5 (with more than 100,000 workers+jobs within a 30-minute auto trip).

The table below lists the resulting trip production rates per person in the age ranges specified. Note that while one might reasonably expect each child to make two school trips per day (to and from), the actual trip rates are somewhat lower: the survey includes days when individual children don't go to school due to school holidays or illness. Furthermore, if children make intermediate stops between school and home, the resulting trips will not appear as HBK trips in the household survey but rather as multiple trips (e.g., OBO and HBO).

TABLE 25: HBK AND HBC TRIP RATES (PER PERSON)

	HBK (Place Types 1 and 2)	HBK (Place Types 3 and 4)	HBC (Place Types 1 and 2)	HBC (Place Types 3 and 4)
Age 0-4	0.15	0.24		
Age 5-14	1.18	1.07		
Age 15-17	0.93	1.06		0.06
Age 18-24	0.07	0.11	0.23	0.24
Age 25-54			0.02	0.02

Attractions and Non-Home-Based Productions

Trip attractions, along with trip productions for non-home-based trips, were estimated using either ordinary linear regression models or partial linear regression models. Unlike ordinary linear regression, partial linear regression can be used even when explanatory variables are strongly correlated with one another. Because the VMIP 2 models include a large number of employment categories highly correlated with one another this model form resulted in more reasonable models than ordinary linear regression for some trip purposes.

Units of analysis for both kinds of regression models were groups of census tracts; the techniques used to group census tracts are described below. The explanatory variables for these models were the total number of jobs in each of the nine employment categories, school enrollment totals at the K-12 and university levels, and the total number of households. The table below lists the nine employment categories used:

TABLE 26: EMPLOYMENT CATEGORIES FOR VMIP 2 MODELS

Category	Description and NAICS code(s)
EMPEDU	Educational Services (61)
EMPFOO	Accommodation and Food Service (72), Art, Entertainment, and Recreation (71),
EMPAGR	Agriculture, Forestry, Fishing and Hunting (11)
EMPOTH	Mining (21), and Manufacturing (31-33)
EMPMED	Health Care and Social Assistance (62)
EMPIND	Utilities (22), Construction (23), Wholesale Trade (42), Transportation and Warehousing (48-49), Other Services (81)
EMPRET	Retail Trade (44-45)

TABLE 26: EMPLOYMENT CATEGORIES FOR VMIP 2 MODELS

Category	Description and NAICS code(s)
EMPOFC	Information (51), Finance and Insurance (52), Real Estate Rental and Leasing (53), Professional, Scientific, and Technical Services (54), Management of Companies and Enterprises (55), and Administrative and Support and Waste Management and Remediation Services (56)
EMPGOV	Public Administration (92)

The units of analysis for these regression models were defined using a combination of geography (census tracts, census designated places, or counties) and place type (as measured by jobs+workers within a 30-minute auto trip). A "rolling up" process was used where the smallest possible analytic units with sufficient sample size were used. Where census tracts attracted at least 50 trips of a given purpose, they were used as analytic units; otherwise census places or full counties, grouped by place type, were used instead.

Data for school enrollments was only available at the full county level. For the home-based school and home-based college trip purposes, this data was used with analytic units equal to counties, despite the fact that this resulted in models with very few analytic units. However, for other trip purposes which used school enrollments as explanatory variables, school enrollments were distributed among those census tracts which had HBK or HBC trip attractions. The countywide total of school enrollments was kept constant, with each tract receiving a portion commensurate with its HBK or HBC trip attractions. The result, while not as accurate as using enrollment data at the tract level, allows trip purposes such as HBO and WBO to have a larger number of analytic units and nevertheless use the school enrollment data.

The table below summarizes the number of analytic units used for each regression model, by trip purpose and attraction (A) versus production (P). For example, the 61 analytic units used for the HBW attractions model includes 6 individual census tracts (with sufficiently many work trips attracted to each), 34 subsets of census places with the same Place Type (e.g., Fresno, type 4; Stockton, type 3; Hanford type 2; Unincorporated Tulare County type 2), and 21 subsets of counties grouped by Place Type (e.g., Sacramento County, types 2 and 3 or San Joaquin County, type 2).

TABLE 27: GEOGRAPHIC UNITS USED IN MODEL ESTIMATION

Trip Purpose	Census Tracts	Census Places by Place Type	Counties by Place Type	Total
HBW (A)	6	34	21	61
HBK (A)	0	0	14	14
HBC (A)	0	0	0	14
HBS (A)	0	24	18	42
HBO (A)	32	78	14	124
WBO (P)	2	21	19	42
WBO (A)	1	20	18	39
OBO (P)	9	43	21	73
OBO (A)	10	47	18	75

Employment data used for model estimation was obtained from the EPA's Smart Location Database (SLD). The employment categories in the SLD do not fully match those in the model, so the model's Construction, Agricultural, and Industrial categories are combined; the resulting trip rate for the combined category is then applied to each of the three model categories. Additional explanatory variables tested include the number of households per tract, and the school enrollment per tract. School enrollment data was obtained from the California Department of Education (K12, public school enrollments only) and from the California Postsecondary Education Commission (college, public and private 2- and 4-year institutions).

All of the regression models estimated were either simple linear regressions with no intercept, or partial linear regressions with no intercept. In the case of non-home-based trips (WBO and OBO), the same variables were used for the production and the attraction models. **Table 28** lists the person trip rates estimated for each model. As an example of interpreting these models, the home-based other attraction model states that each retail, service, and public sector job will attract roughly 2 HBO trips, each K-12 school enrollment will attract roughly 1.5 HBO trips, and each household will attract roughly 1.1 HBO trips.

TABLE 28: ESTIMATED ATTRACTION AND NON-HOME BASED PRODUCTION MODELS

	HBW-A	HBS-A	HBK-A	HBC-A	HBO-A	WBO-P	WBO-A	OBO-P	OBO-A
AGR employment	1.17				0.34				
EDU employment	1.17								
FOO employment	1.17	2.15			1.25	0.12	0.12	8.19	7.66
GOV employment	1.17					0.07	0.09	0.16	0.22
IND employment	1.17				0.34				
MED employment	1.17				3.45	0.18	0.18	0.16	0.22
OFC employment	1.17				5.16	0.33	0.41	0.16	0.22
OTH employment	1.17				0.34				
RET employment	1.17	5.76			1.2	0.15	0.16	8.19	7.66
ELEM enrollment			1.1		0.66	0.8	0.76	0.14	0.05
HS enrollment			1.1		0.66	0.8	0.76	0.14	0.05
COLLEGE enrollment				0.35					
Total households					0.95				

HBW Segmentation by Household Income

Following trip generation, HBW trips were further segmented by household income. On the production side, this segmentation was already achieved by virtue of the fact that household income was one of the variables present in cross-classification. On the attraction side, HBW trip attractions for each employment category were separated into high, medium, and low income based on the percentages in the table below.

Proportion of II, IX, and XI Trips

Once the base trip production and attraction rates were established, trip productions for each TAZ were further segmented into II and IX trips, while trip attractions were further segmented into II and XI trips. This segmentation was calculated separately for each trip purpose and each CDP as described below. Note this segmentation simply describes the proportion of trips which enter or leave the county from each listed CDP; it does not govern the location of those trips, which is still determined by the trip distribution model.

First, all CHTS trip ends and households were associated with a CDP or were determined to fall in unincorporated areas. This process was made more complicated by the fact that the publicly-available version of the CHTS has all locations geocoded by census tract; however, census tract boundaries may not align with CDP boundaries, and each census tract may have multiple CDPs associated with it. In cases where multiple CDPs are associated with a single census tract, the CDP with the largest population in the tract (identified at the census block level) is used. If the largest population in the tract is outside all named CDPs, the tract is identified as an unincorporated portion of the relevant county. Note that some named CDPs are not the largest population center in any census tract, and thus do not appear in the summaries of CHTS data, having been aggregated into either neighboring CDPs or the unincorporated portion of the county.

Next, trip productions for each CDP and trip purpose were segmented into II and IX trips; while trip attractions were segmented into II and XI trips. In cases where the CHTS contains fewer than 30 trips for the place/purpose combination, the county-wide average II versus IX or II versus XI percentage was substituted.

TRIP DISTRIBUTION

The current gravity model trip distribution process and factors for each existing MPO model was mostly maintained for consistency. The required revisions are:

- Add friction factors for additional trip purposes resulting in the jobs housing relationship – segmenting by income level as well as by IX and XI parameters.

- Ensure friction factors for non-work trips do not screen out short trips which are likely candidates for non-motorized travel, particularly in models which have only used vehicle trip generation.

For models without mode choice components, the composite travel time will be estimated using walk time based on distance and an average of walk and drive time for origin-destination pairs where walk is competitive with auto. In addition, the sub-TAZ level of detail available in the GIS network will be used in combination with TAZ size.

The required revisions are listed below.

- Add friction factors for additional trip purposes and income group for home-work.

- Revise friction factors to be continuous and better match survey data.

- Adjust impedance inputs to be based on a composite of person travel times by all modes as well as travel costs, instead of just travel time by auto.

MODE CHOICE

In general, the mode choice functionality is the same as the VMIP 1 model. The primary changes to the mode choice model are listed below.

- The number of transit sub-modes in the model has been expanded from two to four. The prior Transit-Walk and Transit-Drive submodes have been replace with the following modes,

  o Transit-Walk-Bus

  o Transit-Walk-Rail (including the possibility of rail access via bus)

  o Transit-Drive-Bus

  o Transit-Drive-Rail (including the possibility of rail access via bus)

- In the current implementation, Transit-Walk-Bus and Transit-Walk-Rail are combined into a single mode prior to assignment; as are Transit-Drive-Bus and Transit-Drive Rail. This report recommends future model updates assign these modes separately, with the Rail submodes requiring the presence of at least one rail leg.

- Accessibility and built environment variables have been incorporated into the mode choice model.

The VMIP 2 mode choice model is segmented by trip purpose and vehicle availability, using three vehicle availability categories as described in the table below:

<p style="text-align:center">TABLE 29: VEHICLE AVAILABILITY SEGMENTS IN VMIP 2 MODE CHOICE MODELS</p>

Name	Description
0veh	Households which own no vehicles
1veh	Households which have one vehicle but more than one person
Others	Households with either one vehicle and one person, or more than one vehicle

The table below lists the modes available in the VMIP 2 models.

<p style="text-align:center">TABLE 30: MODES AVAILABLE IN VMIP 2 MODE CHOICE MODELS</p>

Category	Name	Segments Available	Trip Purposes	Description
Auto	da	1Veh, Other	All	Drive alone
	s2	All	All	Shared ride, 2 persons
	s3	All	All	Shared ride, 3+ persons
Transit	twb	All	All	Transit, walk-access, bus
	tdb	All	All	Transit, drive-access, bus
	twr	All	All but HBK, HBC	Transit, walk-access, rail
	tdr	All	All but HBK, HBC	Transit, drive-access, rail
	sb	All	HBK only	School bus
Active	walk	All	All	Walk
	bike	All	All	Bike

The variables used in each of the mode choice model segments are listed in the table below. Not all variables are used in all trip purposes models. For the accessibility and built environment variables, the table notes whether the variable is measured at the trip production (P) or trip attraction (A). Note that value of time is a direct consequence of the relationship between in-vehicle time and cost. As such, it is not estimated directly but is instead a consequence of the in-vehicle time (IVT) and cost coefficients. For model implementation purposes, only value of time (VOT) is used in the mode choice utility equation; for clarity, both are reported in the tables below.

TABLE 31: VARIABLES IN VMIP 2 MODE CHOICE MODELS

Variable	Purposes	Description
(Constants)	All	Alternative-specific constants
IVT	All	In-vehicle time
OVT	All	Out-of-vehicle time (access, transfer, egress, and waiting times)
Cost	All	Total cost, including auto operating cost, parking cost and tolls, and transit fares.
VOT	All	Value of time (conversion between cost variables and time variables)
TransitAccess	HBW, WBO, OBO	Jobs available within 30 minutes via transit, decay-weighted (P)
LogEmpDensity	HBW, HBS, HBO	Log (employment density of block group) (A)
IntDensity	HBK, HBC	Pedestrian-oriented intersection density (A)

The form of the VMIP 2 mode choice models is multinomial logit. A nested logit form might have been preferred for theoretical reasons, given the strong relationships among drive, transit, and active modes. However, no satisfactory nested logit models were estimated, likely because of severe constraints on the amount of transit data available. Multinomial logit models produced generally more sensible results and were used instead. Even the multinomial logit models produced some un-intuitive results. Rather than use un-intuitive coefficients, these were replaced by results from VMIP 1 mode choice models, pooled models involving multiple segments or multiple trip purposes, or were omitted altogether.

Home-Based Work

The table below lists model coefficients for HBW segments. Drive-alone was used as a reference mode for all segments, including the 0-vehicle segment where this mode is not permitted. In this segment, utility calculations were carried out without the drive alone mode.

TABLE 32: HBW MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	0	0
	s2	0.710	-1.839	-2.340
	s3	-0.229	-2.587	-2.936

TABLE 32: HBW MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	twb	-1.900	-1.602	-2.754
	tdb	-1.900	-1.602	0.000
	twr	-1.900	-4.173	-5.937
	tdr	-1.900	-0.444	-5.432
	bike	-2.438	-2.898	-3.763
	walk	1.477	0.030	-1.075
IVT	All	-0.035	-0.040	-0.040
OVT	All	-0.070	-0.080	-0.080
OVT/IVT	All	2	2	2
Cost	All	-0.003	-0.002	-0.001
VOT	All	6	10.055	18
	da	x	0	0
	s2	0.828	0.329	0.506
	s3	0.458	0.408	0.506
	twb	1.873	0.586	1.066
LogEmpDensity	tdb	1.873	0.586	1.066
	twr	1.202	0.850	1.202
	tdr	1.066	0.189	1.202
	bike	2.147	0.765	0.506
	walk	1.025	0.178	0.005
	da	0	0	0
	s2	0.013	0.013	0.005
	s3	0.013	0.013	0.005
TransitAccess	twb	0.158	0.027	0.032
	tdb	0.158	0.027	0.032
	twr	0.158	0.027	0.032

TABLE 32: HBW MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	tdr	0.158	0.027	0.032
	bike	0.136	0.031	0.062
	walk	0.136	0.031	0.062

Home-Based Shop

The table below lists model coefficients for HBS segments. Drive-alone was used as a reference mode for the 1-vehicle and 2-vehicle segments, while walk was used as a reference mode for the 0-vehicle segment.

TABLE 33: HBS MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	0	0
	s2	-3.420	-0.495	-0.889
	s3	-4.269	-0.380	-1.009
	twb	-2.439	-3.542	-5.834
	tdb	-2.439	-3.542	-5.834
	twr	-2.439	-3.542	-5.834
	tdr	-2.439	-3.542	-6.961
	bike	-5.341	-3.756	-2.972
	walk	0	2.191	-0.684
IVT	All	-0.025	-0.025	-0.025
OVT	All	-0.050	-0.050	-0.050
OVT/IVT	All	2	2	2
Cost	All	-0.005	-0.003	-0.002
VOT	All	3	6	6.319
LogEmpDensity	da	x	0	0

TABLE 33: HBS MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	s2	-0.040	0.297	0.161
	s3	0.957	0.026	0.161
	twb	0.732	0.916	1.141
	tdb	0.732	0.916	1.141
	twr	0.866	0.866	0.750
	tdr	0.866	0.866	0.750
	bike	1.274	1.171	0.594
	walk	0	0.190	0.458

Home-Based School (K-12)

The table below lists model coefficients for HBK segments. The reference mode for the 0- and 1-vehicle segments is walk; the reference mode for the 2-vehicle segment is shared-ride 3.

TABLE 34: HBK MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	-4.874	-2.110
	s2	-3.560	-1.710	-0.703
	s3	-3.115	-1.540	0
	twb	-0.887	-7.657	0.316
	tdb	-0.887	-7.657	0.316
	bike	-4.456	-4.456	-2.876
	walk	0	0	0.273
	sb	-1.198	-1.346	0.449
IVT	All	-0.025	-0.025	-0.025
OVT	All	-0.050	-0.050	-0.050

TABLE 34: HBK MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
OVT/IVT	All	2	2	2
Cost	All	-0.005	-0.003	-0.002
VOT	All	3	6	9
IntDensity	da	x	-0.004	0
	s2	0	-0.004	0.004
	s3	0	-0.004	-0.019
	twb	-0.019	0.003	0.004
	tdb	0	0	0
	bike	0.003	0.009	0.005
	walk	-0.008	0.000	0.005
	sb	-0.012	-0.004	-0.003

Home-Based College

The table below lists model coefficients for HBC segments. Because of the very small number of HBC trips in the household survey data, all vehicle ownership segments were pooled for model estimation purposes, with distinctions between segments left for adjustment during model calibration. Drive-alone was used as a reference mode. In the 0-vehicle segment, utility calculations were carried out without the drive alone mode.

TABLE 35: HBC MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	0	0
	s2	-2.230	-2.230	-2.230
	s3	-2.396	-2.396	-2.396
	twb	-0.521	-0.521	-0.521
	tdb	-0.521	-0.521	-0.521

TABLE 35: HBC MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	bike	-3.848	-3.848	-3.848
	walk	-1.126	-1.126	-1.126
IVT	All	-0.025	-0.025	-0.025
OVT	All	-0.050	-0.050	-0.050
OVT/IVT	All	2	2	2
Cost	All	-0.005	-0.003	-0.002
VOT	All	3	6	9
	da	x	0	0
	s2	-0.004	0.004	0.004
	s3	-0.004	-0.019	-0.019
IntDensity	twb	0.003	0.004	0.004
	tdb	0	0	0
	bike	0.009	0.005	0.005
	walk	0	0.005	0.005

Home-Based Other

The table below lists model coefficients for HBO segments. Drive-alone was used as a reference mode for the 2-vehicle segment, while walk was used as a reference mode for the 0- and 1-vehicle segments.

TABLE 36: HBO MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	da	x	-1.538	0
	s2	-3.032	-1.086	-0.151
Constant	s3	-3.354	-1.250	0.014
	twb	-4.518	-3.406	-3.174

TABLE 36: HBO MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
	tdb	-8.953	-5.947	-3.341
	twr	-6.684	-6.405	-7.221
	tdr	-6.684	-6.405	-7.221
	bike	-3.368	-3.596	-1.963
	walk	0	0	0.561
IVT	All	-0.025	-0.025	-0.025
OVT	All	-0.050	-0.050	-0.050
OVT/IVT	All	2	2	2
Cost	All	-0.005	-0.003	-0.002
VOT	All	3	6	9
	da	x	-0.455	0
	s2	-0.455	-0.455	0
	s3	-0.614	-0.614	0
	twb	0.387	0.277	0.315
LogEmpDensity	tdb	0.924	0.277	0.315
	twr	-0.407	0.277	0.363
	tdr	-0.407	0.277	0.363
	bike	-0.143	0.559	0.455
	walk	0	0	0.455

Work-Based Other

The table below lists model coefficients for WBO segments. Because of the small number of WBO, 0-vehicle household trips in the household survey data, the 0-vehicle and 1-vehicle segments were pooled for model estimation purposes, with distinctions between them left for adjustment during model calibration. Drive-alone was used as a reference mode. In the 0-vehicle segment, utility calculations were carried out without the drive alone mode.

TABLE 37: WBO MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	0	0
	s2	-1.226	-1.226	-1.308
	s3	-1.857	-1.857	-1.969
	twb	0.000	0.000	-2.453
	tdb	-4.305	-4.305	-2.453
	twr	-3.518	-3.518	-3.285
	tdr	-3.518	-3.518	-2.497
	bike	-3.424	-3.424	-5.431
	walk	-2.108	-2.108	-2.153
IVT	All	-0.035	-0.035	-0.030
OVT	All	-0.089	-0.089	-0.076
OVT/IVT	All	2.515	2.515	2.515
Cost	All	-0.004	-0.001	-0.001
VOT	All	6.076	16.618	18
TransitAccess	da	0	0	0
	s2	0	0	0
	s3	0	0	0
	twb	0.084	0.084	0.023
	tdb	0.084	0.084	0.023
	twr	0.144	0.144	0.062
	tdr	0.144	0.144	0.078
	bike	0.063	0.063	0.045
	walk	0.063	0.063	0.072

Other-Based Other

The table below lists model coefficients for OBO segments. Walk was used as a reference mode for the 0- and 1-vehicle segments; drive-alone was used as a reference mode for the 2-vehicle segment.

TABLE 38: OBO MODE CHOICE MODEL COEFFICIENTS

Variable	Mode	0-Vehicle	1-Vehicle, 2+ person HH	All Others
Constant	da	x	-0.732	0
	s2	-1.975	-0.223	-0.228
	s3	-2.353	-0.732	-0.388
	twb	-2.764	-3.899	-4.442
	tdb	-2.764	-3.899	-4.442
	twr	-4.017	-3.899	-5.409
	tdr	-4.017	-3.899	-5.409
	bike	-3.036	-4.219	-3.627
	walk	0	0	-0.444
IVT	All	-0.030	-0.030	-0.074
OVT	All	-0.061	-0.061	-0.147
OVT/IVT	All	2	2	2
Cost	All	-0.004	-0.003	-0.005
VOT	All	5.191	6	9
TransitAccess	da	x	-0.200	0
	s2	-0.200	-0.200	0
	s3	-0.369	-0.369	0
	twb	0.027	0.097	0.025
	tdb	0.027	0.097	0.025
	twr	0.027	0.097	0.025
	tdr	0.027	0.097	0.025
	bike	0.043	0.150	0.039
	walk	0	0	0.039

PRICING

The auto operating cost was updated based on the Big 4 MPO methodology. The change includes the non-fuel pricing, fuel cost and vehicle fleet determined for each individual county, and a constant price increase for fuel and non-fuel costs applied to forecast the future. More details are found in the memo from the Big 4 in Appendix K: Memo on Auto Operating Cost.

The household income and commute cost was also included in the model for the auto ownership. More details on this are included in the estimation section.

TRIP ASSIGNMENT

Trip assignment includes traffic and transit assignments.

Traffic Assignment

The traffic assignment process in each model was reviewed. During implementation of VMIP 1 it was noticed the addition of distance to the path assignment resulted in routes that did not reflect traffic counts or local knowledge. For VMIP 2, the traffic assignment method was modified to include congested travel time and link or node costs, removing distance.

To allow for a different value of time, traffic assignments by vehicle availability was implemented for a multi-class assignment which separately evaluates and reports the following five vehicle types:

- Drive Alone
- Drive Alone Toll
- Shared Ride 2
- Shared Ride 3+
- Truck

Traffic assignment was modified to remove distance from the path cost function, leaving time and pricing (converted to time using the value of time).

