California Department of Transportation

2000-2001 California Statewide Household Travel Survey

Final Report

(June 2002)

NuStats

California Department of Transportation

Caltrans

2000-2001
California Statewide
Household Travel Survey
Final Report

Table of Contents

List of Tables and Figures

List of Tables and Figures (Cont.)

Executive Summary

The California Department of Transportation (Caltrans) maintains a statewide database of household socioeconomic and travel information, which is used in regional and statewide travel demand forecasting. The most recent database, prior to this survey, contained data from the last statewide survey that was conducted in 1991. The 2000-2001 California Statewide Household Travel Survey was conducted to update the database and will be used to help refine travel estimates, models, and forecasts throughout the State. The resultant data set will be used to estimate and forecast trip generation and distribution, mode choice, and assignments, as well as for vehicle emissions analyses and estimates.

The 2000-2001 survey was conducted between October 2000 and December 2001 among households located in each of the 58 counties throughout the State. A total of 17,040 households participated in the survey. Household socioeconomic data gathered in this survey includes information on household size, income, vehicle ownership, employment status of each household member, and housing unit type among other data. Travel information was also collected including trip times, mode, activity at location, origin and destination, and vehicle occupancy among other travel-related data.

The survey was conducted among randomly selected households using the telephone recruitment/diary mail-out/telephone trip retrieval method. The statistics presented in this report are adjusted for trip under reporting as determined by the Global Positioning System (GPS) analysis as described in Chapter 7. The overall trip under reporting correction factor includes both the Caltrans and Southern California Association of Governments (SCAG) Household Travel Survey GPS efforts. Caution must be taken when comparing results in this report to the 1991 California Statewide Travel Survey results for two reasons:

1) The 1991 Survey results were *not* adjusted for trip under reporting and

2) The 1991 Survey utilized one weight factor – own/rent status by vehicles available while the 2000-2001 Survey results were weighted by four weight factors – own/rent status by vehicle availability, households by county distribution, household size and income.

All trip-level results presented in this section and throughout the main report are based on unlinked trips.

Key Statewide Statistics

This report presents socioeconomic and travel characteristics from the 2000-2001 California Statewide Household Travel Survey. The tables presented in this section include statewide statistics – weighted and expanded. A summary of the key statistics is shown in the following tables. Table A summarizes the key household statistics. Table B summarizes the key travel statistics.

As shown in Table A, the mean household size is 2.8 and the mean number of vehicles available to each household is 1.9. The statewide median household income is $54,946. In 2000, there were 11,502,870 occupied housing units in the State of which just over two-thirds (68%) were single housing units.

Table A: 2000-2001 California Statewide Household Travel Survey Key Household Statistics

Household Vehicles Available	21,448,770
Vehicles in Use on Average Weekday (71%)	15,252,463
Full-time Employees	10,130,359
Licensed Drivers	19,696,497
Occupied Housing Units	11,502,870
Single Housing Units	68%
Multiple and Other Housing Units	31%
Median Household Income	$54,946
Persons Per Household	2.8

Vehicles Per Household	1.9
No Vehicles	9.3%
One Vehicle	29.7%
Two Vehicles	37.7%
Three or More Vehicles	23.4%
Licensed Drivers Per Household	1.7

Table B summarizes the survey trip characteristics for the State of California[1]. On an average weekday (Monday through Friday travel data), there are just over 136 million trips generated throughout the State. The mean number of trips per household is 11.9.

On a per person basis, 5.3 person trips are generated on an average weekday. Among all trips, the average vehicle occupancy is 1.5 persons. During the peak morning commute time (7-9 a.m.), the vehicle occupancy is 1.5 persons. Home-to-work trips have on average 1.1 persons per vehicle during all times of the day and during peak morning commute times.