Transit Assignment

The transit assignment has not changed from VMIP 1 and includes the following variables:

- Transit networks, real or synthetic
- Transit attributes (mode, operator, vehicle type)

- Transit access links (coded into network? How does this work)

- Fares

- User classes (this needs to reflect types of MPO questions, such as sensitivity to fares or value of time)

- Transfer and wait rules

FEEDBACK LOOP

The feedback loop ensures the travel times used as input to trip distribution are consistent with the travel times on the final reported congested road network, as required for air quality conformity analysis. No changes were made during VMIP 2.

INTER-REGIONAL COORDINATION

In VMIP 1, each of the eight SJV counties used its own estimates of travel growth at the county boundaries and the proportions of through traffic. These forecasts of growth and through trips may be very different, even for adjacent counties, making it difficult to consistently identify inter-regional travel and possibly consolidate travel forecasts from multiple MPOs. The basis of the inter-regional coordination in VMIP 2 is the California Statewide Travel Demand Model (CSTDM), which provides a baseline distribution of passenger vehicle trips entering, leaving, or passing through each model area. The statewide model may not need to be re-run for every scenario run in a VMIP 2 model; the process illustrated and described below shows the decision process for whether the statewide model needs to be re-run.

Interregional Process
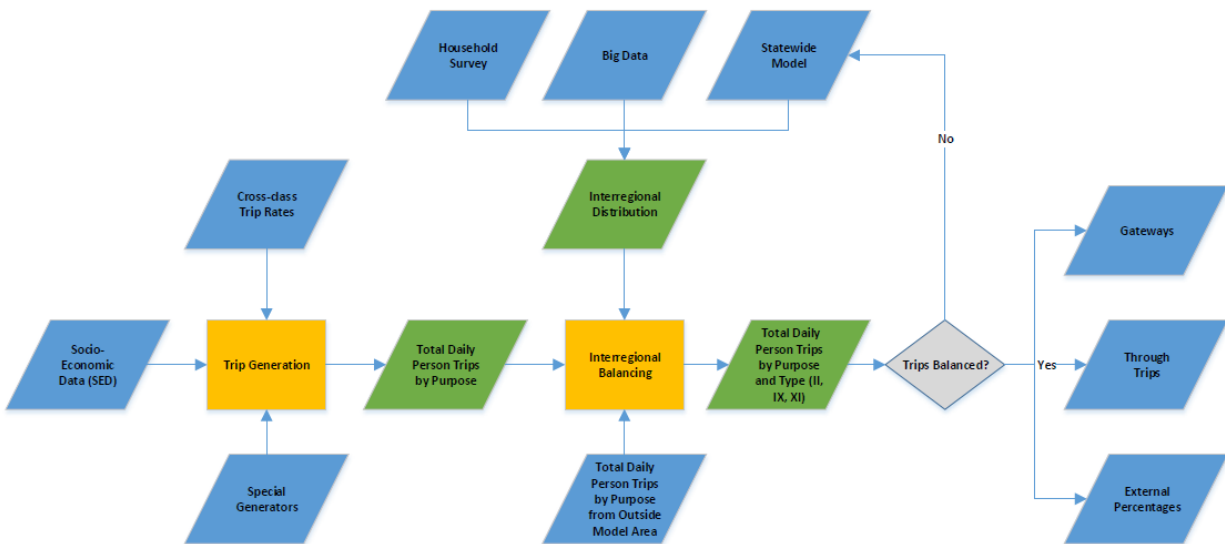
- First, trips internal to the model (ii trips) should be balanced to one another.

- Next, inter-regional trips produced and attracted to the model are compared to the number of IX and XI trips passing through model gateways. If balance can be achieved by re-distributing IX and XI trips among gateways, then there is no need to re-run the statewide model.

- However, if the number of IX trips produced by the model varies significantly from the number of IX trips attracted to gateways, or the number of XI trips attracted by the model is too different from the number of XI trips produced from gateways, then the statewide model must be re-run to account for land use changes which have changed inter-regional travel patterns.

The process outlined above was only partly implemented during VMIP 2 since the CSTDM has not been updated recently and does not include the land use developed for the RTP/SCS for any of the MPOs in the SJV. This report recommends that once the CSTDM (passenger) and California Statewide Freight Forecasting Model (CSFFM) are updated, new through trip tables are implemented in the model.

MODEL CALIBRATION

Calibration is an iterative process where model settings are adjusted so the output of the model matches observed travel patterns. Model calibration helps overcome issues of data quality, sample size, or aggregation bias and results in model outcomes tailored to local travel characteristics.

CALIBRATION TARGETS

The first calibration step is to verify the model is producing reasonable travel behavior across household dimensions:

- Household size

- Household income

A cross-classification comparison of the model outcomes and validation behavior for each of the household dimensions is prepared. The model is calibrated in an iterative method by reducing or increasing the 2012 ACS values until the household cross-classification totals from the model match the validation data source totals.

Model-Specific Calibration Targets

To verify that acceptable levels of calibration have been achieved, the model output for each step or submodel is compared to observed data. This comparison is referred to as validation.

- Vehicle availability was validated using census vehicle ownership cross-classified by household size and income.

- Trip generation was validated for trip productions, attractions, and trip balancing.

  - Trip production: A comparison of model total trips by purpose and observed totals from the expanded 2012 CHTS data. A secondary comparison, if needed, can be HBW trips from more aggregate sources such as the CTPP or NHTS. These sources are used with caution since they report "usual" workplace locations and are not directly comparable to model generated workplace locations. Convert person trip rates to ITE rates using Ave Veh Occ by purpose.

  - Trip attraction: Compare HBW attractions to total jobs in zone, range of 1.2-1.5 HBW attractions per employee in zone (source TFResource.org).

  - Trip balancing: PA totals, within +-10% of totals and totals by purpose.

- The trip distribution gravity model and any associated friction factors (k-factors) were calibrated iteratively to match average trip lengths by purpose and trip length frequencies by purpose are compared with the CHTS.

- The mode choice model was validated against CHTS mode shares.

The calibrated parameters used in the model are reported in Appendix L: Calibrated Parameters and summarized in the 1_Inputs\Support\ VMIP2_Kern COG_Parameters.xlsx.

MODEL STATIC VALIDATION

In the static validation tests, we ran the model to ensure the model output matches available traffic counts and ridership counts, and assessed the model's ability to replicate roadway speeds. This process starts with measuring the model traffic volume flows across screenlines composed of several roadways to ensure overall traffic flows in specific directions are accurately captured. Then, model volumes on individual links are compared to traffic counts. As part of the static validation procedure, elements of the trip generation, trip distribution, and traffic assignment modules were adjusted. Validation results are in the 01_Documents\Validation directory included with the model.

TRIP GENERATION

Trip generation validation consisted of the total production to attraction ratio (P/A) by purpose and the total trips generated per household. As we can see from the table, the P/A ratios are quite close to 1 for all the trip purposes and well within the 10% guideline. When applying the model for future years or land use scenarios, the P/A ratio should be reviewed along with the trips per household to ensure the model results reasonably reflect the scenario. The User Guide contains additional detail on checking the land use, trip balancing, and adjusting the inter-regional factors if needed.

TABLE 39: TRIP GENERATION – PRODUCTION (P)/ATTRACTION (A) BALANCE

Trip Purpose	Evaluation Criterion	Productions	Attractions	P/A Ratio	Difference	Percent Difference
HBW	+/- 10%	413,205	411,768	1.00	-1,437	-0.3%
HBS	+/- 10%	309,374	309,408	1.00	34	0.0%
HBO	+/- 10%	957,052	918,174	1.04	-38,878	-4.1%
NHB	+/- 10%	1,047,452	1,033,642	1.01	-13,810	-1.3%

TABLE 40: WEEKDAY PERSON TRIPS PER HOUSEHOLD

CHTS	Model
10.6	10.6

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P

VEHICLE AVAILABILITY

Next, we checked weekday person trips per household as shown in the table below. Again, the model output matches very closely with the data from the 2012 CHTS. Similarly, vehicle availability from the model as shown in the table below matches with the CHTS data.

TABLE 41: VEHICLE AVAILABILITY

0		1		2		3+	
CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model
7%	8%	31%	33%	39%	40%	22%	18%

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P

MODE SPLIT

When it comes to mode split by purpose, including modes such as drive alone, shared ride 2, transit and walking as well as purposes such as home based work (HBW) and non-home based work (NHB), outputs from the model are once again very close to the CHTS data.

TABLE 42: MODE SPLIT BY PURPOSE

Purpose	Total		Drove Alone		Shared Ride 2		Shared Ride 3+		Transit		Walk		Bike		Other	
	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model
HBW	14%	14%	82%	82%	9%	9%	3%	3%	2%	2$	3%	3%	1%	1%	0%	0%
HBO	56%	55%	28%	31%	21%	19%	32%	28%	1%	1%	15%	16%	1%	1%	2%	4%
NHB	30%	31%	38%	40%	26%	25%	29%	29%	1%	1%	4%	4%	0%	0%	1%	0%
Total	100%	100%	39%	41%	21%	20%	27%	25%	1%	1%	10%	10%	1%	1%	1%	2%

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P. Includes only internal-to-internal, weekday person trips for all modes. School bus trips are categorized as Other.

Model output for trip purposes by mode also falls close to the survey results as clearly shown in the table below. The transit data in the survey were 100% HBO which was not realistic, so the other modes were calibrated and the resulting transit mode in the model was retained rather than forcing 0% transit for other purposes. Additional on-board or similar surveys could be used to verify the results for transit.

TABLE 43: PURPOSE BY MODE

Purpose	Total		Drove Alone		Shared Ride 2		Shared Ride 3+		Transit		Walk		Bike	
	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model	CHTS	Model
HBW	14%	14%	29%	28%	6%	6%	1%	1%	26%	27%	4%	4%	19%	19%
HBO	56%	55%	41%	42%	56%	54%	66%	63%	50%	55%	83%	83%	71%	72%
NHB	30%	31%	30%	30%	38%	40%	33%	36%	24%	18%	13%	13%	10%	9%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P. Includes only internal-to-internal, weekday person trips for all modes. School bus trips are categorized as Other.

DISTRIBUTION – TRAVEL TIME

During the model estimation process the individual household survey records were evaluated. In many cases the reported travel time, level of congestion in the area, and travel distance were inconsistent for a given trip. Rather than using trip distance, the model uses travel time for distribution so future congestion or changes in travel time between modes influences overall travel. The results of the average travel time from the model are close to those observed, with the model being slightly lower than CHTS average times.

TABLE 44: TRIP ASSIGNMENT – AVERAGE TRAVEL TIME (IN MINUTES) BY TRIP PURPOSE

Trip Purpose					
HBW		HBO		NHB	
CHTS	Model	CHTS	Model	CHTS	Model
21.1	19.7	14.2	12.2	13.5	11.5

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P. Includes only internal-to-internal, weekday person trips for all modes.

VEHICLE MILES TRAVELED

Data from Highway Performance Management System (HPMS) were used as a benchmark for comparison of Vehicle Miles Traveled (VMT) within the model area. Although HPMS is an estimate of VMT based on sampled count data throughout the county, it is a standard method and a point of comparison often referenced especially for air quality analysis. The model is within rounding error (0% difference) compared to HPMS and is within the allowed deviation. Based upon VMT being within the estimate from HPMS combined with the travel time distribution and the lack of significant congestion within the region, the distribution portion of the model seems reasonable.

TABLE 45: TRIP ASSIGNMENT – VMT

Evaluation Criterion	HPMS	Model	% Deviation
+-3%	22,523,190	22,530,737	0%

Notes: Daily Vehicle Miles Traveled. Highway Performance Management System – 2014 California Public Road Data, Table 6.

DISTRIBUTION – INTER-REGIONAL TRAVEL

We also looked at model trip distribution and compared it with CHTS survey data. As shown in the table below, the model is close to the survey data for each trip type.

TABLE 46: TRIP DISTRIBUTION – BY PURPOSE (ALL MODES)

Trip Purpose								
Trip Type	**Total**		**HBW**		**HBO**		**NHB**	
	CHTS	**Model**	**CHTS**	**Model**	**CHTS**	**Model**	**CHTS**	**Model**
II	93%	94%	89%	90%	95%	97%	93%	92%
IX	3%	3%	5%	6%	2%	2%	3%	5%
XI	4%	2%	6%	4%	3%	1%	4%	2%

Notes: 2012 California Household Travel Survey, Weekday Trips, re-weighted by F&P. Includes only internal-to-internal, weekday person trips for all modes.

ROADWAY ASSIGNMENT – TRAFFIC VOLUMES

For the Kern COG model, weekday traffic counts were compared to the model assigned volume for total vehicle trips. Kern COG collects a large number of counts on county roads and city streets. Of the over 1,400 counts, nearly 900 were on roadways represented by the travel model. Due to the large number of construction projects, two separate static validation reports were prepared for roadway assignment. The first report includes all 900 counts noted above while the second report filters the count locations to remove those influenced by construction activity (two years for large projects opening and one year for smaller projects). **Table 47** summarizes the static validation tests for both sets of counts. In general, screening out counts influenced by construction activity improves the link level validation. The Assignment Validation Dashboard for All Counts and Filtered Counts are on the following pages.

TABLE 47: SUMMARY OF TRAFFIC ASSIGNMENT VALIDATION – DAILY CONDITIONS

Evaluation Criterion	Guidelines(1)	VMIP 1	VMIP 2 (all Counts)	VMIP 2 (filter Counts)
Number of count locations	N/A	542	938	610
Model/Count Ratio	+/- 10%	0.92	0.85	0.91
Percent within Caltrans Deviation	>75%	66%	57%	62%
Percent Root Mean Square Error	< 40%	45%	56%	44%
Correlation Coefficient	> 0.88	95%	92%	94%
Screenlines within Caltrans Deviation	100%	100%	58%	73%

Notes: (1) *2017 Regional Transportation Plan Guidelines for Metropolitan Planning Organizations*, California Transportation Commission, January 18, 2017 and *Travel Forecasting Guidelines*, State of California Department of Transportation, 1992.
Source: VMIP 1 data provided by Kern COG from the 7/17/13 model validation, with the Caltrans deviation based on all roadways. VMIP 2 data from Fehr & Peers based on the validation results in this document and the 2/3/2017 model run, with the Caltrans deviation based on principal arterials and higher.

The VMIP 2 model does not pass all of the static validation tests even after filtering the counts for construction activity. The model/count ratio shows the model tends to under-predict observed counts despite matching CHTS trip generation rates almost identically. This may suggest increases in trip rates could be justified since surveys may not fully capture existing travel behavior from busier households that are difficult to recruit, plus the potential to not report some short distance trips. The percent of links (and screenlines) within acceptable Caltrans deviations is also lower than the recommended guideline. The percentage root mean square error of 56% is higher than the recommended value of 40%, but this same statistic is reasonable for higher volume roadways above 25,000 as shown in the supplemental dashboards below.

The static validation results improve with the count filtering associated with construction activity. With so many significant construction projects during the time surrounding the validation year, the influence of construction on model validation was an important consideration. Most significant was the construction near SR 58, Truxton, and the Westside Parkway. The figures below show the construction projects and dates on the left with the model volumes on the right. The diversion of traffic away from construction areas, the magnitude of volume on each of the major facilities, and the number of count locations cause the skewing of validation statistics.

To evaluate the shift in traffic statistically, a screenline just west of SR 99 was developed. The allowed screenline deviation is 0.21 and the error in the numbers below is 0.06. As shown in the screenline table below, the further away from the construction area the more accurate the model (i.e., Stockdale is the furthest south and is good but the two roadways on either side are not). To rectify the count data, Kern COG could code the base year roadway network to reflect the 2014 conditions when counts were taken and update the future networks to reflect changes.

TABLE 48: SCREENLINE OF CONSTRUCTION INFLUENCE

CNT_ID	Location	Count	Model
10015	SR 58 (Rosedale Hwy)	54,626	70,273
2454	Truxton	43,866	21,376
3367	California	19.408	40,110
2451	Stockdale	25,933	21,388
Screenline Total		143,853	153,147

The model validation results demonstrate the model performs acceptably at a regional scale especially for key metrics such as VMT and higher volume roadways. At a local scale, sub-area refinements and validation should be performed before using the model for project applications. Refinements may include adding zonal or network detail to the model along with modifications to centroid loadings, network inputs (i.e., speeds), land use inputs, and demographic inputs. To help identify or target sub-regional areas requiring more refinements, users should review the map of daily validation locations. Any applications forecasts should also use an appropriate forecasting approach as described by National Cooperative Highway Research Program (NCHRP) Report 255 or 716 rather than using model forecast volumes directly.

INSERT HIGHWAY ASSIGNMENT DASHBOARD AFTER PDF IS CREATED

INSERT HIGHWAY ASSIGNMENT DASHBOARD AFTER PDF IS CREATED

The figures below show all count locations not within the Caltrans deviation as red, and locations within Caltrans deviation as green. Locations without counts are grey. Note that the figure shows all locations, not only those that are principal arterials and higher. All locations are shown as informational purposes only to assist with sub-area validation and application on development projects.
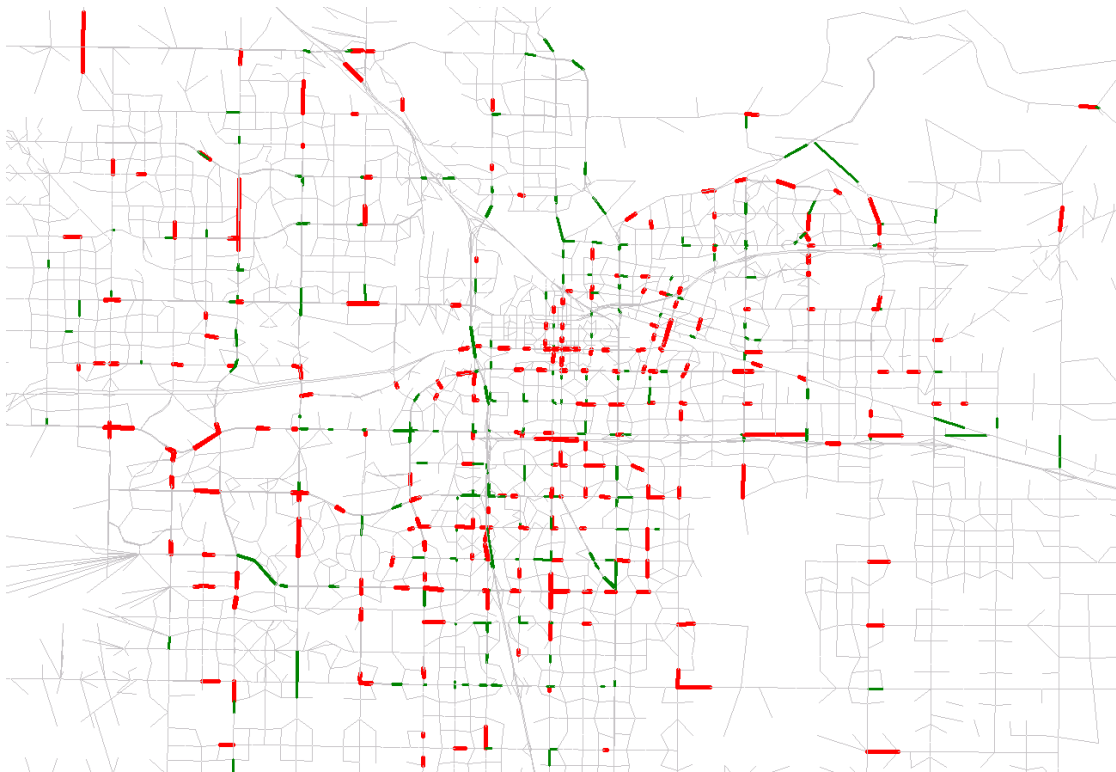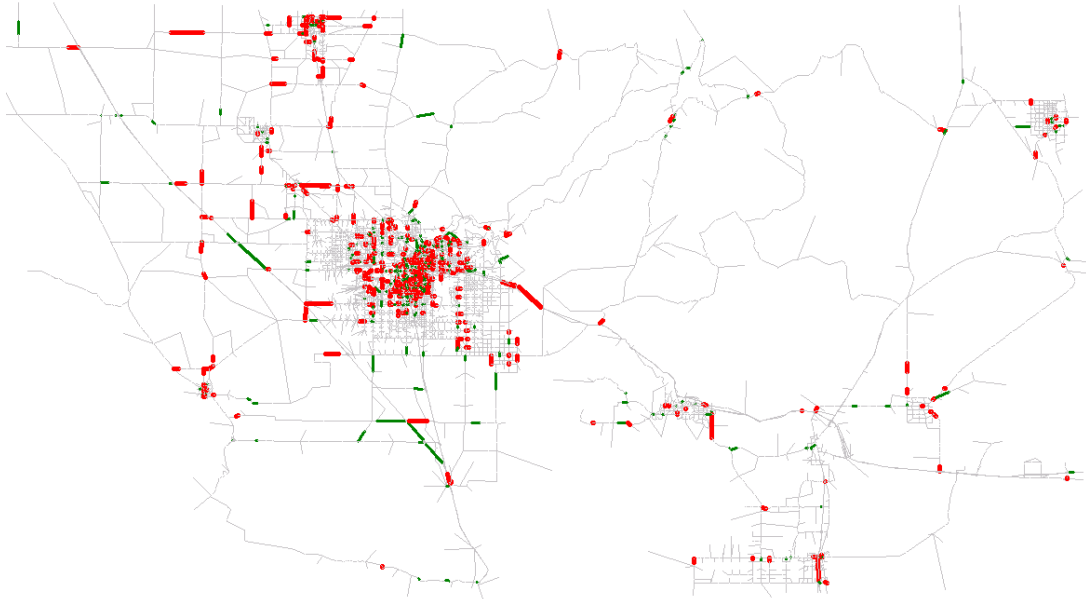
ROADWAY ASSIGNMENT – CONGESTED SPEED

In addition to roadway volume, traffic assignment results produce congested speed by time period. The congested speed is used within the model to determine auto distribution, influences bus in-vehicle time and assignment of auto and truck traffic to each route, and reporting VMT by speed range for air quality analysis. Performance based planning is also a requirement of the Fixing America's Surface Transportation (FAST) Act, with congestion, travel time, and reliability being key indicators. FHWA provides speed data for the National Highway System in the form of the National Performance Research Data Set (NPMRDS) for performance planning. This data set was used to compare the congested observed speeds with those from the travel model. The figure below compares the AM peak hour NPMRDS data to the model. The AM peak was selected since the model feedback loop uses the AM to determine distribution.