**Table B: 2000-2001 California Statewide Household Travel Survey Key Trip Statistics
(Unlinked Trips)**

Travel Data (Unlinked Trips)[1]	Average Weekday (Monday through Friday travel)
Total Household Trips[2]	136,078,168
Household Person Trips[3]	126,812,085
Household Driver Trips[4]	97,108,923
Total Trips Per Household	11.9
Person Trips Per Household	11.0
Person Trips Per Person (All ages)	5.3
Person Trips Per Person Five+ Years of Age[5]	5.6
Driver Trips Per Household	8.5
Driver Trips Per Vehicle Available	4.6
Driver Trips Per Vehicle In Use on Travel Day	6.5
Vehicle Occupancy	
All Trips (24 hours)	1.5
All Trips (7-9 a.m.)	1.5
Home to Work Trips (24 hours)	1.1
Home to Work Trips (7-9 a.m.)	1.1
Mean Travel Time (Respondent Reported)	
Trip Length (All Trips)	20.0 minutes
Weekday Trip Length (Home to Work Trips)	27.6 minutes

---

[1] Adjustment of +64.7% was applied to relevant weekday driver trip rates based on the Global Positioning System (GPS) trip under reporting correction factor. The adjustment factors are derived from both Caltrans and Southern California Association of Governments (SCAG) Household Travel Survey GPS components.
[2] Total trips include all household trips by all modes of travel.
[3] Person trips include all household trips except walk, bicycle, airplane, and "other" mode trips.
[4] Driver trips include household automobile, pickup, RV, van, motorcycle, and truck driver trips.
[5] Only includes trips made by persons five years of age or older.

As expected, over eight in ten trips are made by personal vehicle (driver or passenger) while the public transportation mode share statewide is two percent. More than eight percent of all trips are walk trips.

Travel Mode Distribution	Weekday Percent
Vehicle Driver Trips	60.2%
Vehicle Passenger Trips	25.8%
Public Transportation Trips	2.2%
Bicycle Trips	0.8%
School Bus Trips	2.0%
Walk Trips	8.4%
All Others	0.7%

Survey journey-to-work statewide data are summarized below and compared to the 2000 Census Supplementary Survey where available. The commute mode distribution and other statistics are similar between sampled households and those that participated in the 2000 Census Supplementary Survey.

Journey-to-Work[6]	2000-2001 Survey	2000 Census Supplementary Survey
Trips 1 Hour or longer	10%	10%
Trips 45 minutes or less	88%	79%
Trips Less than 30 Minutes	58%	61%
Trips Less than 20 Minutes	42%	41%
Mean Commute Time (minutes) – 24 hours	27.6	26.7
Mean Commute Time (minutes) – 7-9 a.m.	24.6	Not Available
Average Vehicle Occupancy[7]	1.2	Not Available
Commute Mode (24 hour)		
Drove Alone	76%	75%
Carpooled/Vanpooled	15%	14%
Public Transit[8]	4%	6%
Walked	3%	3%
Bicycled	1%	1%
Other	1%	1%
Commute Mode (7-9 a.m.)		

[6] In the 2000 Census Supplementary Survey, "travel time to work" was reported in terms of the total number of minutes that it *usually took* (emphasis here) for the respondent to get from home to work during the reference week which was the week prior to the survey. In the 2000-2001 California Statewide Household Travel Survey, the travel time to work was calculated by subtracting the actual arrival time at work from the actual departure time from home during the assigned travel date (not usual but actual).
[7] Unlinked trips
[8] Public transit trips include modes such as local public bus, dial-a-ride/paratransit, light rail/streetcar, cable car, BART, Caltrain, and ferry. (Note that intercity bus, AMTRAK, and commercial airplane trips are not included as public transit trips in this report.)

Drove Alone	75%	Not Available
Carpooled/Vanpooled	14%	Not Available
Public Transit	5%	Not Available
Walked	4%	Not Available
Bicycled	1%	Not Available
Other	1%	Not Available

1. Survey Objectives and Overall Approach

This report documents the design, implementation and results of the 2000-2001 California Statewide Household Travel Survey (Statewide Travel Survey), conducted between October 2000 through December 2001. The survey is an essential element in determining statewide and regional travel patterns. The project was conducted under the auspices of California Department of Transportation (Caltrans).

The purpose of the study was to update the statewide database of household socioeconomic and travel information. In turn, updated data will be used to help refine travel estimates, models, and forecasts throughout the State. The study area consists of all 58 counties of the State grouped into 17 regions shown in the following table. Two survey instruments for each the household recruitment and data retrieval stages were used (see Appendices B and D). The resultant data set will be used to fulfill the model's functions of estimating trip generation and distribution, mode choice, and assignments.

Table 1.1
Region Definition

Region	County	Region	County
Western Slope/ Sierra Nevada	Amador	San Joaquin	San Joaquin
	Calaveras		
	Mariposa	San Luis Obispo	San Luis Obispo
	Tuolumne		
AMBAG	Monterey	Santa Barbara	Santa Barbara
	San Benito		
	Santa Cruz	Shasta	Shasta
MTC	Alameda		
	Contra Costa	Stanislaus	Stanislaus
	Marin		
	Napa	Tulare	Tulare
	San Francisco		
	San Mateo	Rural (all others)	Humboldt
	Santa Clara		Madera
	Solano		Nevada
	Sonoma		Sierra
SACOG	El Dorado		Kings
	Placer		Mendocino
	Sacramento		Lake
	Sutter		Tehama
	Yolo		Siskiyou
	Yuba		Inyo
SCAG	Imperial		Mono
	Los Angeles		Alpine
	Orange		Lassen
	Riverside		Modoc
	San Bernardino		Del Norte
	Ventura		Glenn
Butte	Butte		Plumas
Fresno	Fresno		Colusa
Kern	Kern		Trinity
Merced	Merced		
San Diego	San Diego		

The 2000-2001 California Statewide Household Travel Survey, like all recent household travel surveys, relied on the willingness of area residents to complete diary records of their daily travel over a 24-hour (weekday) or 48-hour period (Friday/Saturday or Sunday/Monday pair). Household recruitment was conducted through a "recruitment interview" in which respondents were informed of the survey, its purpose and to request respondents to complete the diaries. Data on households and household members were also collected during the recruitment interview.

Participating households were assigned a specific "travel day" or days to record their travel, which typically occurred 10-12 days after recruitment and during which household members were asked to record travel information in their diaries for a specified 24-hour or 48-hour period. Beginning the day after the assigned date, attempts to contact households were made to retrieve the diary information. A total of 24,049 households were recruited to participate in the study. Of these, 17,040 households (71%) completed travel diaries, and the information was retrieved from all household members regardless of age. The 17,040 households represent 40,146 persons, 33,540 vehicles and 134,173 trips. These particular figures are the actual number of records (unweighted and unexpanded) in the respective data files.

While the sample is a good representation of households within the state and within each region, weights were applied to bring the households into proportion to the distribution of households, by county, according to the 2000 Census. These weights were also based on household size, income, and vehicle ownership by housing own/rent status as obtained from the 2000 Census data files. A detailed description of the weights is provided in Chapter 6 - Survey Data Weighting and Expansion of this report. Except when noted, all data in this report are weighted. A separate independent weight, Trip Underreporting Factor, is detailed in the Global Positioning System (GPS) section of this report (Chapter 7).

The survey used a scientifically formulated sample design, appropriate instruments for data collection, a package of written materials and internet-based methods to communicate with survey respondents, a toll-free survey hotline, and data collection, processing, and reporting procedures.

The final report presents the results and describes survey execution. It is organized into chapters by major topic. The chapters include:

1. Survey Objectives and Overall Approach
2. Survey Instruments and Materials
3. Sampling Design and Procedures
4. Survey Pretest and Final Survey Design
5. Interviewer Training, Survey Methods, and Quality Control Procedures
6. Survey Data Weighting and Expansion
7. Global Positioning System (GPS) and Adjustment Factors for Trip Underreporting
8. Survey Results
9. Survey Quality Assessment
10. Survey Limitations
11. Recommendations for Future Survey Improvements

2. Survey Instruments and Materials

This chapter details the survey instruments and materials used in the project. The survey instruments were developed based on Caltrans' modeling needs and the required data variables were specified prior to the start of the project. The project included one survey instrument for each of the three data collection stages and materials were mailed to respondents immediately after two of the stages. The survey instruments and materials used in each stage are detailed below. The Global Positioning System (GPS) sub-task procedures are detailed in the GPS and Adjustment Factors for Trip Underreporting chapter of this report.