The model generally reflects the congested locations. The magnitude of congested speed is generally higher in the model due to the static assignment method not capturing physical capacity (i.e., allows volume to exceed capacity) and does not include intersection or lane change capacity/speed reductions. Overall, the model speeds are reasonable for a regional model application and for air quality analysis. For specific corridor projects, refinements to the speeds or implementing dynamic traffic assignment to account for travel and departure time and operational capacity not being exceeded are necessary.
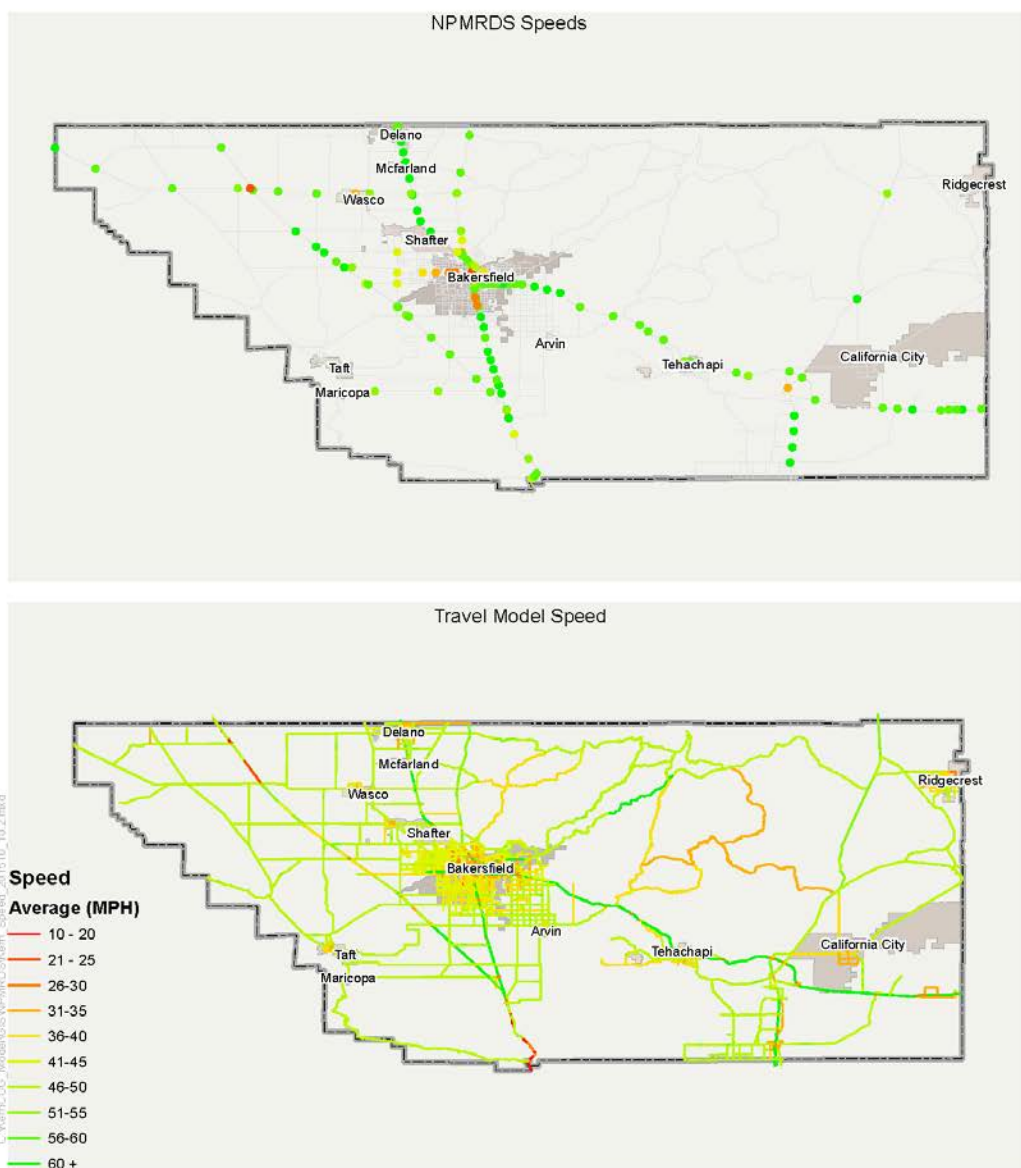
Figure 1

AM Peak Hour Speed Comparison

TRANSIT ASSIGNMENT – SYSTEM RIDERSHIP

As shown in the table below, the total transit system ridership is slightly high compared to the observed ridership. With transit mode share for transit less than 3%, minor differences in mode result in a noticeable difference in transit riders. The person trips per household match survey data, although surveys often don't fully represent all households, especially larger households with demanding schedules.

TABLE 49: DAILY TRANSIT ASSIGNMENT

Validation Statistic	Evaluation Criterion	Observed Ridership	Model Ridership	Percentage
Difference between actual ridership to model results for entire system	+/- 20%	27,035	30,059	11%

Notes: Observed Ridership includes Golden Empire Transit (GET), Kern Regional Transit (KRTS), and Delano Area Regional Transit average weekday unlinked trips for 2014

TRANSIT ASSIGNMENT – LINE RIDERSHIP

In addition to overall system ridership, the higher ridership routes are reported individually on the following page. The top 10 high volume routes account for approximately 67% of total system ridership. Routes with few routes per day or few days per week and dial-a-ride were excluded. The Travel Model Improvement Program (TMIP) does not have specific criteria for transit validation but does provide examples of what some regions across the country report relating to transit line ridership. The Florida State Urban Transportation Modeling System (FSUTMS line level guidelines by transit line ridership provided by TMIP were used for the validation of the Kern COG Model.

TABLE 50: DAILY TRANSIT ASSIGNMENT BY ROUTE

Route	Observed	Model	Acceptable[1]	Percent Difference
21	2,409	2,273	+/- 65%	5.6%
22	3,820	3,491	+/- 65%	8.6%
41	1,493	2,477	+/- 100%	-65.9%
42	1,036	4,643	+/- 100%	-348.2%
43	1,318	1,865	+/- 100%	-41.5%
44	1,995	2,207	+/- 100%	-10.6%
45	1,880	4,645	+/- 100%	-147.1%
46	835	992	+/- 150%	-18.8%
61	863	1,897	+/- 150%	-119.8%
81	699	175	+/- 150%	75.0%

Notes: 1. Acceptable criteria as reported by TMIP based on *FSUTMS-Cube Framework Phase II - Model Calibration and Validation Standards: Model Validation Guidelines and Standards*, prepared by Cambridge Systematics, Inc., for Florida Department of Transportation Systems Planning Office, December 31, 2007.

THROUGH TRIPS

Although the through trips have not been updated, enhancements to travel behavior within the model include more reasonable internal trip rates and estimates consistent with the 2012 CHTS. As discussed in the inter-regional coordination section, the CSTDM has not been updated to reflect the SJV MPO current RTPs. As such, the XX trips, derived from the CSTDM, were not adjusted upward. Further, XX truck trips in VMIP 2 were converted from passenger car equivalents (PCEs) to vehicles since the assignment accounts for PCEs and the counts (passenger vehicles plus trucks) are also in terms of vehicles. The volumes at the gateway are a combination of IXXI and XX and increasing either\both will increase VMT. Based on the comparison of count estimates to the Kern COG model and the level of confidence in each data source, Kern COG should decide which they have more confidence in and make adjustments as needed.

Given the model volumes being lower than counts presented in the highway validation tables above, the percentage of through trips was investigated. The through trips in the model were based on the CSTDM estimate of XX trips through the county. The CSTDM has a base year of 2008 and a forecast year of 2035. The model's XX trip estimation procedure uses the CSTDM data along with interpolation/extrapolation to produce XX trip tables for the various model years. The table below shows the base and future trips at I-5 gateways from the CSTDM. As shown in the table below, the 2008 CSTDM significantly over-estimates the volume at I-5 south gateway, resulting in a reduction of volume between 2008 and 2035 in the model. To correct this, the model development process used a growth rate of 2% applied to the count to replace the negative growth rate in the CSTDM. This growth rate should be reviewed for reasonableness given the lower overall population and employment growth rates in California since the 2008/09 recession. For I-5 north, the overall vehicle trips estimated increase and the through trips decrease as a percentage of total trips.

TABLE 51: CSTDM DAILY GATEWAY FORECAST

Gateway Location	2008 Count	2008 CSTDM			2035 CSTDM		
		Volume	XX Trips	XX %	Volume	XX Trips	XX %
I-5 South	67,000	98,326	10,989	11.2%	78,583	13,524	17.2%
I-5 North	27,000	24,343	10,093	41.5%	41,518	12,208	29.4%

The interpolated 2015 values from the CSTDM are shown in the table below and include the 2% growth rather than negative growth at the I-5 south gateway. As shown in the table, the 2015 Kern COG model volumes are low compared to 2015 counts and 2015 CSTDM. However, the numeric value and percentage of through trips in the Kern COG 2015 model are significantly higher than the CSTDM estimated 2015 values.

TABLE 52: CSTDM DAILY GATEWAY FORECAST

Gateway Location	2015 Count	2015 CSTDM Estimate			2015 Kern COG Model		
		Volume	XX Trips	XX %	Volume	XX Trips	XX %
I-5 South	70,043	73,625	14,879	20.2%	44,710	33,370	74.6%
I-5 North	31,157	30,688	8,976	29.2%	27,681	17,874	64.6%

Given the trip generation rate for the 2015 Kern COG model being low compared to expectations (i.e., matching 2012 CHTS but knowing that household surveys often underestimate trips per household), we investigated the IX/XI and II percentages compared to 2012 CHTS. As shown in Table 46, the percentage of IXXI trips is lower than 2012 CHTS targets by a few percent depending on purpose. With the current trip generation being low, if the trip generation rate was increased and the IXXI percentages were increased, the magnitude of trips internal to the model will stay roughly the same and the difference would exit the model on the major routes. Increasing trip generation and adjusting the IXXI factors would result in increased VMT and volume on the freeways, which would likely improve the freeway static validation but may result in increased VMT. The VMT is currently 0% deviation from HPMS, allowing room for increase while still being within the allowed deviation.

Kern COG may also want to investigate changing the current process for estimating XX trips originally developed for the VMIP 1 models. It may be more appropriate to take the direct XX trip tables from the CSTDM as the starting point for the Kern COG model XX trip tables and then adjust accordingly.

Since the gateway trips are a combination of XX trips and IXXI trips, we performed a test that increased the trips to/from the I-5 south gateway and the overall trip generation. See the table below for the summary of the results for the key gateway locations. With the 2012 CHTS often underreporting trip generation, the trip generation was increased from 10.6 trips per household (matching 2012 CHTS) to 11.0 trips per household. The through trips were based on the CSTDM and interpolated to 2015. The result was that gateway volumes increased along with VMT, which was 6% higher compared to the HPMS estimate. Further iterations of model refinement could improve the overall model performance and may be warranted if Kern COG has sufficient confidence the traffic counts or HPMS VMT estimate is accurate. The HPMS VMT value is an estimate while the traffic counts reflect observed data.

The Caltrans estimated volumes at gateways, HPMS, CHTS, and the CSTDM all have different levels of accuracy. When determining if more model refinements are needed, Kern COG should determine the level

of confidence associated with current model inputs/outputs and the validation data. This determination would help inform which components of the model are the most reasonable to adjust.

TABLE 53: SUMMARY OF KERN COG MODEL GATEWAY VOLUMES

Gateway	Location	Validated Model	Test Model
64	I-5 (N)	27,682	31,244
67	SR 43	2,923	2,994
69	SR 99	31,354	35,547
72	SR 65	5,285	5,908
29	SR 395 (N)	2,114	2,147
30	SR 178	1,533	1,590
31	US 395 (S)	2,049	2,111
32	SR 58 (E)	11,178	11,377
81	Lancaster Bl	1,775	1,840
33	Sierra Hwy	2,607	2,712
34	SR 14	11,665	12,108
37	I-5 (S)	44,711	53,968
39	SR 33 (S)	1,616	1,840
40	SR 58 (W)	117	120
41	SR 46	3,157	4,259
Screenline Total		149,766	169,765

APPENDIX A:

PREPARATION OF CALIFORNIA HOUSEHOLD TRAVEL SURVEY DATA

MEMORANDUM

Date: June 23, 2015

To: Users of CHTS data prepared by Fehr & Peers

From: Jennifer Ziebarth

Subject: Cleaning and Weighting of California Household Travel Survey Data

WC14-3115

The purpose of this memo is to document the steps undertaken to prepare the 2012 California Household Travel Survey (CHTS) for use in the Valley Model Improvement Program, Phase 2 (VMIP 2) project.

The 2012 CHTS is a statewide dataset of multi-modal travel behavior and household demographics. The survey includes data from a total of 42,431 households, collected using telephone surveys and GPS devices from all counties in California. The dataset includes travel patterns, including activity purpose, duration, travel distance, travel time, and mode choice. Demographics include household size, income, vehicle availability, and the additional characteristics of the individuals within the household.

Data preparation included the following steps:

1. Identify and repair unreasonable or missing trip distances.
2. Identify and consolidate transit trip chains.
3. Identify trip purposes.
4. Impute missing household income data.
5. Calculate a set of household-level weights to replace those provided with the CHTS.
6. Recode certain variables
7. Attach MPO and Census Designated Place information to trip and household records
8. Aggregate information about persons in the household to the household record
9. Attach person-level data to the trip records

TRIP DISTANCE CLEANING

The California Household Travel Survey provides trip distances in two formats: an "as-traveled" distance intended to be the actual distance traveled, and an "air distance" reflecting the straight-line distance between the trip's origin and destination. However, the as-traveled distance was missing from some trip records and was unreasonable in others. The graph below shows the relationship between air distance and as-traveled distance for all non-airplane trips in the CHTS. Trips whose as-traveled distance deviate too much from their air distance are candidates for providing a "cleaned" distance.
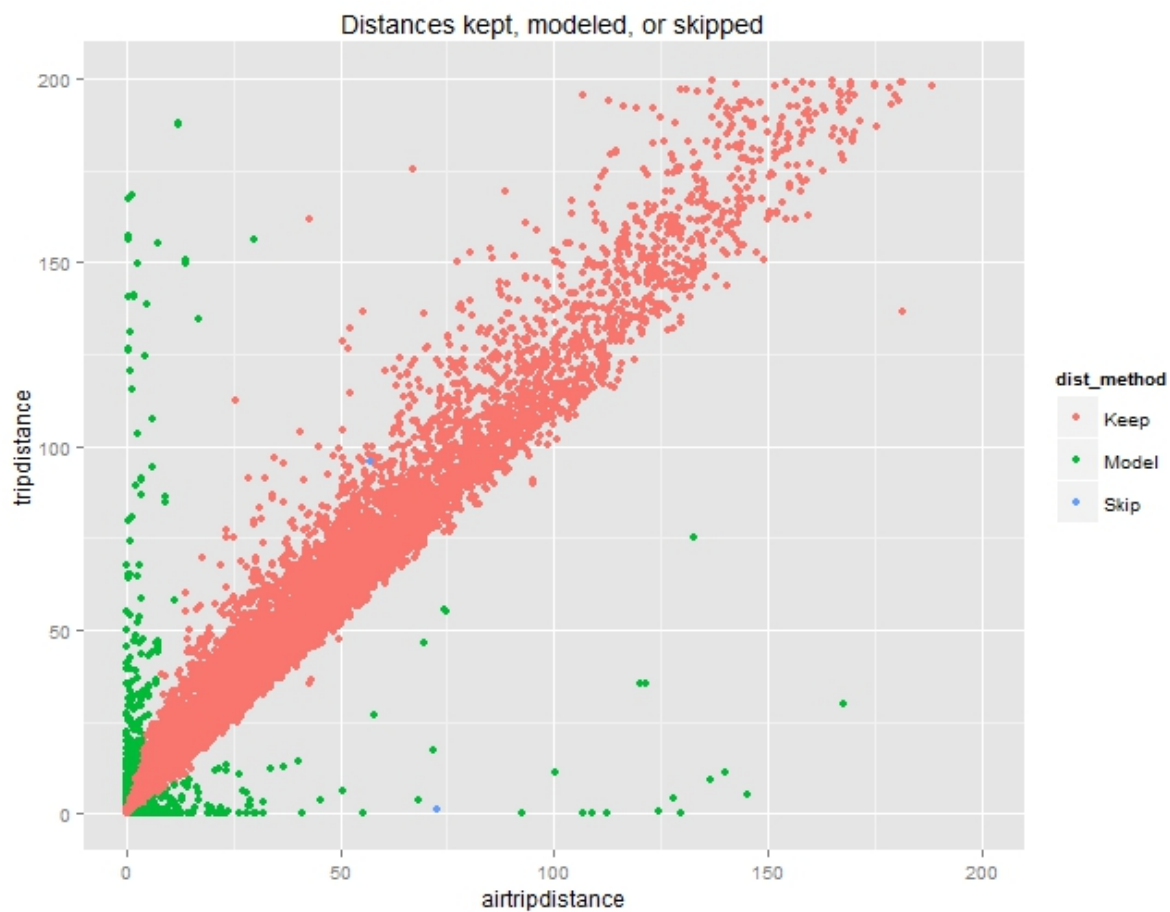
To provide "cleaned" trip distances, a simple linear regression was performed separately for each travel mode based only on the data where the as-traveled distance is deemed reasonable.

IDENTIFY TRIP PURPOSES

To identify trip purposes, both the activity purpose from the CHTS activities file and the place name from the CHTS places file were used. The activity codes provided in the CHTS data are as follows:

1. PERSONAL ACTIVITIES (SLEEPING, PERSONAL CARE, LEISURE, CHORES)

2. PREPARING MEALS/EATING

3. HOSTING VISITORS/ENTERTAINING GUESTS

4. EXERCISE (WITH OR WITHOUT EQUIPMENT)/PLAYING SPORTS

5. STUDY / SCHOOLWORK

6. WORK FOR PAY AT HOME USING TELECOMMUNICATIONS EQUIPMENT

7. USING COMPUTER/TELEPHONE/CELL OR SMART PHONE OR OTHER COMMUNICATIONS DEVICE FOR PERSONAL ACTIVITIES

8. ALL OTHER ACTIVITIES AT MY HOME

9. WORK/JOB DUTIES

10. TRAINING

11. MEALS AT WORK

12. WORK-SPONSORED SOCIAL ACTIVITIES (HOLIDAY OR BIRTHDAY CELEBRATIONS, ETC)

13. NON-WORK RELATED ACTIVITIES (SOCIAL CLUBS, ETC)

14. EXERCISE/SPORTS

15. VOLUNTEER WORK/ACTIVITIES

16. ALL OTHER WORK-RELATED ACTIVITIES AT MY WORK

17. IN SCHOOL/CLASSROOM/LABORATORY

18. MEALS AT SCHOOL/COLLEGE

19. AFTER SCHOOL OR NON-CLASS-RELATED SPORTS/PHYSICAL ACTIVITY

20. ALL OTHER AFTER SCHOOL OR NON-CLASS RELATED ACTIVITIES (LIBRARY, BAND REHEARSAL, CLUBS, ETC)

21. CHANGE TYPE OF TRANSPORTATION/TRANSFER (WALK TO BUS, WALK TO/FROM PARKED CAR)

22. PICKUP/DROP OFF PASSENGER(S)

23. DRIVE THROUGH MEALS (SNACKS, COFFEE, ETC.) [SHOW IF PTYPE <> 1 (HOME)]

24. DRIVE THROUGH OTHER (ATM, BANK) [SHOW IF PTYPE <> 1]

25. WORK-RELATED (MEETING, SALES CALL, DELIVERY)

26. SERVICE PRIVATE VEHICLE (GAS, OIL, LUBE, REPAIRS)

27. ROUTINE SHOPPING (GROCERIES, CLOTHING, CONVENIENCE STORE, HH MAINTENANCE)

28. SHOPPING FOR MAJOR PURCHASES OR SPECIALTY ITEMS (APPLIANCE, ELECTRONICS, NEW VEHICLE, MAJOR HH REPAIRS)

29. HOUSEHOLD ERRANDS (BANK, DRY CLEANING, ETC.)

30. PERSONAL BUSINESS (VISIT GOVERNMENT OFFICE, ATTORNEY, ACCOUNTANT)

31. EAT MEAL AT RESTAURANT/DINER

32. HEALTH CARE (DOCTOR, DENTIST, EYE CARE, HIROPRACTOR, VETERINARIAN)

33. CIVIC/RELIGIOUS ACTIVITIES

34. OUTDOOR EXERCISE (PLAYING SPORTS/JOGGING, BICYCLING, WALKING, WALKING THE DOG, ETC.)

35. INDOOR EXERCISE (GYM, YOGA, ETC.)

36. ENTERTAINMENT (MOVIES, WATCH SPORTS, ETC)

37. SOCIAL/VISIT FRIENDS/RELATIVES

38. OTHER (SPECIFY) [NOTE: LISTED ON DIARY] (O_APURP)

39. LOOP TRIP (FOR INTERVIEWER ONLY-NOT LISTED ON DIARY)

99. DONT KNOW/REFUSED

Each place visited was assigned a place based on the following criteria:

- If the place name is "HOME," then the place is "HOME," regardless of the activity purposes.

- If the place includes an activity with purpose code between 9 and 16, the place is "WORK."

- If the place includes an activity with purpose code between 17 and 20, then:

  o If the place name includes identifying strings such as "COLLEGE," "UNIV," "UCLA," or "USC," the place is "COLLEGE."

  o If the place name includes "PRESCHOOL" or "DAYCARE," the place is "OTHER".

- o   Otherwise the place is "K12."

- If the place includes an activity with purpose code 27 or 28, then the place is "SHOP."

- Otherwise, the place is "OTHER."

Once the purpose for each place has been determined, assigning a purpose to each trip is straightforward. For non-transit trips, the purpose at the trip origin is the purpose of the immediately preceding place record, and the purpose at the trip destination is the purpose of the place record itself. Then:

- If one end of the trip is "HOME" and the other is "WORK," the trip is home-based work ("HBW").

- If one end of the trip is "HOME" and the other is "K12," the trip is home-based K-12 ("HBK").

- If one end of the trip is "HOME" and the other is "COLLEGE," the trip is home-based college ("HBC").

- If one end of the trip is "HOME" and the other is "SHOP," the trip is home-based shop ("HBS").

- If one end of the trip is "HOME" and the other is either "OTHER" or "HOME," the trip is home-based other ("HBO").

- If one end of the trip is "WORK" and the other end is anything but "HOME," the trip is work-based other ("WBO").