The survey followed a six-step process. 1) The recruitment call secured the household's participation in the study and obtained demographic information, which was used to 2) prepare personalized travel diaries for all household members. 3) The diaries were mailed to each member of a household and used during their assigned travel day to record all of their travel. 4) In addition, a reminder call was made to confirm receipt of the packet and answer any last minute questions prior to the assigned travel days. 5) Following the assigned travel day, a retrieval call was made to obtain the recorded information. 6) The retrieved data was edited and processed, then reported locations were geocoded to x/y coordinates.

Recruitment Interview. The purpose of the recruitment interview was to secure household participation in the study. The interview was also conducted using CATI. The questionnaire introduction was specifically designed to obtain agreement on participation. The other objectives of the recruitment questionnaire were to collect information on the characteristics of the household and the individual people in the household. The recruitment questionnaire is included in Appendix B of this report.

Respondent Material Mailing. The day following recruitment, the demographic information was used to prepare personalized diaries (either 24- or 48-hour) to send to each member of the household. The diary was designed to be used by the respondent as a memory jogger during the retrieval interview.

A personalized cover letter was also prepared and included in the packet, along with a "reminder sheet" reminding the household of its assigned travel date and to record their trips and activities in their diaries. These materials are included in Appendix C of this report. NuStats' Fulfillment Department mailed the packets from its office in Austin, Texas.

Reminder Call. The night prior to the assigned travel day, a reminder call was made to each household to confirm receipt of the packet and answer any last minute questions. If the packet was not received by this time, the address was re-confirmed and a new travel date was assigned and the diary packet re-sent.

Retrieval Interview. Using CATI, the interviewers collected all travel information recorded by respondents for the designated 24-hour or 48-hour travel diary period. The CATI program prompted interviewers to gather all pertinent information, as well as reference the same trips made by other household members. A look-up table of frequently visited locations aided with the retrieval process. The retrieval questionnaire is included in Appendix D of this report.

3. Sampling Design and Procedures

This chapter provides documentation of the sample design and procedure used in the survey. Details in this chapter include the definition of the sample universe for each of the 17 regions of the state and required sample size, sample selection, sample frame generation and sample preparation.

Survey Universe and Sample Size

The universe for the Statewide Survey is defined as all households located within the State of California. The state is also grouped into 17 survey regions. During the Request for Proposals (RFP) process, Caltrans specified the required number of samples for each of the 17 regions. Table 3.1 includes each region, corresponding Year 2000 estimated number of households, and required sample size for both weekday (24-hour, Monday through Friday) and weekend (48-hour, Friday and Saturday or Sunday and Monday) samples.

The required number of samples for each county was based on its proportional distribution of Year 2000 estimated number of households within its region. The only exception was for the SCAG region. The SCAG region used a disproportionate sample with an equal number (500) of samples required for each county.

Table 3.1
Region Definition and Sample Size

Region	County	Year 2000 Estimated Households	Weekday Sample Size (goal)	Weekend Sample Size
Western Slope/ Sierra Nevada	Amador	12,819	111	16
	Calaveras	17,553	152	9
	Mariposa	6,815	59	8
	Tuolumne	20,505	178	20
	Region Total	57,692	500	63
AMBAG	Monterey	108,675	255	38
	San Benito	15,410	36	18
	Santa Cruz	88,687	448	50
	Region Total	212,772	740	106
MTC	Alameda	515,170	209	113
	Contra Costa	353,547	144	98
	Marin	98,159	40	60
	Napa	45,115	18	52
	San Francisco	311,191	127	56
	San Mateo	266,130	108	39
	Santa Clara	568,257	231	46
	Solano	131,947	54	22
	Sonoma	169,923	69	15
	Region Total	2,459,437	1,000	501
SACOG	El Dorado	64,827	92	81
	Placer	90,431	128	28
	Sacramento	443,139	627	78
	Sutter	31,300	44	14
	Yolo	55,817	79	13
	Yuba	21,804	31	7
	Region Total	707,317	1,000	221