- In all other cases, the trip is non-home-based ("NHB").

In some cases it is useful to consolidate these trips into a simpler scheme:

- Home-based work ("HBW") is the same as above.

- Home-based other ("HBO") includes "HBO," "HBK," "HBC," and "HBS" above.

- Non-home-based ("NHB") includes "WBO" and "NHB" above.

For transit trips, the purpose identification is slightly more complex and first requires identification of chains of transit trips (see below).

JOINT TRAVEL AMONG HOUSEHOLD MEMBERS

When multiple household members travel together in a single vehicle, the trip is considered a joint trip. Such trips are identified using arrival and departure times as well as person codes for household members on the trip. If the only purpose of the trip is to drop off or pick up household members, the trip is flagged as an escort trip.

This coding allows flexibility in how escort trips are counted when CHTS records are summarized. To avoid losing potentially important information, no trip purposes are changed.

IDENTIFY AND CONSOLIDATE TRANSIT TRIP CHAINS

In recording transit trips, the California Household Travel Survey treats each portion of the transit trip chain as a separate trip. For example, a trip in which the traveler drives to a rail station, takes the train to a second rail station, and then walks to a workplace is listed in the survey as three separate, consecutive trips, with three separate modes. This method of record-keeping makes it possible to track the mode of access and egress for a transit trip, but for most travel behavior analyses it is preferable to consider these three trips as a single unit. . Thus, a necessary step of data preparation is identification and consolidation of chains which make up a single linked transit trip.

To identify chains of transit trips, trips are flagged as transit access, transit egress, or transit transfer using the following criteria. A transit access trip is one which:

- Immediately precedes a trip whose mode is a transit mode,

- Does not itself use a transit mode, and either

  o Has an activity of "change to type of transportation / transfer" coded, or

  o Has an activity duration less than 30 minutes and a location whose name contains a keyword suggesting a transit stop, such as "station," "bus," "subway," etc.

- Does not end at the traveler's home.

A transit egress trip is one which:

- Immediately follows a trip whose mode is a transit mode,

- Does not itself use a transit mode, and either

  o Has an activity of "change to type of transportation / transfer" coded, or

  o Has an activity duration less than 30 minutes and a location whose name contains a keyword suggesting a transit stop, such as "station," "bus," "subway," etc.

- Does not depart from the traveler's home.

A trip which fits both sets of criteria, appearing to be both transit access and transit egress, is considered a transit transfer.

Once potential access, transfer, and egress trips have been identified, the first and last legs of transit trip chains are identified according to the following criteria. The first leg of a transit trip chain is one which:

- Is flagged as a transit access trip, or

- Is a transit trip whose preceding trip is not transit and does not have an activity of "change to type of transportation" coded, and whose previous activity duration is greater than 30 minutes.

The last leg of a transit trip chain is one which:

- Is flagged as a transit egress trip, or

- Is a transit trip which does not have an activity of "change to type of transportation" coded, whose following trip is not transit and whose activity duration is greater than 30 minutes.

Note the actual criteria are slightly more involved; for details see the R code. For validation of this process, it was confirmed no person has a different number of trips flagged as the first in a transit chain than flagged as the last in a transit chain.

Once transit trip chains have been identified, a trip purpose can be assigned to the chain as a whole. The chain origin is the origin for the first trip in the chain, that is, the purpose of the immediately preceding place. The chain destination is the destination for the final trip in the chain. The same categorization of trip purposes is used as described in the previous section.

COMPARISON OF TRIP MODES

The modes reported in the cleaned CHTS data are slightly simplified from those reported in the original CHTS data. In addition, mode categories in the cleaned CHTS data reflect vehicle occupancy of drive modes and mode of access for transit modes. The comparison between the original mode reported in the CHTS and the simplified mode in the cleaned data is as follows:

Simplified mode	Original modes
Walk	Walk; Wheelchair / Mobility Scooter Other Non-Motorized
Bike	Bike

Simplified mode	Original modes
Drive Alone	Auto / Van / Truck Driver Auto / Van / Truck Passenger Carpool / Vanpool Motorcycle / Scooter / Moped Rental Car / Vehicle
Drive Shared 2	Auto / Van / Truck Driver Auto / Van / Truck Passenger Carpool / Vanpool Motorcycle / Scooter / Moped Rental Car / Vehicle
Drive Shared 3	Auto / Van / Truck Driver Auto / Van / Truck Passenger Carpool / Vanpool Motorcycle / Scooter / Moped Rental Car / Vehicle
Drive Shared 4+	Auto / Van / Truck Driver Auto / Van / Truck Passenger Carpool / Vanpool Motorcycle / Scooter / Moped Rental Car / Vehicle
Taxi	Taxi / Hired Car / Limo
Shuttle	Private shuttle (SuperShuttle, employer, hotel, etc.) Other Private Transit
Walk to Bus	Greyhound Bus Local Bus, Rapid Bus Express Bus / Commuter Bus (AC Transbay, Golden Gate Transit, etc.) Premium Bus (Metro Orange / Silver Line) Public Transit Shuttle (DASH, Emery Go Round, etc.) AirBART / LAX FlyAway Amtrak Bus Other Bus
Drive to Bus	Greyhound Bus Local Bus, Rapid Bus Express Bus / Commuter Bus (AC Transbay, Golden Gate Transit, etc.) Premium Bus (Metro Orange / Silver Line) Public Transit Shuttle (DASH, Emery Go Round, etc.) AirBART / LAX FlyAway Amtrak Bus Other Bus

Simplified mode	Original modes
Walk to Rail	BART, Metro Red / Purple Line ACE, Amtrak, Caltrain, Coaster, Metrolink Metro Blue / Green / Gold Line, Muni Metro, Sacramento Light Rail, San Diego Sprinter / Trolley / Orange/Blue/Green, VTA Light Rail Street Car / Cable Car Other Rail
Drive to Rail	BART, Metro Red / Purple Line ACE, Amtrak, Caltrain, Coaster, Metrolink Metro Blue / Green / Gold Line, Muni Metro, Sacramento Light Rail, San Diego Sprinter / Trolley / Orange/Blue/Green, VTA Light Rail Street Car / Cable Car Other Rail
Walk to Ferry	Ferry / Boat
Drive to Ferry	Ferry / Boat
School Bus	School Bus
Paratransit	Dial-a-Ride / Paratransit (Access Services, etc.)
(removed from cleaned data)	Plane
NA	RF

IMPUTATION OF MISSING DATA

Although the household records are largely complete, certain key variables are missing for a small number of records. Variables used to estimate household weights (see next section) are imputed if they are missing. Additional variables were created to flag households whose data is imputed rather than reported in the original survey. The imputation process for these variables is described below.

HOUSEHOLD INCOME

Household income was not reported for 3,642 (8.6%) of households. For these households, the most likely income was calculated by comparing households of the same size, number of vehicles owned, and tenure type (own versus rent). The imputed household income is the average income category of the comparable households. For cases where fewer than ten households were considered comparable, households were grouped to provide a larger sample.

HOUSEHOLD RESIDENTIAL TYPE

The residential unit type was not available for 69 households (0.2% of the full CHTS). Residential unit type was imputed for these households by examining the residential unit types of households with the same size, number of vehicles owned, and household income category. The imputed residential unit type (single family, multi-family, or other) is set to be the most common residential unit type for matching households.

AGE OF HEAD OF HOUSEHOLD

Age of the head of household could not be determined for one household. This household was assumed to have a head in the age 25-64 category.

ESTIMATION OF SURVEY WEIGHTS

Surveys are meant to capture the characteristics of an entire population by randomly sampling a small proportion of the population. Often, a perfectly random sample is hard to achieve — some groups are difficult to survey and are under-represented, other groups are over-represented. To balance this bias, sample weights are estimated to "reshape" the sample. Fehr & Peers estimated household sample weights for the CHTS to balance the survey sample to match county-level percentages for several variables as reported in the 2012 American Community Survey 5-year estimates. Variables used as controls for the re-weighting are:

- Household size (one to seven or more)

- Household income (nine income categories)

- Number of workers per household (zero to three or more)

- Number of vehicles owned per household (zero to four or more)

- Household residential unit type (three categories)

- Household size (one to five or more) cross-classified by household income (five categories)

- Household size (one to five or more) cross-classified by number of vehicles per household (zero to four or more)

- Household size (one to five or more) cross-classified by number of workers per household (zero to three or more)

Counties were weighted either individually or, in the case of counties with fewer CHTS households, in groups of at most four adjacent counties weighted as a single unit. The multi-county groups used for weighting where single-county sample sizes were insufficient were:

- Lake and Mendocino Counties

- Del Norte, Siskiyou, Lassen, Modoc, Plumas, Sierra, and Nevada Counties

- Shasta, Tehama, Trinity, Glenn, and Colusa Counties

- Yolo, Yuba, and Sutter Counties

- Alpine, Amador, Calaveras, Mariposa, Tuolomne, Inyo, and Mono Counties

- Monterey and San Benito Counties

Expansion weights, suitable for expanding CHTS data to represent the full population of a county, were calculated for each county individually. Separate expansion weights exist for all households, and for households whose travel day is a weekday.

Weighting reports for each of the eight San Joaquin Valley counties is in the appendix to this memo.

ATTACH MPO AND CENSUS DESIGNATED PLACE INFORMATION

Fields are added to the household record listing the MPO and the Census Designated Place (CDP) of the household location; fields are added to the trip record listing the MPO and CDP of the trip origin and destination. Many MPOs in California are a single county; in this case, the MPO code is identical to the county FIP code. Multi-county MPOs are coded as follows:

1. AMBAG: Santa Cruz, Monterey, and San Benito Counties
2. MTC: Alameda, Contra Costa, Solano, Napa, Sonoma, Marin, San Francisco, San Mateo, and Santa Clara Counties
3. SACOG: Sacramento, Yolo, Yuba, Sutter, and portions of El Dorado and Placer counties
4. SCAG: Los Angeles, Ventura, Orange, Riverside, Imperial, and San Bernardino counties
5. TMPO: Portions of El Dorado and Placer counties

El Dorado and Placer counties are divided between two MPOs: the Tahoe Basin area lies in TMPO while the remainder of the counties are part of SACOG. Records are coded into the proper MPO using their census tract.

ATTACH PERSON DATA

A limited amount of data from the raw CHTS person file is attached to the final household and trip records. Demographic information such as the traveler's age, racial identity, worker, and student status is attached to the trip record. Fields indicating the number of household members in various age categories are added to the household record, along with a field indicating the age category of the head of household. The age categories used are:

- Age 0-2
- Age 3-4
- Age 5-14
- Age 15-17
- Age 18-24
- Age 25-34
- Age 35-44
- Age 45-54
- Age 55-64
- Age 65-74
- Age 75 and up

APPENDIX B:

CALIFORNIA HOUSEHOLD TRAVEL SURVEY DATA DICTIONARY

MEMORANDUM

Date: April 21, 2015

To: File

From: Jennifer Ziebarth

Subject: **Instructions for using CHTS cleaned data**

WC14-3115

The purpose of this memo is to provide instructions for using the cleaned and re-weighted California Household Travel Survey data. It includes data dictionaries for both the household and trip files, and important instructions regarding the use of household and trip weights.

JOINING THE HOUSEHOLD AND TRIP FILES

The "sampno" variable is a household ID code which can be used to join the household and trip files.

USING THE WEIGHTS

Please note that the CHTS data comes with survey weights which must be correctly applied to yield accurate summaries.

There are three types of weights included with the cleaned CHTS data:

- Household-level weights (hhweight and hhexpweight)

- Trip-level weights (tripweight and tripexpweight)

- Trip correction factor (tcf)

In order to use CHTS data accurately, one or more of these weights must be applied. The following instructions describe when to use each type of weight, and explain and give examples of using the weights.

DETERMINING WHICH WEIGHTS TO USE

To determine which weights to use, consider the following criteria:

- When summing or averaging values that pertain to households, use the household weights *hhweight* or *hhexpweight*. Examples include calculating the percentage of 0-vehicle households in a region, calculating the average number of licensed drivers per household, or calculating the number of households in a region with school-aged children. The *hhweight* weighting factor will weight households relative to one another and is useful for computing percentages, while the *hhexpweight* factor will also provide estimates of the total number of households.

- When summing or averaging values that pertain to trips from different households, use the trip weights *tripweight* or *tripexpweight*. Examples include calculating the average distance per vehicle trip, calculating mode shares, or calculating the distribution of travel times. As with the household weights, *tripweight* will weight trips relative to one another and is useful for computing percentages, while the *tripexpweight* factor will also provide estimates of the total number of trips.

- When summing or averaging values that pertain to trips within a single household, use the trip correction factor *tcf*. Often this is not done on its own but as the first of a two-step process; an example is calculating average VMT per household: first sum the VMT per household using the *tcf* weight, then average each household's VMT using either the *hhweight* or the *hhexpweight* weight. Similar two-step processes should be used to calculate the number of person-trips per household and the number of vehicle-trips per household.

- **When in doubt about which weight to use, please contact Jennifer Ziebarth.** I'm more than happy to help or to double-check that you've chosen the right weighting factor for your situation.

EXAMPLE 1: PROPORTION OF 2-OR-MORE VEHICLE HOUSEHOLDS

To calculate the proportion of households with two or more vehicles, sum the weights of households with two or more vehicles, then divide by the sum of all household weights. In equation form:

$$Proportion\ of\ 2-vehicle\ households = \frac{\sum_{2\ or\ more\ vehicle\ households}(household\ weight)}{\sum_{all\ households}(household\ weight)}$$

To do this in Excel, use the SUMIF and SUM functions:

The spreadsheet shows the formula:

`=SUMIF(K2:K34,">=2",P2:P34)/SUM(P2:P34)`

Code	placeName	ctfip	countyNam	MPOcode	MPOname	income	incomeImp	hhveh	hhbic	restype	restypeImp	headAge	hhweight	hhe
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	1	0	1	0	AGE75	0.177266	5;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	3	1	1	0	AGE2564	0.430407	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	3	0	3	0	1	0	AGE6574	0.633395	18
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	1	0	5	0	AGE2564	0.918663	2;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	1	0	0	0	1	0	AGE6574	0.337288	9;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	2	5	1	0	AGE2564	0.432	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	10	0	2	2	1	0	AGE2564	0.361505	1(
1144	Chowchilla	6039	Madera Co	6039	Madera Co	7	0	3	0	1	0	AGE2564	0.649022	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	3	0	2	3	1	0	AGE2564	0.326413	9(
1144	Chowchilla	6039	Madera Co	6039	Madera Co	6	0	2	0	1	0	AGE6574	0.260418	7;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	0	1	0	AGE6574	0.50301	;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	3	1	0	AGE2564	0.226261	(
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	1	1	0	AGE2564	0.22044	6;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	4	0	2	0	2	0	AGE2564	0.172157	5;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	4	0	1	0	1	0	AGE6574	0.495801	1z
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	4	8	1	0	AGE2564	0.330312	9;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	1	0	1	0	AGE2564	0.404656	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	4	0	2	0	1	0	AGE2564	0.172157	5;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	1	1	0	AGE2564	0.38262	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	2	1	0	AGE2564	0.576103	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	8	0	3	0	1	0	AGE2564	0.31765	9z
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	2	0	1	0	AGE2564	0.312371	9;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	2	0	1	0	1	0	AGE6574	0.50301	;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	3	0	3	2	1	0	AGE2564	0.247052	;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	1	1	0	1	0	AGE75	0.463271	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	6	0	3	5	1	0	AGE2564	0.293335	8(
1144	Chowchilla	6039	Madera Co	6039	Madera Co	7	0	1	0	1	0	AGE6574	0.730451	2;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	2	0	1	0	AGE2564	0.211476	6;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	5	0	1	0	1	0	AGE2564	0.233042	6;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	6	0	1	3	6	0	AGE2564	0.386914	1;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	1	0	3	0	1	0	AGE6574	0.351482	1(
1144	Chowchilla	6039	Madera Co	6039	Madera Co	9	0	2	1	1	0	AGE2564	0.24474	7;
1144	Chowchilla	6039	Madera Co	6039	Madera Co	3	0	3	0	1	0	AGE6574	0.633395	18
			Proportion of 2+ vehicle households:					=SUMIF(K2;						

To do this in R, use the sum function, identifying the subset of households with at least two vehicles in the numerator and all households in the denominator.

```
> prop_2plus <- sum(chowchilla$hhweight[chowchilla$hhveh>=2]) / sum(chowchilla$hhweight)
> prop_2plus
[1] 0.4930628
> |
```

EXAMPLE 2: AVERAGE TRIP DISTANCE

To calculate average trip distance for a collection of trips, sum the products of each trip distance multiplied by the trip weight, then divide by the sum of all trip weights. In equation form:

$$Average\ trip\ distance = \frac{\sum_{trips}(trip\ distance)*(trip\ weight)}{\sum_{trips}(trip\ weight)}$$

To do this in Excel, use the SUMPRODUCT and SUM functions:

Font | Alignment | Number | Styles | Cells

`=sumproduct(R2:R94,BA2:BA94)/sum(BA2:BA94)`

	D	E	F	G	H	Q	R	AB	AO	AZ	BA	BB	BC
	dTract	oPlace	oPlaceNam	dPlace	dPlaceNam	tripPurp	totalDist	modeString	age	tcf	tripweight	tripexpweight	
)1	300	1144	Chowchilla	1144	Chowchilla	NHB	5.664943	Drive Alone	62	1.085538	0.186883	67.43078	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	0.664659	Drive Alone	62	1.331198	0.229175	82.69053	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	0.926246	Drive Shared 2	62	1.331198	0.229175	82.69053	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	0.872687	Drive Alone	63	1.08895	0.187471	67.64272	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	0.872687	Drive Alone	63	1.08895	0.187471	67.64272	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	0.926246	Drive Shared 2	63	1.331198	0.229175	82.69053	
)0	300	1144	Chowchilla	1144	Chowchilla	HBS	0.64899	Drive Shared 2	60	2.174432	0.83198	300.1934	
)0	300	1144	Chowchilla	1144	Chowchilla	HBS	0.644164	Drive Shared 2	60	2.174432	0.83198	300.1934	
)0	300	1144	Chowchilla	1144	Chowchilla	HBS	0.64899	Drive Shared 2	58	2.174432	0.83198	300.1934	
)0	300	1144	Chowchilla	1144	Chowchilla	HBS	0.64899	Drive Shared 2	58	2.174432	0.83198	300.1934	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	2.994741	Walk	18	0.696995	0.401541	144.8832	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	2.994741	Drive Shared 2	18	0.988271	0.569346	205.4302	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	5.840582	Drive Alone	77	1.188906	0.377656	136.265	
)0	300	1144	Chowchilla	1144	Chowchilla	NHB	0.913314	Drive Shared 3	27	1.331198	0.328875	118.6639	
)0	300	1144	Chowchilla	1144	Chowchilla	HBO	1.999179	Drive Alone	27	1.331198	0.328875	118.6639	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	0.644646	Drive Alone	37	1.08895	0.319427	115.255	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	0.642672	Drive Shared 3	37	1.08895	0.319427	115.255	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	1.388131	Drive Shared 2	37	1.331198	0.390487	140.8946	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	1.388131	Drive Alone	37	1.331198	0.390487	140.8946	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	0.644646	Drive Shared 3	31	1.331198	0.390487	140.8946	
)0	300	1144	Chowchilla	1144	Chowchilla	NHB	2.387291	Drive Alone	31	1.331198	0.390487	140.8946	
)0	300	1144	Chowchilla	1144	Chowchilla	NHB	1.514051	Drive Alone	31	1.331198	0.390487	140.8946	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	1.388131	Drive Shared 2	31	1.331198	0.390487	140.8946	
)2	201	1144	Chowchilla	1144	Chowchilla	HBK	6.436982	School Bus	10	1.365993	0.400693	144.5773	
)1	300	1144	Chowchilla	1144	Chowchilla	NHB	5.901662	School Bus	10	0.988271	0.289894	104.5991	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	0.642672	Drive Shared 3	10	1.188906	0.348748	125.8344	
)2	202	1144	Chowchilla	1144	Chowchilla	HBK	2.771092	School Bus	8	1.208368	0.354457	127.8942	
)2	202	1144	Chowchilla	1144	Chowchilla	HBK	2.773391	School Bus	8	1.173979	0.344369	124.2544	
)2	300	1144	Chowchilla	1144	Chowchilla	HBO	1.388131	Drive Shared 2	8	1.08895	0.319427	115.255	
)0	202	1144	Chowchilla	1144	Chowchilla	HBO	1.388131	Drive Shared 2	8	1.08895	0.319427	115.255	
)2	300	1144	Chowchilla	1144	Chowchilla	HBW	2.674768	Drive Alone	65	1.079488	0.228286	82.3697	
)0	300	1144	Chowchilla	1144	Chowchilla	WBO	0.242534	Drive Alone	65	1.273109	0.269232	97.14383	
)0	202	1144	Chowchilla	1144	Chowchilla	HBS	2.490271	Drive Alone	65	1.273109	0.269232	97.14383	
)2	300	1144	Chowchilla	1144	Chowchilla	HBW	2.674768	Drive Alone	65	1.079488	0.228286	82.3697	
)0	202	1144	Chowchilla	1144	Chowchilla	HBW	3.015911	Drive Alone	65	1.079488	0.228286	82.3697	
)2	201	1144	Chowchilla	1144	Chowchilla	HBW	6.905395	Drive Alone	59	0.988271	0.208996	75.40944	
)1	202	1144	Chowchilla	1144	Chowchilla	HBW	6.905395	Drive Alone	59	1.188906	0.251425	90.71877	
				Average trip distance			`=sumproduct(R2:R94,BA2:BA94)/sum(BA2:BA94)`						

To do this in R, use the weighted.mean function:

```
> weighted.mean(chowchilla_ii_trips$totalDist,chowchilla_ii_trips$tripweight)
[1] 2.282369
>
```

EXAMPLE 3: VMT PER HOUSEHOLD

To calculate the average VMT per household requires working with both the trips and households data, and using two different weights at different steps of the process. Note the "sampno" variable is a household ID which can be used to join the household and trip data to each other.

The first step in calculating VMT per household is to find the sum of all vehicle trip distances for each household, using the trip correction factor as a weight. Note that to select vehicle trips you can select trips for which autoDriver=1; this will select each vehicle trip exactly once. The total VMT per household is the sum $VMT = \sum_{vehicle\ trips}(trip\ distance) * (tcf)$.

The second step in calculating VMT per household is to find the weighted average of all of the household VMTs just calculated. Because we're working per household, we need to use the household weights:

$$Average\ VMT\ per\ household = \frac{\sum_{households}(household\ VMT) * (household\ weight)}{\sum_{households}(household\ weight)}$$

DATA DICTIONARY: HOUSEHOLDS

The following table documents the variables in the cleaned household data file.

HOUSEHOLDS FILE DATA DICTIONARY

Variable	Description
sampno	Household ID
hctract	Census tract of household residence. A 10-digit ID which includes the county FIP as well as the census tract.
placeCode, placeName	Census Designated Place of household residence
ctfip,countyName	County of household residence
MPOcode, MPOname	MPO of household residence. Same as county for 1-county MPOs.
servicepop	Service population: Jobs + workers within 45 minutes by auto (time-decay-weighted)
income, incomeImputed	Household income category, flag for imputed data 1 = Less than $10,000 2 = $10,000 - $24,999 3 = $25,000 - $34,999 4 = $35,000 - $49,999 5 = $50,000 - $74,999 6 = $75,000 - $99,999 7 = $100,000 - $149,999 8 = $150,000 - $199,999 9 = $200,000 or more
hhsize	Number of household residents
hhemp, hhstu, hhlic	Number of household workers, students, driver's license holders
hhveh, hhbic	Number of vehicles and number of bicycles owned by household
restype, restypeImputed	Residential unit type, flag for imputed data

HOUSEHOLDS FILE DATA DICTIONARY

Variable	Description
headAge, headAgeImputed	Age category of HH head, flag for imputed data
tripMonth	Month of travel day
tripDay	Day of week for travel day
householdTrips	Total number of person-trips taken by household members on the travel day
Age0002, Age0304, Age0514, Age1517, Age1824, Age2534, Age3544, Age4554, Age5564, Age6574, Age75	The number of household residents in each age category
hhweight	Household weight
hhexpweight, hhexpweight_weekday	Household expansion weight for all households and for weekday subset of households

Data sources: 2012 CHTS household and person files, as cleaned and prepared by F&P; for details see the CHTS data preparation memo.

DATA DICTIONARY: TRIPS

The following table documents the variables in the cleaned trips data file.

TRIPS FILE DATA DICTIONARY

Variable	Description
sampno, perno	Household ID, person ID
chainno, numLegs	Trip chain ID, number of legs in trip chain
dep_hr, dep_min, arr_hr, arr_min	Time of trip departure & arrival (hour, minute)
tripPurp	Trip purpose (7 categories)
modeString	Trip mode (16 categories)
totalDist, totalTime	Total trip distance (miles) and time (minutes)
oTract, dTract	Census tract of trip origin and destination. (10-digit number, includes county FIP code)
pTract, aTract	Census tract of trip production and attraction
oPlace, oPlaceName, dPlace, dPlaceName	Census Designated Place of trip origin and destination
pPlace, pPlaceName, aPlace, aPlaceName	Census Designated Place of trip production and attraction

TRIPS FILE DATA DICTIONARY

Variable	Description
oFIP, oCountyName, dFIP, dCountyName	County of trip origin and destination
pFIP, pCountyName, aFIP, aCountyName	County of trip production and attraction
oMPO, oMPOname, dMPO, dMPOname	MPO of trip origin & destination (same as county for one-county MPOs)
pMPO, pMPOname, aMPO, aMPOname	MPO of trip production and attraction
oServicePop, dServicePop	Service population (jobs + workers within 45 minutes by auto, time-decay-weighted) at trip origin and destination
opurp, dpurp	Purpose recorded at trip origin and destination
opurp1,opurp2,opurp3,dpurp1,dpurp2,dpurp3	Detailed activity purpose codes at trip origin and destination
totalDist	Total trip distance (including transit access/egress)
accessDist, xferDist, egressDist	Transit access, transfer , egress distances
IVT, accessTime, xferTime, egressTime, waitTime	In-vehicle time, transit access, transfer, egress, and wait times
dwellTime	Time spent at trip destination
autoDriver	Flag for driver of auto trips
nonHHDriver	Flag for trips where the respondent is a passenger on a trip where a non-HH member is the driver
hhmem, nonhhmem	Count of HH and non-HH passengers on trip (not including the driver)
escortFlag	Flag for trip whose only discernable purpose is to escort another person
accMode, egrMode	Transit access and egress modes
accOcc, egrOcc	Vehicle occupancy of access and egress modes
age	Age of trip-maker
gender,ntvty, hisp,race,disab	Gender, nativity, Hispanic & racial identity, disability status of trip-maker
worker,student, schoolType	Worker & student status, and school type of trip-maker
license, transPass	Driver's license, transit pass status of trip-maker
tcf, tripweight	Trip correction factor , trip weight

Data sources: Data sources: 2012 CHTS person, place, and activity files, as cleaned and prepared by F&P; for details see the CHTS data preparation memo.

APPENDIX C:

SIMPLE SUMMARIES OF CHTS DATA

MEMORANDUM

Date: December 29, 2015

To: File

From: Jennifer Ziebarth

Subject: Data dictionary for CHTS simple summaries

WC14-3115

The purpose of this memo is to provide a data dictionary for the "simple" summaries of CHTS data. These summaries come in both Excel (.xlsx) and csv (.csv) formats. The summaries have one record for each geographic unit and are suitable for joining to a shapefile for visualization in GIS. The data summarized here includes the most commonly requested data from the CHTS including mode shares, trip purposes, trip distance, and trip time.

DATA DICTIONARY: CHTS SIMPLE SUMMARIES

Grouping	Variable	Description
Geography	geogCode, geogName, geogType, lookup	Code, name, and type of geography (e.g., state, county, MPO, or "place" (city or named place recognized by census). The lookup field is useful for creating VLOOKUPs in Excel, and helps to distinguish between cities and counties with the same name (e.g., Alameda_place is the city of Alameda; Alameda_county is the county.)

DATA DICTIONARY: CHTS SIMPLE SUMMARIES

Grouping	Variable	Description
Households, Trips, and Sample Sizes	HHsampleSize, PTsampleSize,VTsampleSize	Number of household, person-trip, and vehicle-trip records in the CHTS for this geography. CAUTION: If there are fewer than 100 households or trips for a geography, then the corresponding summaries should be used with caution. If there are fewer than 30 households for a given geography, it is excluded from this summary. Consult Jennifer Ziebarth for advice on how to proceed.
Households, Trips, and Sample Sizes	numHH, numPersonTrips, numVehTrips	The total number of households, person-trips, and vehicle trips represented by the CHTS for this geography.
Person-Trips per Household	PersonTrips_per_HH, PersonTrips_per_HH_HBW, PersonTrips_per_HH_HBO, PersonTrips_per_HH_NHB	The average number of person-trips per household, total and by trip purpose. Includes all travel modes, and all trips regardless of o/d.
Person-Trips per Household	PMT_per_HH, PMT_per_HH_HBW, PMT_per_HH_HBO, PMT_per_HH_NHB	The average number of person-miles traveled per household, total and by trip purpose. Includes all travel modes, and all trips regardless of o/d.
Person-Trips per Household	PHT_per_HH, PHT_per_HH_HBW, PHT_per_HH_HBO, PHT_per_HH_NHB	The average number of person-hours traveled per household, total and by trip purpose. Includes all travel modes, and all trips regardless of o/d.
Vehicle-Trips per Household	VehicleTrips_per_HH, VehicleTrips_per_HH_HBW, VehicleTrips_per_HH_HBO, VehicleTrips_per_HH_NHB	The average number of vehicle-trips per household, total and by trip purpose. Includes all trips regardless of o/d.
Vehicle-Trips per Household	VMT_per_HH, VMT_per_HH_HBW, VMT_per_HH_HBO, VMT_per_HH_NHB	The average number of vehicle-miles traveled per household, total and by trip purpose. Includes all trips regardless of o/d.
Vehicle-Trips per Household	VHT_per_HH, VHT_per_HH_HBW, VHT_per_HH_HBO, VHT_per_HH_NHB	The average number of vehicle-hours traveled per household, total and by trip purpose. Includes all trips regardless of o/d.

DATA DICTIONARY: CHTS SIMPLE SUMMARIES

Grouping	Variable	Description
Person-Trips per Household (ii only)	PersonTrips_per_HH_ii, PersonTrips_per_HH_HBW_ii, PersonTrips_per_HH_HBO_ii, PersonTrips_per_HH_NHB_ii	The average number of person-trips per household, total and by trip purpose. Includes all travel modes, but only trips *within the named geography*.
Person-Trips per Household (ii only)	PMT_per_HH_ii, PMT_per_HH_HBW_ii, PMT_per_HH_HBO_ii, PMT_per_HH_NHB_ii	The average number of person-miles traveled per household, total and by trip purpose. Includes all travel modes, but only trips *within the named geography*.
Person-Trips per Household (ii only)	PHT_per_HH_ii, PHT_per_HH_HBW_ii, PHT_per_HH_HBO_ii, PHT_per_HH_NHB_ii	The average number of person-hours traveled per household, total and by trip purpose. Includes all travel modes, but only trips *within the named geography*.
Vehicle-Trips per Household (ii only)	VehicleTrips_per_HH_ii, VehicleTrips_per_HH_HBW_ii, VehicleTrips_per_HH_HBO_ii, VehicleTrips_per_HH_NHB_ii	The average number of vehicle-trips per household, total and by trip purpose. Includes only trips *within the named geography*.
Vehicle-Trips per Household (ii only)	VMT_per_HH_ii, VMT_per_HH_HBW_ii, VMT_per_HH_HBO_ii, VMT_per_HH_NHB_ii	The average number of vehicle-miles traveled per household, total and by trip purpose. Includes only trips *within the named geography*.
Vehicle-Trips per Household (ii only)	VHT_per_HH_ii, VHT_per_HH_HBW_ii, VHT_per_HH_HBO_ii, VHT_per_HH_NHB_ii	The average number of vehicle-hours traveled per household, total and by trip purpose**.** Includes only trips *within the named geography*.
Person-Trip Distance by mode & purpose	PersonTrip_Avg_Distance_*mode-_purpose*	Average person-trip distance (miles) for each combination of mode and purpose. Includes ii trips (trips internal to the named geography) only.
Person-Trip Time by mode & purpose	PersonTrip_Avg_Time_*mode_purpose*	Average person-trip time (minutes) for each combination of mode and purpose. Includes ii trips (trips internal to the named geography) only.

DATA DICTIONARY: CHTS SIMPLE SUMMARIES

Grouping	Variable	Description
Daily mode shares	modeShare_*mode_purpose*	Average daily mode share for the listed mode within all trips of the listed purpose. If no purpose is listed, mode share is for trips of all purposes. Includes ii trips (trips internal to the named geography) only.
Peak period mode shares	modeShare_*mode_purpose*_peak	Average peak period mode share for the listed mode within all trips of the listed purpose. For purposes of this summary, peak period is defined as 6-9 AM and 4-7 PM. If no purpose is listed, mode share is for trips of all purposes. Includes ii trips (trips internal to the named geography) only.
Daily purpose shares	purpShare_*mode_purpose*	Average daily purpose share for the listed purpose within all trips of the listed mode. Includes ii trips (trips internal to the named geography) only.
Peak period purpose shares	purpShare_*mode_purpose*_peak	Average peak period purpose share for the listed purpose within all trips of the listed mode. For purposes of this summary, peak period is defined as 6-9 AM and 4-7 PM. Includes ii trips (trips internal to the named geography) only.
Direction Share	dirShare_*direction_purpose*	Average daily share of trips by direction: internal (ii), outgoing (ix), and incoming (xi), within all trips of the given purpose. If no purpose is listed, then share of trips by direction for all purposes combined.

Data sources: 2012 CHTS household, person, place, and activity files, with F&P modifications

Summarized using script MasterCHTSSummaries.R

APPENDIX D:
FLAT SUMMARIES OF CHTS DATA

MEMORANDUM

Date: April 22, 2015

To: File

From: Jennifer Ziebarth

Subject: Data dictionary for CHTS flat summaries

WC14-3115

The purpose of this memo is to provide a data dictionary for the "flat" summaries of CHTS data. These summaries come in both Excel (.xlsx) and csv (.csv) formats. The summaries have one record for each geographic unit and are suitable for joining to a shapefile for visualization in GIS.

DATA DICTIONARY: CHTS FLAT SUMMARIES

Grouping	Variable	Description
Geography	geogCode, geogName, geogType	Code, name, and type of geography (e.g., state, county, MPO, or "place" (city or named place recognized by census)
Number of Households and Trips	numHH, HHsampleSize, HH_Warning	Number of households represented by the CHTS for this geography, CHTS household sample size for this geography, and warning indicating whether data should be used with caution (*, 100 households or fewer) or used only when aggregated to include more households (**, 30 households or fewer).

DATA DICTIONARY: CHTS FLAT SUMMARIES

Grouping	Variable	Description
Number of Households and Trips	numVehTrips, VTsampleSize, vehTripWarning	Number of vehicle trips represented by the CHTS for this geography, CHTS vehicle trip sample size for this geography, and warning indicating whether data should be used with caution (*, 100 vehicle trips or fewer) or used only when aggregated to include more vehicle trips (**, 30 vehicle trips or fewer).
Number of Households and Trips	numPersonTrips, PTsampleSize, personTripWarning	Number of person trips represented by the CHTS for this geography, CHTS person trip sample size for this geography, and warning indicating whether data should be used with caution (*, 100 person trips or fewer) or used only when aggregated to include more person trips (**, 30 person trips or fewer).
Demographics	HH1, HH2, HH3, HH4, HH5, hhsize	Percentage of households with 1, 2, 3, 4, or 5+ members; average number of persons per household
Demographics	Veh0,Veh1,Veh2,Veh3,Veh4; hhveh	Percentage of households with 0,1,2,3, or 4+ autos; average number of vehicles per household
Demographics	Inc1, Inc2, Inc3, Inc4, Inc5, Inc6, Inc7, Inc8, Inc9	Percentage of households in each income category: 1. Less than $10,000 2. $10,000 to $24,999 3. $25,000 to $34,999 4. $35,000 to $49,999 5. $50,000 to $74,999 6. $75,000 to $99,999 7. $100,000 to $149,999 8. $150,000 to $199,999 9. $200,000 or more
Demographics	RUG1, RUG3, RUG6	Percentage of households by residential type. RUG1 = Single family; RUG3=Multi-family; RUG6 = Other (e.g., Mobile home, RV, boat)
Demographics	Age1824,Age2564,Age6574, Age75	Percentage of households by age category of household head

DATA DICTIONARY: CHTS FLAT SUMMARIES

Grouping	Variable	Description
Demographics	Pop0005, Pop0514, Pop1517, Pop1824, Pop2554, Pop5564, Pop6574, Pop75	Average number of residents per HH in each category
Household Summaries	VMT_per_HH_*purpose_mode*	Average VMT per Household by purpose and mode.
Household Summaries	VehicleTrips_per_HH_*purpose_mode*	Average Vehicle Trips per Household by purpose and mode
Household Summaries	PersonTrips_per_HH_*purpose_mode*	Average Person Trips per Household by purpose and Mode
Vehicle Trip Summaries	numVehTrips_*purpose_mode_distribution*	Total number of vehicle trips represented for each combination of purpose, mode, distribution
Vehicle Trip Summaries	vehDist_*purpose_mode_distribution*	Average vehicle trip distance for each combination of purpose, mode, distribution
Vehicle Trip Summaries	vehTime_*purpose_mode_distribution*	Average vehicle trip time for each combination of purpose, mode, distribution
Vehicle Trip Summaries	vehOcc_*purpose_mode_distribution*	Average vehicle occupancy for each combination of purpose, mode, distribution
Person Trip Summaries	numPersonTrips_*purpose_mode_distribution*	Total number of person trips represented for each combination of purpose, mode, distribution
Person Trip Summaries	PersDist_*purpose_mode_distribution*	Average person trip distance for each combination of purpose, mode, distribution
Person Trip Summaries	PersTime_*purpose_mode_distribution*	Average person trip time for each combination of purpose, mode, distribution

Data sources: 2012 CHTS household and person files, with F&P modifications

Summarized using script MasterCHTSSummaries.R

APPENDIX E:

FILTERABLE SUMMARIES OF CHTS DATA

MEMORANDUM

Date: December 29, 2015

To: File

From: Jennifer Ziebarth

Subject: Data dictionary for CHTS filterable summaries

WC14-3115

The purpose of this memo is to provide instructions for using the "filterable" summaries of CHTS data. Unlike the "flat" summaries, which are comparatively small in size, the "filterable" summaries allow for filtering based on multiple criteria, and as such they are quite large files. To simplify the summaries and allow for somewhat smaller file sizes, the filterable summaries are separated into two files, household summaries and trip summaries, which are described below.

INSTRUCTIONS AND HINTS

The filterable summaries allow CHTS data to be viewed by geography as well as selecting households or trips with certain demographic or travel profiles, such as households with two or more vehicles owned, or trips internal to the geography.

In most cases it is possible to select any combination of filter variables and see a summary of the relevant CHTS data. However, note that for some combinations the sample size of CHTS households, vehicle trips, or person trips may be quite small. Warning fields indicate whether the data can be used on its own, should be viewed with caution, or used only when aggregated with other data.

	Large enough sample size for confident reporting.
*	Use with caution: sample size may be not be large enough for statistical confidence.
**	Do not use in isolation. Sample size is too small for this result to stand on its own.

OTHER TIPS

- Non-vehicle modes such as bike, walk, or transit always have 0 vehicle trips per household in the household summaries, and 0 vehicle trips in the trip summaries, because these modes do not generate vehicle trips.

- Mode shares (and other "share" variables) are measured relative to mode= "All," with all other filters identical.

- Note that in some cases cities and counties share a name, so you may need to filter on both geogName and geogType to get the result you're looking for.

EXAMPLES

The examples below shows some of the tips above:

- The summary shows both the city of Tulare and the county of Tulare; the CHTS has 464 households in the county, but only 57 households in the city. Thus, summaries for the city should be used with caution.

- Vehicle trips, VMT, and VHT per household are 0 for all modes except the drive modes.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	\multicolumn{3}{Geography}			\multicolumn{8}{Filter variables}							\multicolumn{6}{Summaries per person-trip}					
	geogCode	geogName	geogType	HH size	HH vehicles	HH income	Trip purpose	Mode	Resident	Direction	Peak	Total Number of Person Trips	Person-trip Sample Size	Person Trip Warning	Person-trip Mode Share	Person-trip Purpose Share	Person-trip Resident Share	Person-trip Direction Share	Person-trip Distance Mean	Person-trip Time Mean
54	1797	Tulare	place	All	All	All	All	All	Res	ii	All	101,614	312		100%	100%	94%	83%	1.4	8.1
86	1797	Tulare	place	All	All	All	All	Bike	Res	ii	All	98	1	**	0%	100%	100%	100%	2.4	24.0
98	1797	Tulare	place	All	All	All	All	DriveAlon	Res	ii	All	34,581	132		34%	100%	90%	74%	1.7	9.0
33	1797	Tulare	place	All	All	All	All	DriveShar	Res	ii	All	31,815	80	*	31%	100%	100%	86%	1.4	8.0
65	1797	Tulare	place	All	All	All	All	DriveShar	Res	ii	All	21,865	56	*	22%	100%	93%	89%	1.2	5.4
94	1797	Tulare	place	All	All	All	All	Other	Res	ii	All	3,537	10	**	3%	100%	100%	100%	2.0	14.9
37	1797	Tulare	place	All	All	All	All	Walk	Res	ii	All	9,718	33	*	10%	100%	100%	99%	0.6	9.0
45	6107	Tulare	county	All	All	All	All	All	Res	ii	All	1,378,601	3,986		100%	100%	99%	96%	3.8	12.5
77	6107	Tulare	county	All	All	All	All	Bike	Res	ii	All	7,461	38	*	1%	100%	100%	100%	1.1	17.7
93	6107	Tulare	county	All	All	All	All	DriveAlon	Res	ii	All	513,362	1,729		37%	100%	99%	94%	4.8	12.8
29	6107	Tulare	county	All	All	All	All	DriveShar	Res	ii	All	337,192	1,042		24%	100%	99%	95%	4.2	11.5
65	6107	Tulare	county	All	All	All	All	DriveShar	Res	ii	All	325,740	734		24%	100%	100%	97%	3.3	11.3
98	6107	Tulare	county	All	All	All	All	Other	Res	ii	All	30,612	79	*	2%	100%	100%	95%	4.5	24.8
25	6107	Tulare	county	All	All	All	All	Transit	Res	ii	All	6,100	29	**	0%	100%	100%	86%	4.4	37.4
47	6107	Tulare	county	All	All	All	All	Walk	Res	ii	All	158,133	335		11%	100%	100%	100%	0.5	12.4

- All visible entries for "purpose share" are 100%, because trip purpose has been filtered to show all trip purposes combined ("All").

- Mode shares for rows where mode= "All" are 100%, while mode shares in other rows are smaller than 100%. The 34% mode share in the third row indicates that that row's mode ("Drive Alone") represents 34% of all person trips with the selected characteristics: In the city of Tulare, all household sizes, vehicles, and incomes, trips by residents only ("Res"), and only trips within Tulare ("ii").

- In many cases shown the number of households or trips is too small to draw any conclusions with the visible data. For example, the second row indicates the CHTS has only one weekday person trip, made by a resident of the city of Tulare, within that city, by bike. The red highlight serves as a warning that this single trip is not enough to draw wider conclusions.

DATA DICTIONARIES

DATA DICTIONARY: CHTS HOUSEHOLD FILTERABLE SUMMARIES

Type	Variable	Description
Geography	geogCode, geogName, geogType	Code, name, and type of geography (e.g., state, county, region/MPO, or "place" (city or named place recognized by census)
Filter	HH size	Household size : HH1=1, HH2=2, HH3=3, HH4=4, HH5=5 or more, HH4+ = 4 or more,
	HH vehicles	Number of vehicles owned by household: Veh0=0, Veh1=1, Veh2=2, Veh3=3, Veh4=4 or more, Veh2+ = 2 or more

DATA DICTIONARY: CHTS HOUSEHOLD FILTERABLE SUMMARIES

Type	Variable	Description
	HH income	Household income by category: Low = $0 - $49,999; Med = $50,000 - $99,999; High = $100,000 or more
	Trip purpose	Trip purpose, 3 categories (HBW, HBO, NHB). "HB" includes both HBW and NHB.
	Mode	Mode (Active, Drive Alone, Drive Shared 2, Drive Shared 3+, Transit, Other)
	Peak	All = All trips; Peak = 6-9am or 4-7pm; Offpeak = all other times
Summaries Per Household	HH total	Total number of households
	HH sample size	Number of CHTS household records
	HH Warning	Warning indicating whether data should be used with caution (*, 100 households or fewer) or used only when aggregated to include more households (**, 30 households or fewer).
	Person Trips per HH Mean	Average number of person trips per household
	PMT per HH Mean	Average Person Miles Traveled per household
	PHT per HH Mean	Average Person Hours Traveled per household
	Vehicle Trips per HH Mean	Average number of vehicle trips per household
	VMT per HH Mean	Average Vehicle Miles Traveled per household
	VHT per HH Mean	Average Vehicle Hours Traveled per household

Data sources: 2012 CHTS, as cleaned and summarized by Fehr & Peers

DATA DICTIONARY: CHTS TRIP FILTERABLE SUMMARIES

Type	Variable	Description
Geography	geogCode, geogName, geogType	Code, name, and type of geography (e.g., state, county, MPO, or "place" (city or named place recognized by census)
Filter	HH size	Household size : HH1=1, HH2=2, HH3=3, HH4=4, HH5=5 or more, HH4+ = 4 or more,

DATA DICTIONARY: CHTS TRIP FILTERABLE SUMMARIES

Type	Variable	Description
	HH vehicles	Number of vehicles owned by household: Veh0=0, Veh1=1, Veh2=2, Veh3=3, Veh4=4 or more, Veh2+ = 2 or more
	HH income	Household income by category: Low = $0 - $49,999; Med = $50,000 - $99,999; High = $100,000 or more
	Trip purpose	Trip purpose, 3 categories (HBW, HBO, NHB). "HB" includes both HBW and NHB.
	Mode	Mode (Active, Drive Alone, Drive Shared 2, Drive Shared 3+, Transit, Other)
	Resident	Restrict to residents of the listed geography? Res= Only residents; Non= Only non-residents; All = Both residents and non-residents
	Direction	Direction of trip, relative to the listed geography. ii =internal trip within the geography. ix = outgoing trip which starts inside and ends outside the geography. xi = incoming trip which begins outside and ends inside the geography.
	Peak	All = All trips; Peak = 6-9am or 4-7pm; Offpeak = all other times
Summaries per Vehicle Trip	Total Number of Vehicle Trips	Total number of vehicle trips
	Vehicle trip sample size	Number of CHTS vehicle trip records
	Vehicle Trip Warning	Warning indicating whether data should be used with caution (*, 100 vehicle trips or fewer) or used only when aggregated to include more vehicle trips (**, 30 vehicle trips or fewer).
	Vehicle Trip Mode Share, Vehicle Trip Purpose Share, Vehicle Trip Resident Share, Vehicle Trip Direction Share	Percent of vehicle trips with the current mode , purpose, residence status, or direction
	Vehicle Trip Distance Mean	Average vehicle trip distance
	Vehicle Trip Time Mean	Average vehicle trip time
	Vehicle Occupancy Mean	Average vehicle occupancy per vehicle trip
	Total Number of Person Trips	Total number of person trips
	Person Trip Sample Size	Number of CHTS person trip records

DATA DICTIONARY: CHTS TRIP FILTERABLE SUMMARIES

Type	Variable	Description
Summaries per Person Trip	Person Trip Warning	Warning indicating whether data should be used with caution (*, 100 person trips or fewer) or used only when aggregated to include more vehicle trips (**, 30 person trips or fewer).
	Person Trip Mode Share, Person Trip Purpose Share, Person Trip Resident Share, Person Trip Direction Share	Percent of person trips with the current mode , purpose, residence status, or direction
	Person Trip Distance Mean	Average person trip distance
	Person Trip Time Mean	Average person trip time

Data sources: 2012 CHTS, as cleaned and summarized by Fehr & Peers

APPENDIX F:

SIMPLIFIED CHTS DATA

MEMORANDUM

Date: October 7, 2015

To: File

From: Jennifer Ziebarth

Subject: How to use simplified CHTS data

WC14-3115

The purpose of this memo is to provide a data dictionary and instructions for using the simplified CHTS data (also known as "pivot summaries"). This data comes in .csv format and is intended to be further processed in Excel.

DATA DICTIONARY

The table below lists the variables present in the simplified CHTS data.

DATA DICTIONARY: SIMPLIFIED CHTS DATA

Grouping	Variables	Description
Location	oTract, dTract, homeTract, workTract	Census tract for trip origin, destination, home location, and (for respondents with a work trip on survey date) work location. Census tracts are listed as 10-digit state+county+tract FIPS code.
Location	oPlace, dPlace, homePlace, workPlace	Census Designated Place (e.g., city or other named place) for trip origin, destination, home location, and (for respondents with a work trip on survey date) work location.
Location	oFIP, dFIP, homeFIP, workFIP; oCounty, dCounty, homeCounty, workCounty	County (both FIPS code and name) for trip origin, destination, home location, and (for respondents with a work trip on survey date) work location.

DATA DICTIONARY: SIMPLIFIED CHTS DATA

Grouping	Variables	Description
Location	oRegion, dRegion, homeRegion, workRegion	Region for trip origin, destination, home location, and (for respondents with a work trip on survey date) work location. Regions are multi-county MPOs or other multi-county regions as listed below: • AMBAG: Monterey, San Benito, and Santa Cruz Counties • MTC: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma Counties • SACOG: El Dorado*, Placer*, Sacramento, Sutter, Yolo, and Yuba Counties, excluding Tahoe Basin area of El Dorado and Placer counties • SCAG: Imperial, Los Angeles, Orange, Riverside, San Bernardino, Ventura Counties • TMPO: Tahoe Basin area of El Dorado and Placer Counties • SJV: Fresno, Kern, Kings, Madera, Merced, San Joaquin, Stanislaus, and Tulare Counties • North: Butte, Colusa, Del Norte, Glenn, Humboldt, Lake, Lassen, Mendocino, Modoc, Nevada, Plumas, Shasta, Sierra, Siskiyou, Tehama, and Trinity Counties • Central Mountains: Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, and Tuolumne Counties • S Central Coast: San Luis Obispo and Santa Barbara Counties • SANDAG: San Diego County
Mode	Mode	One of the following travel modes: • DriveAlone, DriveShared • Bus, Rail, Ferry • Walk, Bike • Other (e.g., taxi, school bus, paratransit, …)
Purpose	Purpose	One of the following trip purposes: • HBW (home-based work) • HBO (home-based other) • NHB (non-home-based)
Distance	Distance	Total trip distance, rounded to the nearest mile. (Trips under half a mile are reported as distance 0). Note that trip distances in the survey are calculated from respondent's origin and destination, and the route used may not match the respondent's actual route.

DATA DICTIONARY: SIMPLIFIED CHTS DATA

Grouping	Variables	Description
Time	Time	Total trip time (including transit access/egress and waiting), rounded to the nearest 5 minutes. (Trips under 2.5 minutes are reported as time 0.) Note that trip times are self-reported by survey respondents.
Person-Trips	numPersTrips	Weighted and expanded number of person-trips for the given origin, destination, home, work, purpose, mode, distance, and time.
Person-Trips	rawPersTrips	Survey sample size for person-trips with the given origin, destination, home, work, purpose, mode, distance, and time.
Vehicle-Trips	numVehTrips	Weighted and expanded number of vehicle-trips for the given origin, destination, home, work, purpose, mode, distance, and time.
Vehicle-Trips	rawVehTrips	Survey sample size for vehicle-trips with the given origin, destination, home, work, purpose, mode, distance, and time.

Data sources: 2012 CHTS household and person files, with F&P modifications

Summarized using script ModeDistTime_PurposeDistrib.R

ON SURVEY WEIGHTING AND EXPANSION

The variables representing the number of person-trips and vehicle-trips are weighted and expanded to represent the total number of household-related trips of the listed type. While the survey is weighted to match household demographics (such as household size, household income, etc.) on a per-county basis, some limitations of the survey should be kept in mind when using the expanded number of trips.

- Because the CHTS is a **household** travel survey, it only measures travel related to (California) households. It does not measure commercial trips, trips made by visitors, or trips made by California residents who are not classified by the census as belonging to households – e.g., residents of group living quarters such as college dormitories, military bases, medical facilities, or correctional facilities.

- The survey weights supplied with the CHTS were judged to be insufficient for Fehr & Peers' purposes and we have therefore re-calculated weights in-house. For more information, see the CHTS data preparation memo or contact Jennifer Ziebarth.

USING THE SIMPLIFIED DATA

The simplified CHTS data is designed to be a flexible format which can produce the most commonly-requested summaries of CHTS data. Within Excel, this data can be filtered, summed, averaged, or brought into pivot tables and pivot charts to create a variety of summaries. Several common examples are detailed below. Two general comments may help you get started:

1. Because the CHTS is a weighted survey, you'll want to use the weighted variables numPersTrips and numVehTrips to count person-trips or vehicle-trips for almost any summary.

2. It's important to always confirm your summary is based on a large enough sample to provide reasonable representation of the population. For this reason, the sample sizes rawPersTrips and rawVehTrips are also provided. In general, caution should be used when summaries are based on less than 100 total (person- or vehicle-) trips; summaries based on a sample of less than 30 total trips should not be used alone, but should rather be pooled with additional data.

EXAMPLES OF COMMONLY REQUESTED SUMMARIES

MODE SHARE BY TRIP PURPOSE

To create a table of mode shares by trip purpose, start by confirming the CHTS has enough records to summarize the characteristics of interest. Create a pivot table with modes as rows, trip purposes as columns, and raw person-trips as values. In the Value Field Settings dialog, summarize values by Sum. Add filters to the pivot table to select other characteristics of interest such as residence or work location, origin, destination, etc. In the example below, we've selected records for respondents who live in Oakland and work in Walnut Creek.

homePlace	Oakland				
workPlace	Walnut Creek				
Sum of rawPersTrips	Column Labels				
Row Labels	HBO		HBW	NHB	Grand Total
DriveAlone	8		12	7	27
Rail			1	1	2
Walk	3			1	4
Grand Total	11		13	9	33

Unsurprisingly, there aren't very many trips in the CHTS with these characteristics, so we should expand our criteria. A good guideline for mode share summaries is at least 100 trips total, and at least 30 trips for each trip purpose.

Once we've confirmed the CHTS has enough responses with the characteristics of interest, create a second pivot table with the same rows, columns, and filters, and with number of person-trips as values. In the Value Field Settings dialog, summarize values by Sum, and show the values as percentage of column total.

| homePlace | Oakland | | | |
| workCounty | Contra Costa | | | |

Sum of numPersTrips	Column Labels			
Row Labels	HBO	HBW	NHB	Grand Total
DriveAlone	59%	84%	68%	70%
DriveShared	29%	0%	15%	14%
Rail	0%	16%	1%	5%
Walk	12%	0%	16%	11%
Grand Total	100%	100%	100%	100%

AVERAGE VEHICLE TRIP LENGTH

To estimate average vehicle-trip length, again start by confirming the CHTS has enough trips with the desired characteristics. Create a pivot table with raw vehicle trips (summarized by sum) in the value field, and any other desired characteristics in filters, rows and columns. Here, we see there are sufficient records for residents of all three AMBAG counties to allow summarizing vehicle trip length.

| homeRegion | AMBAG | | | |

Sum of rawVehTrips	Column Labels			
Row Labels	HBO	HBW	NHB	Grand Total
Monterey	1,597	827	997	3,421
San Benito	429	225	279	933
Santa Cruz	1,170	521	849	2,540
Grand Total	3,196	1,573	2,125	6,894

To determine average vehicle trip length by trip purpose, it's easier not to use a pivot table but to work with the relevant portion of the data directly. Set filters for the desired characteristics, and create a new column multiplying trip distance by the number of vehicle trips.

	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	homeRegic	workTr	workPl	workCc	workRe	mode	purpos	distanc	time	numPe	rawPer	numVe	rawVel	numVT*distance	
	AMBAG	NA	NA	NA	NA	DriveAlon	HBO	109	150	0.09	1	0.09	1	=W817*S817	
	AMBAG	6E+09	San Leand	Alameda	MTC	DriveShar	HBW	92	170	1.07	1	1.07	1		
	AMBAG	6E+09	Fremont	Alameda	MTC	DriveAlon	HBW	90	115	0.1	1	0.1	1		
	AMBAG	NA	NA	NA	NA	DriveShar	NHB	111	270	0.19	2	0.09	1		
	AMBAG	NA	NA	NA	NA	DriveAlon	HBO	333	340	0.06	1	0.06	1		
	AMBAG	NA	NA	NA	NA	DriveShar	NHB	84	110	0.18	1	0.18	1		
	AMBAG	NA	NA	NA	NA	DriveAlon	HBO	55	55	0.11	1	0.11	1		
	AMBAG	6.05E+09	Los Banos	Merced	SJV	DriveShar	NHB	2	5	1.36	2	0.67	1		
	AMBAG	6.05E+09	Los Banos	Merced	SJV	DriveShar	HBO	71	70	1.06	2	0.53	1		
	AMBAG	NA	NA	NA	NA	DriveShar	NHB	111	140	0.19	2	0.09	1		

Then, create sums for both the number of vehicle trips and vehicle trips * distance. Because we want to calculate average vehicle trip length for residents of the three AMBAG counties separately, SUMIF statements will help to sum only the values we're interested in.

								fx		=SUMIF($K1:$K195510,K$195618,$W1:$W195510)					

J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
homeP	homeC	homeRegic	workTr	workPl	workCc	workRe	mode	purpos	distanc	time	numPe	rawPer	numVe	rawVel	numVT*dist
Unincorpc	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveShar	NHB	87	170	1.01	2	0.52	1	45.24
Santa Cru:	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveAlon	NHB	90	120	0.05	1	0.05	1	4.5
Soquel	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveShar	NHB	41	70	0.66	2	0.33	1	13.53
Santa Cru:	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveShar	NHB	168	360	0.02	2	0.01	1	1.68
Felton	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveShar	NHB	88	30	0.91	1	0.91	1	80.08
Felton	Santa Cru:	AMBAG	6.8E+09	Other US	Other US	Other US	DriveShar	NHB	1	15	0.86	1	0.86	1	0.86
Felton	Santa Cru:	AMBAG	6.8E+09	Other US	Other US	Other US	DriveShar	NHB	1	30	0.69	1	0.69	1	0.69
Rio del Mi	Santa Cru:	AMBAG	NA	NA	NA	NA	DriveShar	NHB	163	240	0.91	4	0.25	1	40.75
Total	Monterey												=SUMIF($		
Total	San Benito														
Total	Santa Cruz														
Total	AMBAG														

Finally, divide the sum of vehicle trips * distance by the sum of vehicle trips, and you have the average vehicle trip distance. Note that this process is creating a weighted average of the trip distance, using the number of vehicle trips as a weight.

J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
homeP	homeC												numVe	rawVel	numVT*distance	
7 Felton	Santa Cru:	A	N	N	N	N	D	N	#	#	1	1	0.91	1	80.08	
8 Felton	Santa Cru:	A	#	O	O	O	D	N	1	#	1	1	0.86	1	0.86	
9 Felton	Santa Cru:	A	#	O	O	O	D	N	1	#	1	1	0.69	1	0.69	
0 Rio del Ma	Santa Cru:	A	N	N	N	N	D	N	#	#	1	4	0.25	1	40.75	
															Average Vehicle Trip Distance	
7																
8 Total	Monterey												1685.18		13083.81	7.76
9 Total	San Benito												250.92		2686.28	10.71
0 Total	Santa Cruz												1287.31		9913.47	7.70
1 Total	AMBAG												3223.41		25683.56	7.97
2																

O/D TABLE

To create an O/D table for a set of geographies, again start by setting up a pivot table with the desired filters, with origins as rows, destinations as columns, and raw trips (either person- or vehicle-trips) as value; this will help you to confirm whether sample sizes are sufficient.

oRegion	SJV							
dRegion	SJV							
Sum of rawVehTrips	Column Labels							
Row Labels	Fresno	Kern	Kings	Madera	Merced	San Joaquin	Stanislaus	Tulare	Grand Total
Fresno	3,576	8	51	107	20	1	3	100	3,866
Kern	11	4,024	1		1	1	1	38	4,077
Kings	55	2	798	1				43	899
Madera	110	1	1	633	18	-	6		769
Merced	17	1		19	1,354	6	85	1	1,483
San Joaquin	2			2	7	2,076	104		2,191
Stanislaus	2	1		4	84	104	1,602	2	1,799
Tulare	99	33	46	4		1	1	2,519	2,703
Grand Total	3,872	4,070	897	770	1,484	2,189	1,802	2,703	17,787

In this example, overall we have plenty of vehicle trips to summarize, but for the pairs with a small number of survey records we shouldn't draw any conclusions beyond the obvious one that these pairs don't experience as much interaction as other pairs.

Create a second pivot table with the same rows, columns, and filters, and with number of trips as values. To help distinguish cells with enough sample size to draw conclusions, cells with sufficient sample size are highlighted in green in the example below.

oRegion	SJV							
dRegion	SJV							
Sum of numVehTrips	Column Labels							
Row Labels	Fresno	Kern	Kings	Madera	Merced	San Joaquin	Stanislaus	Tulare	Grand Total
Fresno	1,716,778	1,962	13,634	18,028	8,853	266	1,077	22,169	1,782,766
Kern	2,265	1,439,497	162		448	162	211	9,538	1,452,284
Kings	14,181	470	215,434	269				8,006	238,360
Madera	20,314	330	269	165,030	3,725	-	1,463		191,130
Merced	9,487	583		3,981	372,138	716	25,554	121	412,581
San Joaquin	247			1,378	833	1,157,843	37,287		1,197,587
Stanislaus	556	621		1,120	25,876	36,474	793,667	500	858,813
Tulare	21,272	7,294	8,705	1,693		264	410	795,079	834,717
Grand Total	1,785,099	1,450,758	238,204	191,498	411,873	1,195,725	859,669	835,413	6,968,238

GRAPH OF TRIP DISTANCE BY MODE

Excel can create pivot tables and pivot charts which appear side-by-side with the same data. As before, confirm there are enough trips in the CHTS to summarize by creating a pivot table with mode as columns, distance as rows, raw person-trips as values (summarized by sum), and any desired filters. In this example, we certainly have enough trips for most modes, but should be cautious about drawing conclusions about

Rail or Other modes. Also, trips of 10 miles or longer are few enough that they should be considered as an aggregate rather than mile-by-mile.

oPlace	Oakland							
dPlace	Oakland							
Sum of rawPersTrips	Column Labels							
Row Labels	Bike	Bus	DriveAlone	DriveShared	Other	Rail	Walk	Grand Total
0	4		58	60			384	506
1	31	16	162	245	10		149	613
2	9	14	110	122	11	1	10	277
3	4	22	85	84		3	4	202
4	1	13	42	38		2	4	100
5		5	53	43		6		107
6		6	36	33	1			76
7		3	20	24	2		1	50
8		7	18	17	1	1		44
9		7	13	12	1	1		34
10		4	6	3		4		17
11		2	5	3				10
12		1	2			1		4
13			1	7		1		9
14				2		1		3
Grand Total	49	100	611	693	26	21	552	2052

To create the graph, change the value field from raw person trips to number of person trips (still summarized by sum). While the default pivot-chart bar chart format conveys some information, it's probably clearer to see if we change the chart type to a line chart:
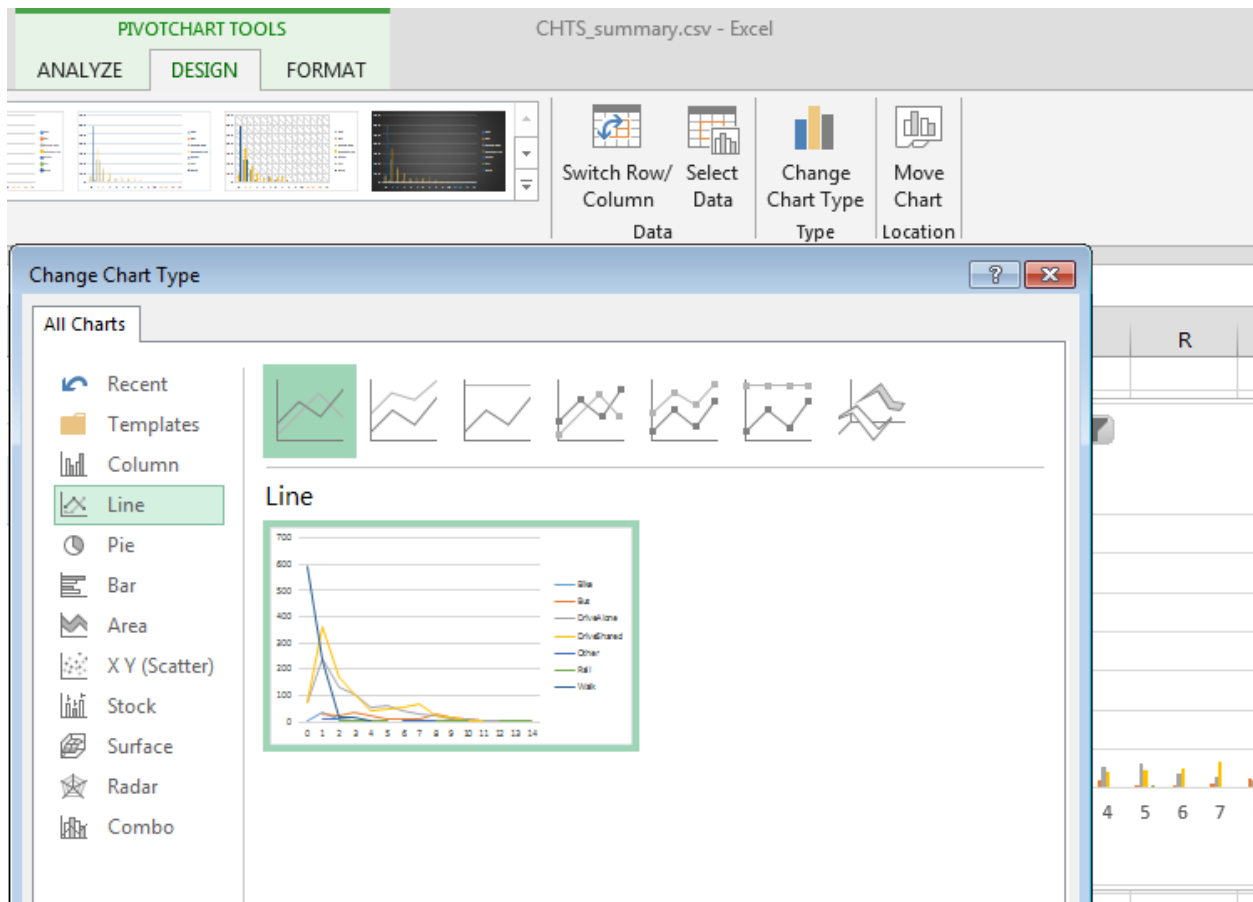
If we'd rather look at mode share for each distance, we can show the values as a percentage of the row total – remembering that trips of 10 miles are longer may show unreasonable variability because there are so few of them in the survey.
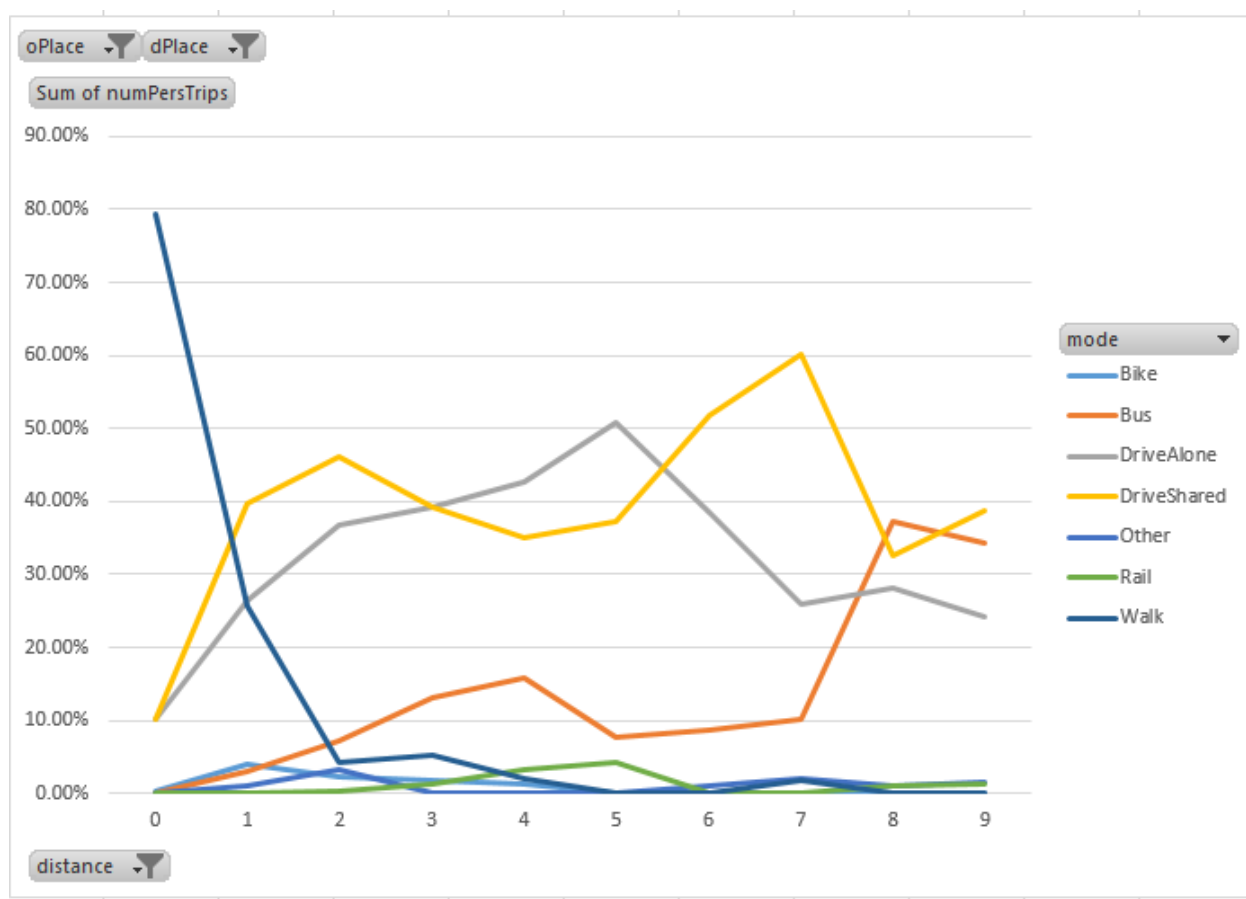
APPENDIX G:

DATA DICTIONARY FOR TAZ DATA INPUTS

The table below is a data dictionary for the elements of the TAZdata.csv model input.

DATA DICTIONARY FOR TAZDATA.CSV

Name	Description
TAZ	Traffic Analysis Zone ID
AIRBASIN	For counties containing multiple air basins,
MID_BNDRY	Middle school boundary
HIGH_BNDRY	High school boundary
GENPARKCOST	Parking cost, general public
EMPCOST	Parking cost, employees
INTDEN	Intersection density (No longer used, replaced by Python script)
WALKPERC	Percentage of TAZ lane miles that are walkable (No longer used, replaced by Python script)
MHHINC	Median household income
AREA_AC	Total area of the TAZ, in acres, including undeveloped land
RESACRE	Total developed area of TAZ devoted to residential uses
EMPACRE	Total developed area of TAZ devoted to non-residential uses
HWYCOM	Percentage of commercial that is highway focused
PTERM	Additional out-of-vehicle time required for drive trip productions to reach vehicle
ATERM	Additional out-of-vehicle time required for drive trip attractions to reach vehicle
PKFREQ	Frequency of peak-period transit service (used for synthetic transit)
OPFREQ	Frequency of off-peak transit service (used for synthetic transit)
EJ	Environmental Justice code

DATA DICTIONARY FOR TAZDATA.CSV

Name	Description
HBWH_ix	Percentage of home-based work (high income) trips produced which leave the model
HBWH_xi	Percentage of home-based work (high income) trips attracted from outside the model
HBWM_ix	Percentage of home-based work (medium income) trips produced which leave the model
HBWM_xi	Percentage of home-based work (medium income) trips attracted from outside the model
HBWL_ix	Percentage of home-based work (low income) trips produced which leave the model
HBWL_xi	Percentage of home-based work (low income) trips attracted from outside the model
HBS_ix	Percentage of home-based shop trips produced which leave the model
HBS_xi	Percentage of home-based shop trips attracted from outside the model
HBK_ix	Percentage of home-based school (K-12) trips produced which leave the model **(NOT used in the model: all HBK trips are assumed to be internal to the model.)**
HBK_xi	Percentage of home-based school (K-12) trips attracted from outside the model **(NOT used in the model: all HBK trips are assumed to be internal to the model.)**
HBC_ix	Percentage of home-based college trips produced which leave the model
HBC_xi	Percentage of home-based college trips attracted from outside the model
HBO_ix	Percentage of home-based other trips produced which leave the model
HBO_xi	Percentage of home-based other trips attracted from outside the model
WBO_ix	Percentage of work-based other trips produced which leave the model
WBO_xi	Percentage of work-based other trips attracted from outside the model

DATA DICTIONARY FOR TAZDATA.CSV

Name	Description
OBO_ix	Percentage of other-based other trips produced which leave the model
OBO_xi	Percentage of other-based other trips attracted from outside the model
EMP_EDUH	Percentage of educational employment that is high-income
EMP_EDUM	Percentage of educational employment that is medium-income
EMP_EDUL	Percentage of educational employment that is low-income
EMP_FOOH	Percentage of food/entertainment employment that is high-income
EMP_FOOM	Percentage of food/entertainment employment that is medium-income
EMP_FOOL	Percentage of food/entertainment employment that is low-income
EMP_GOVH	Percentage of government employment that is high-income
EMP_GOVM	Percentage of government employment that is medium-income
EMP_GOVL	Percentage of government employment that is low-income
EMP_INDH	Percentage of industrial employment that is high-income
EMP_INDM	Percentage of industrial employment that is medium-income
EMP_INDL	Percentage of industrial employment that is low-income
EMP_MEDH	Percentage of medical employment that is high-income
EMP_MEDM	Percentage of medical employment that is medium-income
EMP_MEDL	Percentage of medical employment that is low-income
EMP_OFCH	Percentage of office employment that is high-income

DATA DICTIONARY FOR TAZDATA.CSV

Name	Description
EMP_OFCM	Percentage of office employment that is medium-income
EMP_OFCL	Percentage of office employment that is low-income
EMP_RETH	Percentage of retail employment that is high-income
EMP_RETM	Percentage of retail employment that is medium-income
EMP_RETL	Percentage of retail employment that is low-income
EMP_OTHH	Percentage of mining/manufacturing employment that is high-income
EMP_OTHM	Percentage of mining/manufacturing employment that is medium-income
EMP_OTHL	Percentage of mining/manufacturing employment that is low-income
EMP_AGRH	Percentage of agricultural employment that is high-income
EMP_AGRM	Percentage of agricultural employment that is medium-income
EMP_AGRL	Percentage of agricultural employment that is low-income

APPENDIX H:

ACCESSIBILITY VARIABLES

The table below lists all of the accessibility and D-variables calculated during the Accessibility portions of the model. Note that the accessibility metrics are calculated during the Input Preparation phase of the model, and updated as the model runs through each iteration.

TABLE H-1: DATA DICTIONARY FOR TAZ-LEVEL ACCESSIBILITY VARIABLES

Variable	Description
ATYPE	Place type, calculated from EMP_30AUT + WRK_30AUT
TOTHH_SF	Total households in single-family residential units
HHPOP_SF	Total household population in single-family residential units
TOTHH_MF	Total households in multi-family residential units
HHPOP_MF	Total household population in multi-family residential units.
WRKPOP	Total working-age population.
INTDEN	Intersection density (intersections per square mile, including undeveloped area)
DIRECT	Not currently used; placeholder for measure of directness
WALK_MI	Miles of walkable roadway links
WALKPERC	Percentage of TAZ which is walkable
RESACRE	Developed acres for residential purposes
EMPACRE	Developed acres for non-residential purposes
HH_05TRN	Households within half-mile of transit
WRK_05TRN	Working-age population within half-mile of transit
EMP_05TRN	Jobs within half-mile of transit
EMP_30TRN	Jobs within 30 minutes by transit
WRK_30TRN	Working-age population within 30 minutes by transit
EMP_1WALK	Jobs within 1-mile walk

TABLE H-1: DATA DICTIONARY FOR TAZ-LEVEL ACCESSIBILITY VARIABLES

Variable	Description
WRK_1WALK	Working-age population within 1-mile walk
EMP_3BIKE	Jobs within 3-mile bike ride
WRK_3BIKE	Working-age population within 3-mile bike ride
EMP_30AUT	Jobs within 30 minutes by auto
WRK_30AUT	Working-age population within 30 minutes by auto
ACT_30AUT	Activity (jobs + working-age population) within 30 minutes by auto
ACT_30TRN	Activity (jobs + working-age population) within 30 minutes by transit
COMMUTECOST	Average annual cost of commuting by auto

APPENDIX I:

COMPARISON OF LAND USE CATEGORIES

The table below shows the residential land use data elements and how the VMIP 2 grouping compares to other data sources including the CHTS, ACS, and VMIP 1 categorization.

TABLE 3.2-8:
RESIDENTIAL AGGREGATION STRUCTURE FOR VMIP 2

	VMIP 2 (grouped)	VMIP 2	2012 CHTS	2012 ACS 5 Year	VMIP 1		CTPP 2010
Residence Type			*resty*	*B25024 (BG)*			
		RUG1 (SF)	01 1, detached	RU1	RU1	1, detached	SF detached (RU1)
			02 1, attached	RU2	RU2	1, attached	SF attached (RU2)
		RUG2 (MF)	04 2-4 units	RU3	RU3	2	MF 2-4 (RU3 + RU4)
				RU4	RU4	3 or 4	
			05 5-19 units	RU5	RU5	5 to 9	MF 5-19 (RU5 + RU6)
				RU6	RU6	10 to 19	
			06 20+ units	RU7	RU7	20 to 49	MF 20-49 (RU7)
				RU8	RU8	50 or more	MF 50+ (RU8)
		RUG3 (OTH)	03 Mobile home	RU9	RU9	Mobile home	MH (RU9)
			07 Boat, RV, van, etc.	RU10	RU10	Boat, RV, van, etc.	Other (RU10)
Household Size			*hhsize*	*B25009 (BG)*			
		HH1		1-person	HH1	HOUSEHOLD SIZE 1	1-person
		HH2		2-person	HH2	HOUSEHOLD SIZE 2	2-person
		HH3		3-person	HH3	HOUSEHOLD SIZE 3	3-person
		HH4		4-person	HH4	HOUSEHOLD SIZE 4	4-or-more-person
		HH5	Range is 1-15	5-person	HH5	HOUSEHOLD SIZE 5	
				6-person	HH6	HOUSEHOLD SIZE 6	
				7-or-more-person	HH7	HOUSEHOLD SIZE 7 or more	
				Total Households	TOTHH	TOTAL HOUSEHOLD	
Household Vehicles			*hhveh*	*B25044 (BG)*			
		Veh0		No vehicle available	Veh0	No vehicle available	0 cars
		Veh1	Range is 0-15	1 vehicle available	Veh1	1 vehicle available	1 car
		Veh2		2 vehicles available	Veh2	2 vehicles available	2 cars

TABLE 3.2-8:
RESIDENTIAL AGGREGATION STRUCTURE FOR VMIP 2

	VMIP 2 (grouped)	VMIP 2	2012 CHTS	2012 ACS 5 Year	VMIP 1		CTPP 2010
		Veh3		3 vehicles available	Veh3	3 vehicles available	3 cars
		Veh4		4 vehicles available	Veh4	4 vehicles available	4-or-more-cars
				5 or more vehicles available	Veh5	5 or more vehicles available	
Household Income			*incom*	*S1901 (BG)*			
	INCLOW	INC1	1 Less than $10,000	Less than $10,000	INC1	Less than $10,000	Less than $15,000
			2 $10,000 - $24,999	$10,000 to $14,999	INC2	$10,000 to $14,999	
				$15,000 to $19,999	INC3	$15,000 to $24,999	$15,000-$24,999
				$20,000 to $24,999			
		INC2	3 $25,000 - $34,999	$25,000 to $29,999	INC4	$25,000 to $34,999	$25,000-$34,999
				$30,000 to $34,999			
			4 $35,000 - $49,999	$35,000 to $39,999	INC5	$35,000 to $49,999	$35,000-$49,999
				$40,000 to $44,999			
				$45,000 to $49,999			
	INCMED	INC3	5 $50,000 - $74,999	$50,000 to $59,999	INC6	$50,000 to $74,999	$50,000-$74,999
				$60,000 to $74,999			
		INC4	6 $75,000 - $99,999	$75,000 to $99,999	INC7	$75,000 to $99,999	$75,000-$99,999
	INCHIGH	INC5	7 $100,000 - $149,999	$100,000 to $124,999	INC8	$100,000 to $149,999	$100,000-$149,999
				$125,000 to $149,999			
			8 $150,000 - $199,999	$150,000 to $199,999	INC9	$150,000 to $199,999	$150,000 or more
			9 $200,000 - $249,999	$200,000 or more	INC10	$200,000 or more	
			10 $250,000 or more				
				Total, household income	TOTINC	TOTAL HH INCOME	Total, household income
Population by Age			*age*	*B01001 (BG)*	AGE		
		POP0005		Under 5 years		People 0 to 5 years	
		POP0514		5 to 9 years		People 5 to 14 years	
				10 to 14 years			
		POP1517	*Range is 0-98, 99 for 99+*	15 to 17 years		People 15 to 17 years	
		POP1824		18 and 19 years		People 18 to 24 years	
				20 years			
				21 years			
				22 to 24 years			

TABLE 3.2-8:
RESIDENTIAL AGGREGATION STRUCTURE FOR VMIP 2

	VMIP 2 (grouped)	VMIP 2	2012 CHTS	2012 ACS 5 Year	VMIP 1		CTPP 2010
		POP2554		25 to 29 years 30 to 34 years 35 to 39 years 40 to 44 years 45 to 49 years 50 to 54 years	People 25 to 54 years		
		POP5564		55 to 59 years 60 and 61 years 62 to 64 years	People 55 to 64 years		
		POP6574		65 and 66 years 67 to 69 years 70 to 74 years	People 65 to 74 years		
		POP75		75 to 79 years 80 to 84 years 85 years and over	People 75 years and over		
Age of head of household			age	SF1-2010 H17 (ACS B19037 has fewer categories)			
		AGE1524		Householder 15 to 24 years	Hage1	Householder 15 to 24 years	Householder 15 to 17 years
		AGE2564	Not a separate variable but does have ages of all household members to use for calculation of this variable	Householder 25 to 34 years	Hage2	Householder 25 to 34 years	Householder 18 to 24 years
				Householder 35 to 44 years	Hage3	Householder 35 to 44 years	Householder 25 to 44 years
				Householder 45 to 54 years	Hage4	Householder 45 to 54 years	Householder 45 to 59 years
				Householder 55 to 59 years	Hage5	Householder 55 to 59 years	Householder 60 to 64 years
				Householder 60 to 64 years	Hage6	Householder 60 to 64 years	Householder 65 to 74 years
		AGE6574		Householder 65 to 74 years	Hage7	Householder 65 to 74 years	Householder 75 years and over
		AGE75		Householder 75 to 84 years	Hage8	Householder 75 to 84 years	
				Householder 85 years and over	Hage9	Householder 85 years and over	
work trip Travel time			totalTime (F&P created)				
			All travel times are measured in minutes; for transit trips totalTime is a sum of IVT, waitTime, accessTime, xferTime,egressTime		TT1	Less than 10 minutes	Less than 5
					TT2	10 to 14 minutes	5 to 9 minutes
					TT3	15 to 19 minutes	15 to 19 minutes
					TT4	20 to 24 minutes	20 to 20 minutes
					TT5	25 to 29 minutes	30 to 44 minutes
					TT6	30 to 34 minutes	45 to 59 minutes
					TT7	35 to 44 minutes	60 to 74 minutes
					TT8	45 to 59 minutes	75 to 89 minutes

TABLE 3.2-8:
RESIDENTIAL AGGREGATION STRUCTURE FOR VMIP 2

	VMIP 2 (grouped)	VMIP 2	2012 CHTS	2012 ACS 5 Year	VMIP 1		CTPP 2010
				*same as VMIP 1	TT9	60 or more minutes *is available as same in CTPP *still looking for place of work by census tract	90 minutes or more **aggregate option

TABLE 3-2.9: NON-RESIDENTIAL LAND USE CATEGORY AGGREGATION STRUCTURE

VMIP 2	VMIP 1	Description	NAICS	CTPP	CSTDM
EMPEDU	EDUCATION	Educational Services (Schools, Junior Colleges, Colleges, Universities, Professional Schools	61	Edu / Health	Education and health
EMPFOO	ACCOMODTNS	Accommodation	721	Arts/Rec/Accom/Food	Leisure and hospitality
	FOOD	Food Services	722	Arts/Rec/Accom/Food	Leisure and hospitality
	ENT_REC	Arts, Entertainment, and Recreation	71	Arts/Rec/Accom/Food	Leisure and hospitality
EMPGOV	PUBLIC	Public Administration	92	Government	Office
EMPIND	CONSTRUCTN	Construction	23	Construction	Primary and Secondary
	UTILITIES	Utilities	22	Trans / Util.	Trans / Util.
	SVC_OTHER	Other Services (except Public Administration)	81	Other Service	Other Service
	WHOLESALE	Wholesale Trade	42	Wholesale	Wholesale
	WAREHOUSE	Transportation and Warehousing	48-49	Trans / Util.	Trans / Util.
EMPMED	HEALTH	Health Care and Social Assistance	62	Edu / Health	Education and health

	TABLE 3-2.9: NON-RESIDENTIAL LAND USE CATEGORY AGGREGATION STRUCTURE					
VMIP 2	VMIP 1	Description	NAICS	CTPP		CSTDM
EMPOFC	INFORMATN	Information	51	Information		Office
	FINAN_INSR	Finance and Insurance	52	FIRE		Office
	REALESTATE	Real Estate and Rental and Leasing	53	FIRE		Office
	SVC_PROF	Professional, Scientific, and Technical Services	54	Prof Sci, Admin		Office
	SVC_MNGMNT	Management of Companies and Enterprises	55	Prof Sci, Admin		Office
	SVC_ADMIN	Administrative and Support and Waste Management and Remediation Services	56	Prof Sci, Admin		Office
EMPRET	RETAIL	Retail Trade	44-45	Retail		Retail
EMPOTH	MANUFACTUR	Manufacturing	31-33	Manufacturing		Primary and Secondary
	MINING	Mining, Quarrying, and Oil and Gas Extraction	21	Ag_Mining		Primary and Secondary
EMPAGR	AGRICULTUR	Agriculture, Forestry, Fishing and Hunting	11	Ag_Mining		Primary and Secondary

APPENDIX J:

GUIDANCE ON STATIC VALIDATION

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
Static Validation					
Transit Assignment	1. Difference between actual ridership to model results for entire system	+/- 20%	2010 RTP Guidelines Daily	Source of actual daily ridership: http://www.ntdprogram.gov/ntdprogram/archives.htm (National transit database for base year, typically 2008) 2010 RTP Guidelines specify difference between actual ridership to model results for a given year by route group (i.e., Local Bus, Express Bus, etc.). However, National transit database only specifies transit ridership for entire system. Valley Transit operators do not use consistent route groups.	Table
Traffic Assignment	2. % of Links within Caltrans Deviation Allowance	At Least 75%	2010 RTP Guidelines *Travel Forecasting Guidelines*, Caltrans, 1992	Source of traffic data: Vehicle count database for each County for comparison Daily, non-directional	Table, Figure of location and deviation color (valid, +1, +2, -1, -2). Graph (model validation scatter plot).
	3. % of Screenlines within Caltrans Deviation Allowance	100%	2010 RTP Guidelines *Travel Forecasting Guidelines*, Caltrans, 1992	Daily, non-directional	Table

[1] Potential solutions to unexpected results may vary-: TMIP Guidelines are the standard reference for troubleshooting and solutions: http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/FHWA-HEP-10-042/FHWA-HEP-10-042.pdf

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
	4. Correlation Coefficient	At Least 0.88	3.2010 RTP Guidelines *Travel Forecasting Guidelines*, Caltrans, 1992	Daily, non-directional	Table
	5. Percent Root Mean Squared Error (RMSE) (model-wide)	Below 40%	2010 RTP Guidelines	Daily, non-directional	Table
	6. Percent Root Mean Squared Error (RMSE) (functional classification)	Below 40%		No specific criteria available Daily, non-directional Functional Class: Freeway Highway Expressway Arterial Collector	Table
	7. Percent Root Mean Squared Error (RMSE) (volume range)	0-4,999 – <116% 5,000 to 9,999 – <43% 10,000 to 19,999 – <28% 20,000 to 39,999 – < 25% 40,000 to 59,000 – < 30% 60,000 to 89,999 – <-19%	Harvey, G., et al. A Manual of Regional Transportation Modeling Practice for Air Quality Analysis for the Natural Association of Regional Councils, Washington, D.C. July 1993	Is there a minimum number of counts in a volume range or functional class range that we want to consider?	Table

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
	8. Model Volume to Count Ratio (model-wide)	General relationship (i.e., high or low) between model volumes and counts	2010 RTP Guidelines	Daily, non-directional *Minimum Travel Demand Model Calibration and Validation Guidelines for State of Tennessee*. FHWA - identifies that model volumes should be within 5-10% of observed traffic volumes on the highway network. This is the range reference in TMIP, *Model Validation and Reasonableness Checking Manual*, 1997 for screenlines	Table
	9. Model Volume to Count Ratio (roadway functional classification)	Freeway – +/- 7% Major Arterial – 10% Minor Arterial – 15% Collector – 25%	TMIP, *Model Validation and Reasonableness Checking Manual*, 1997	Daily, non-directional Percent difference targets for daily traffic volumes by facility type.	Table
	XX. Distribution of Class by Time of Day	Comparison to collected count data		Total vehicles trips stratified by class and time of day.	Table
	XX. Distribution of Time of Day by Class	Comparison to collected count data		Total vehicles trips stratified by time of day and class.	Table

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
	10. Model Volume to Count Ratio (volume range)	<1,000 < 60% 1,000-2,500 < 47% 2,500-5,000 – <36% 5,000-10,000 – <29% 10,000-25,000 – <25% 25,000-50,000 – <22% >50,000 – <21%	TMIP, *Model Validation and Reasonableness Checking Manual,* 1997	Percent difference targets for daily traffic volumes for individual links.	Table
Reasonableness Checks					
Highway and Transit Networks	11. General roadway network and transit line coding	Reasonableness Check	TDF Model	Centerline	
Trip Generation	12. PA Balance	+/- 10% by purpose and overall	TDF Model	after including IX/XI trips	Table or bar chart comparing balance before and after adjustment
Trip Distribution	13. Zonal Trip Distribution		TDF Model	Select link assignment for gateways, TAZ near gateway, and TAZ central to model network.	Network bandwidth plots.

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
Vehicle Availability	14.		2010 ACS (Surveys from 2006-2010) and CHTS http://www.dot.ca.gov/hq/tsip/tab/documents/travelsurveys/Final2001_StwTravelSurveyWkdayRpt.pdf	County level comparison Compare percent of households (single and multiple) with 0, 1, 2, 3+ autos CHTS includes survey data for Fresno, Kern, Merced, San Joaquin, Stanislaus, and Tulare counties. (Table 4, Pages 26 – 30)	
Feedback Loop	15.			Convergence	
Comparisons					
Land Use	16. Total Population	Within 3% (based on RHNA criteria)	Census	by income group	Bar chart comparing model to census data.
	17. Total Households	Ideally within 3% (RHNA criteria)	Census or Department of Finance	RHNA allocations are not anticipated until mid-2013	Bar chart comparing model to census data.
	18. Total Employment	Note	Department of Finance	Check reasonableness of retail jobs per household and non-retail jobs per household. Job mix?	Bar chart comparing model to census data.
Trip Generation	19. Person trip rates		CHTS, ITE	Convert person trip rates to ITE rates using Ave Veh Occ by purpose	Table
Trip Distribution	20. Average Trip Length by Purpose		CHTS	3-County model also has OD survey	Table
	21. Trip Length Frequency Distribution by Purpose		CHTS	3-County model also has OD survey	Graph for each purpose

TABLE A-1:
DRAFT SUMMARY OF MODEL PERFORMANCE – STATIC VALIDATION

Model Component	Validation Statistic	Evaluation Criterion	Source	Notes, further guidance[1]	Documentation
	XX. Percentage of IX/XI/XX trips for long-distance trips		Cellphone Inter-regional Data	Compare percentage of II/IX/XI trips from model trip tables with percentage of II/IX/XI trips from cellphone inter-regional travel data.	Table and/or Map
Trip Assignment	22. Vehicle class		Count data	Percent by class for each period Percent by time period for each class	Table
	23. VMT	+/- 5%	HPMS http://www.dot.ca.gov/hq/tsip/hpms/hpmslibrary	Compare countywide daily VMT estimate from HPMS (Table 10, Page 80) Reasonableness of comparison should be based on how the model compares to HMPS estimates. In general, The model should be VMT forecasts should be lower than the HPMS estimate, since HPMS VMT is estimated for local streets that are not in the model networks.	Table
	24. Travel Speed by Functional Classification		Existing Data	Compare by functional classification based on observed data. For all classifications, summarize average speed, minimum, and maximum. If observed data is not available, compare relative congested speed by functional class.	Table
	25. Average Travel Time by Trip Purpose		CHTS	Daily CHTS provide travel time for HBW trips and total trips. http://www.dot.ca.gov/hq/tsip/tab/documents/travelsurveys/Final2001_StwTravelSurveyWkdayRpt.pdf	Table
Mode Split	26. Mode split by purpose		CHTS	Daily	Pie chart

Source: Fehr & Peers, 2016

APPENDIX K:

MEMO ON AUTO OPERATING COST

MEMORANDUM

To: Ken Kirkey, MTC; Huasha Liu, SCAG; Gordan Garry, SACOG; Muggs Stoll, SANDAG

From: David Ory, MTC; Guoxiong Huang, SCAG; Bruce Griesenbeck, SACOG; Clint Daniels, SANDAG

Re: Automobile Operating Cost for the Second Round of Sustainable Communities Strategies

Date: October 13, 2014

This memorandum summarizes our collective thinking regarding fuel price assumptions for the second round of sustainable communities strategies (SCSs)[2].

Background

The Regional Targets Advisory Committee (or RTAC) formed by the California Air Resources Board (ARB) recommended that MPOs use "consistent long-range planning assumptions statewide, to the degree practicable, including … existing and forecasted fuel prices and automobile operating costs."[3] For the first round of sustainable communities strategies, we agreed to use the following sets of assumptions:

- Base Year Fuel Price: Region-specific, set during model calibration
- Year 2020 Fuel Price: $4.74 (Year 2009 dollars, $2009);
- Year 2035 Fuel Price: $5.24 ($2009);
- Effective Fleet-wide Fuel Efficiency: Region-specific, derived from ARB's Emission Factor (EMFAC) software;
- Year 2020 Non-fuel-related Operating Cost (if included in region-specific automobile operating cost calculations): $0.09 ($2009);

---

[2] The first round beginning with SANDAG's 2011 RTP/SCS; the second round beginning with SANDAG's 2015 RTP/SCS.
[3] See page 10 of *Recommendations of the Regional Targets Advisory Committee Pursuant to Senate Bill 375: A Report to the California Air Resources Board*.

- Year 2035 Non-fuel-related Operating Cost (if included in region-specific automobile operating cost calculation): $0.11 ($2009).

This set of assumptions were used to compute the assumed perceived automobile operating cost for each MPO. The resulting values are shown in **Table 54**.

Table 54: Assumed Perceived Automobile Operating Costs ($2009) for First Round of SCSs

MPO	Base Year Cost (year)	Year 2020 Cost	Year 2035 Cost	Avg Annual Growth (Base to 2035)
SCAG	$0.23 (2005)	$0.32	$0.32	1.1%
MTC	$0.18 (2010)	$0.28	$0.28	1.8%
SACOG	$0.21 (2008)	$0.27	$0.29	1.2%
SANDAG	$0.19 (2008)	$0.22	$0.21	0.4%

Using the above assumptions, we achieved consistency in forecast year fuel price as well as the approach to computing perceived automobile operating cost. Unfortunately, we were not able to achieve consistency in base year assumptions. Achieving consistency across MPOs for base year input is more difficult than achieving consistency across forecast year input because base year input is part of the expensive and time consuming model development process.

The result of using consistent forecast year assumptions and inconsistent base year assumptions were uneven changes in the assumed increase in perceived automobile operating cost across MPOs. For example, between 2010 and 2035, MTC assumes a 1.8 percent average annual increase in perceived automobile operating cost; between 2008 and 2035, SANDAG assumes a 0.4 percent average annual increase. It is worth noting that the base year differences may reflect actual base year differences (i.e., fuel prices changing from 2005 to 2010) and do reflect regional differences in the assumed average fleet-wide fuel efficiency. In any case, the differences in growth rates make it difficult to claim that the perceived automobile operating costs were handled in a consistent manner.

Proposed Approach

Our proposed remedy for the above-described problem is *not* to try and achieve consistent base year assumptions. The model calibration process is difficult enough without adding the constraint of a single perceived automobile operating cost introduced at an unknown time in the model development cycle. Rather, we propose using a consistent growth in fuel price between the SB 375 base year of 2005 and the forecast years used in the SCS, specifically the target years 2020, and 2035. In addition, we propose using a consistent non-fuel-related operating cost as well as consistent data sources for effective fleet-wide fuel efficiency and base year gas price.

The following subsections outline the approach. Note that the below assumptions do not account for potential increases in fuel costs from California's Cap-and-Trade program.

Fuel Price Assumptions

The Department of Energy issues an annual forecast of motor vehicle gasoline prices. The 2013 forecast[4] is paired with historical information from 2005 to compute a consistent fuel price ratio that will be used by each MPO. The target value for the calculation is not the midpoint between the low and high forecast, but rather three-quarters of the way between the low and high forecasts, plus 32 cents ($2010) – the 32 cents accounts for gasoline generally being more expensive in California than the rest of the nation. These calculations are shown in **Table 55**.

Table 55: Department of Energy Forecasts and Resulting Growth Ratio (Prices in Year 2010 Dollars)

Year	Low	High	Low plus 75% Diff + 32 cents	Ratio to 2005
2005	---	---	$2.82*	---
2015	$2.70	$3.77	$3.82	1.35
2020	$2.54	$4.17	$4.08	1.45
2025	$2.53	$4.39	$4.25	1.51
2030	$2.52	$4.77	$4.53	1.61
2035	$2.53	$5.18	$4.84	1.72
2040	$2.57	$5.70	$5.24	1.86

* – Historical price taken from http://www.eia.gov/dnav/pet/pet_pri_gnd_a_epm0_pte_dpgal_a.htm, and converted to year 2010 dollars.

To compute an MPO-specific forecast year fuel price, the growth ratios in **Table 55** are paired with base year prices. We propose using base year prices from a consistent source, specifically the retail gasoline price data from the Oil Price Information Service (OPIS); these prices will be introduced during our next round of model development activities. The assumed base year prices are shown in **Table 56** for each of the MPO areas for years 2005 through 2012. These prices will be used in subsequent model development activities[5].

Table 56: Historical Gas Prices per OPIS (All prices in Year 2010 dollars)

Year*	MTC	SCAG	SACOG	SANDAG
2005	$2.83	$2.85	$2.74	$2.84

---

[4] The data is here: http://www.eia.gov/forecasts/archive/aeo13/source_oil.cfm.

[5] Some MPOs will be recalibrating their models and generating a "new" "forecasts" (or "backcasts") of year 2005. Others will not. Those generating new forecasts will use the fuel prices listed in **Table 56**; those not generating new forecasts will leave their prices as they were set in their model development processes.

2008	$3.68	$3.53	$3.53	$3.35
2010	$3.17	n/a	$3.09	$2.92
2012	$3.87	$3.90	$3.85	$3.64

* - The base year prices are only shown (and, in some cases, only purchased) for 2005 and potential model calibration years. For example, SCAG intends to use a 2012 calibration year, and, as such, did not purchase the year 2010 prices from OPIS.

Non-Fuel-Related Operating Costs

As noted above, the calculation of perceived automobile operating cost is assumed to have two components: fuel costs and non-fuel-related costs. Similar to the base year fuel price, we propose using base year non-fuel-related operating costs from a consistent source, specifically the American Automobile Association (AAA). The assumed non-fuel-related base year prices are shown in **Table 57**; these are national estimates that we'll assume apply to each of the MPO areas. These prices will be used in subsequent model development activities.

Table 57: Non-Fuel-Related Operating Costs (Prices in Year 2010 dollars per mile)

Year	Maintenance	Tires	Maint. + Tires
2005	$0.0437	$0.0062	$0.05
2006	$0.0453	$0.0065	$0.05
2007	$0.0437	$0.0069	$0.05
2008	$0.0452	$0.0076	$0.05
2009	$0.0447	$0.0082	$0.05
2010	$0.0444	$0.0096	$0.05
2011	$0.0461	$0.0103	$0.06
2012	$0.0524	$0.0105	$0.06

The above data can be used to estimate forecast-year non-fuel-related costs. Using a simple linear regression and extrapolation, the forecast year values shown in **Table 58** can be computed. Similar to the gasoline price, the MPOs will use the computed ratio to calculate the forecast year values from whatever values were or are assumed for year 2005.

Table 58: Forecast Year Non-Fuel-Related Operating Costs Ratios (Prices in Year 2010 dollars)

Year	Estimate	Ratio to 2005
2005	$0.050	---
2012	$0.063	1.26
2015	$0.062	1.25
2020	$0.069	1.38
2025	$0.075	1.50
2030	$0.081	1.62
2035	$0.087	1.75
2040	$0.093	1.87

Effective Fleet-wide Fuel Efficiency

The computation of perceived automobile operating cost requires an assumption be made about the effective passenger-vehicle[6] fuel efficiency. ARB's EMFAC software provides two estimates of carbon dioxide (CO_2) emissions. The first estimate is for a hypothetical future in which fuel and vehicle regulations are not enacted; this hypothetical future is used only for computing emissions for SB 375 purposes (method A). The second estimate is for the expected future in which fuel and vehicle regulations are enacted (method B). This future is assumed for all non-SB 375 purposes, including federally-mandated conformity analyses. Unfortunately, the EMFAC software only provides a fuel consumption result for the first set (method A) of CO_2 emissions. The effective fleet-wide fuel efficiency needs to be calculated from the second estimate. Each MPO will use the following equation to compute the effective fleet-wide fuel efficiency:

$$FE = \frac{VMT}{\frac{(CO_2)_B \cdot FLCFS}{(CO_2)_A} \cdot FC_A}$$

where VMT is passenger-vehicle miles traveled, $(CO_2)_A$ is the passenger-vehicle CO_2 estimate from method A, $(CO_2)_B$ is the passenger-vehicle CO_2 estimate from method B, and FC_A is the passenger-vehicle fuel consumption from method A. FLCFS is an adjustment factor to account for Low Carbon Fuel Standards (LCFS) CO_2 reduction factors assumed in EMFAC 2011. LCFS is a fuel standard that requires a reduction of at least 10 percent in the carbon intensity of California's transportation fuels by 2020 (see Table 5-2, http://www.arb.ca.gov/msei/emfac2011-technical-documentation-final-updated-0712-v03.pdf). FLCFS is set at 1.11 to offset this reduction factor in the fuel efficiency calculations as the reduction

---

[6] Defined as EMFAC vehicle types LDA, LDT1, LDT2, and MDV.

from LCFS is related to carbon content rather than fuel consumption. The calculation assumes a linear relationship between CO_2 emissions and fuel consumption.

Using the effective fuel efficiency derived from EMFAC presents a "chicken or egg" problem, as one cannot generate the fuel-efficiency estimate unless an input assumption about operating cost is made, but the operating cost assumption requires a fuel-efficiency estimate. In practice, each MPO will select a representative fuel efficiency estimate during the SCS development process that will be carried through SCS adoption.

Region-Specific Calculations

Detailed calculations are provided below for each of the MPO regions. The regions differ as to whether they will update the year 2005 simulation results using the prices presented in **Table 56** and **Table 57**; either way, consistent ratios for fuel prices (presented in **Table 55**) and non-fuel-related prices (**Table 58**) are applied to either the updated or non-updated 2005 assumptions.

MTC: Assuming updated Year 2005 Simulation Results

Using the above information, MTC will compute the year 2005, 2020, and 2035 perceived automobile operating cost estimates using the approach detailed in **Table 59**.

Table 59: MTC Region Example Calculations Assuming Updated 2005 Results (Prices in Year 2010 dollars)

Year	Quantity	Value
2005	Region-specific fuel price (Table 56, dollars per mile)	$2.83
	Non-fuel-related price (Table 57, dollars per mile)	$0.05
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	20.09
	Perceived automobile operating cost (cents per mile)	19.1¢
2020	Consistent fuel price ratio (Table 55)	1.45
	Region-specific fuel price (Ratio x 2005 price)	$4.09
	Consistent non-fuel-related price ratio (Table 58)	1.38
	Region-specific non-fuel-related price	$0.07
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	25.15[†]
	Perceived automobile operating cost (cents per mile)	23.1¢
2035	Consistent fuel price ratio (Table 55)	1.72
	Region-specific fuel price (Ratio x 2005 price)	$4.85
	Consistent non-fuel-related price ratio (Table 58)	1.75

Region-specific non-fuel-related price	$0.09
Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	28.85[†]
Perceived automobile operating cost (cents per mile)	25.6¢

[†] - Value may change during the planning process.

SCAG: Assuming Updated Year 2005 Simulation Results

Using the information contained in this memorandum, SCAG will compute the year 2020 and 2035 perceived automobile operating cost estimates using the approach detailed in **Table 61**.

Table 60: SCAG Region Example Calculations (Prices in Year 2010 dollars)

Year	Quantity	Value
2005	Region-specific fuel price (Table 56, dollars per gallon)	$2.85
	Non-fuel-related price (Table 57, dollars per mile)	$0.05
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	18.63
	Perceived automobile operating cost (cents per mile)	20.3¢
2020	Consistent fuel price ratio (Table 55)	1.45
	Region-specific fuel price (Ratio x 2005 price)	$4.12
	Consistent non-fuel-related price ratio (Table 58)	1.38
	Region-specific non-fuel-related price	$0.07
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	23.63[†]
	Perceived automobile operating cost (cents per mile)	24.3¢
2035	Consistent fuel price ratio (Table 55)	1.72
	Region-specific fuel price (Ratio x 2005 price)	$4.89
	Consistent non-fuel-related price ratio (Table 58)	1.75
	Region-specific non-fuel-related price	$0.09
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	26.40[†]
	Perceived automobile operating cost (cents per mile)	27.3¢

[†] - Value may change during the planning process.

SACOG: Assuming Static Year 2005 Simulation Results

Using the information contained in this memorandum, SACOG will compute the year 2020 and 2035 perceived automobile operating cost estimates using the approach detailed in **Table 61**.

Table 61: SACOG Region Example Calculations (Prices in Year 2010 dollars)

Year	Quantity	Value
2005	Region-specific fuel price (Table 3, dollars per gallon)	$2.74
	Non-fuel-related price (Table 4, dollars per mile)	$0.05
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	19.50
	Perceived automobile operating cost (cents per mile)	19.1¢
2020	Consistent fuel price ratio (Table 55)	1.45
	Region-specific fuel price (Ratio x 2005 price)	$3.96
	Consistent non-fuel-related price ratio (Table 58)	1.38
	Region-specific non-fuel-related price	$0.07
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	24.92[†]
	Perceived automobile operating cost (cents per mile)	22.8¢
2035	Consistent fuel price ratio (Table 55)	1.72
	Region-specific fuel price (Ratio x 2005 price)	$4.70
	Consistent non-fuel-related price ratio (Table 58)	1.75
	Region-specific non-fuel-related price	$0.09
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	28.30[†]
	Perceived automobile operating cost (cents per mile)	25.4¢

[†] - Value may change during the planning process.

SANDAG: Assuming Static Year 2005 Simulation Results

Using the information contained in this memorandum, SANDAG will compute the year 2020 and 2035 perceived automobile operating cost estimates using the approach detailed in **Table 62**.

Table 62: SANDAG Region Example Calculations (Prices in Year 2010 dollars)

Year	Quantity	Value
2005	Region-specific fuel price (Table 56, dollars per gallon)	$2.84
	Non-fuel-related price (Table 57, dollars per mile)	$0.05
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	18.89
	Perceived automobile operating cost (cents per mile)	20.0¢
2020	Consistent fuel price ratio (Table 55)	1.45
	Region-specific fuel price (Ratio x 2005 price)	$4.11
	Consistent non-fuel-related price ratio (Table 58)	1.38
	Region-specific non-fuel-related price	$0.07
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	23.98[†]
	Perceived automobile operating cost (cents per mile)	24.0¢
2035	Consistent fuel price ratio (Table 55)	1.72
	Region-specific fuel price (Ratio x 2005 price)	$4.87
	Consistent non-fuel-related price ratio (Table 58)	1.75
	Region-specific non-fuel-related price	$0.09
	Effective passenger vehicle fuel efficiency (EMFAC, miles per gallon)	27.20[†]
	Perceived automobile operating cost (cents per mile)	26.7¢

[†] - Value may change during the planning process.

Comparisons across SCS Rounds

Table 63 compares the fuel price and resulting automobile operating cost results across SCS rounds for each MPO *assuming* the effective fleet-wide fuel efficiency number remains unchanged from the first to second round – this number will change during the planning process.

Table 63: Fuel Price and Automobile Operating Cost Comparison across SCS Rounds (Prices in Year 2010 Dollars)

Year	Quantity	MTC		SCAG		SANDAG		SACOG	
		Rnd 1	Rnd 2	Rnd 1	Rnd 2	Rnd 1	Rnd 2	Rnd 1	Rnd 2
2005	Fuel price	$2.79	$2.83	$2.83	$2.85	$2.68	$2.84	$2.70	$2.74
	Auto. Oper. Cost	21.2¢	19.1¢	23.8¢	20.3¢	19.2¢	18.9¢	19.7¢	19.1¢
2020	Fuel price	$4.74	$4.09	$4.74	$4.12	$4.74	$4.11	$4.74	$3.96
	Auto. Oper. cost	28.7¢	23.1¢	31.9¢	24.3¢	22.6¢	24.0¢	27.0¢	22.8¢
2035	Fuel price	$5.24	$4.85	$5.24	$4.89	$5.24	$4.87	$5.24	$4.70
	Auto. Oper. cost	28.6¢	25.6¢	32.3¢	27.3¢	21.7¢	26.7¢	28.9¢	25.4¢
Ratios	2020 to 2005	1.34	1.21	1.34	1.20	1.18	1.20	1.37	1.20
	2035 to 2005	1.33	1.34	1.36	1.34	1.13	1.33	1.47	1.33

Next Steps

This memorandum proposes a consistent approach for computing fuel price for each of our MPOs for the second round of sustainable community strategies. After collecting your feedback and modifying our approach accordingly, we will share this approach with ARB and the other MPOs across the state.

APPENDIX L:

CALIBRATED PARAMETERS