Table 3.1 (Continued)
Region Definition and Sample Size

Region	County	Year 2000 Estimated Households	Weekday Sample Size	Weekend Sample Size
SCAG	Imperial	52,405	500	0
	Los Angeles	3,178,460	500	0
	Orange	935,603	500	0
	Riverside	558,796	500	0
	San Bernardino	569,105	500	0
	Ventura	244,749	500	0
	Region Total	5,539,119	3,000	0
Butte	Butte	80,383	500	67
Fresno	Fresno	269,554	500	97
Kern	Kern	233,013	500	91
Merced	Merced	65,703	500	64
San Diego	San Diego	994,253	1,000	267
San Joaquin	San Joaquin	187,966	500	84
San Luis Obispo	San Luis Obispo	87,003	500	68
Santa Barbara	Santa Barbara	138,217	750	103
Shasta	Shasta	67,051	500	64
Stanislaus	Stanislaus	154,167	500	78
Tulare	Tulare	121,000	500	73
Rural	All Others	330,095	2,500	52
	TOTALS	11,704,742		2,000

The survey produced a sample size of 17,040 randomly selected households with an overall reliability of ± 0.8 percentage points at the 95% confidence level with respect to household level attributes. Table 3.2 summarizes the standard error for various sample sizes.

Table 3.2
Standard Error Rates at the 95% Confidence Level

Sample Size	Standard Error
500	± 4.5
750	± 3.7
1,000	± 3.2
2,500	± 2.0
3,000	± 1.8
17,000	± 0.8

Sample Selection

The survey employed a probability sample selection process to select households for inclusion in the study. The major requirement for probability samples is that the relative probability (or chance) that any household in the universe will be included is known. Once the sampling procedure is determined, selection of specific households for inclusion in the sample is left entirely to chance.

The type of probability sample used is stratified sampling in which the sample elements were drawn proportionately to the number of households for each county within a region. The sample was randomly generated across all telephone exchanges within each county.

The 2000 Census reveals that the overall percentage of households with telephones in the study area is 97.0%. As discussed in the weighting chapter of this report, the actual percentage of truly non-telephone households (non-episodic) is about half of the Census percentage. The Census takes a snap-shot picture (survey of a single day) of whether a household has phone service or not. Most households that did not have

telephone service on the day of the Census more than likely did not have service due to service interruption (non-payment) and had the service reinstated at a later date. Table 3.3 shows the percentage of households with telephones, by region.

Table 3.3
Telephone Coverage

Region	Year 2000 Estimated Number of Households	Percentage of Households w/Phone*
Western Slope /Sierra Nevada	57,692	96.2%
AMBAG	212,772	97.6%
MTC	2,459,437	98.2%
SACOG	707,317	97.0%
SCAG	5,539,119	96.8%
Butte	80,383	96.4%
Fresno	269,554	95.2%
Kern	233,013	94.1%
Merced	65,703	94.9%
San Diego	994,253	97.7%
San Joaquin	187,966	95.7%
San Luis Obispo	87,003	97.7%
Santa Barbara	138,217	97.6%
Shasta	67,051	95.9%
Stanislaus	154,167	96.9%
Tulare	121,000	93.3%
Rural	330,094	93.5%
Overall	11,704,742	97.0%

Source: U.S. Census, 1990

Sample Frame Generation

The sample frame for the survey included listed and unlisted telephone numbers. A "listed" telephone number is a telephone number for which a household address can be identified; an "unlisted" telephone number is one for which a household address can not be identified.

Both the listed and unlisted telephone numbers were generated using random digit dial (RDD) procedures. Using a telephone database that contains the universe of listed business and residential telephone numbers, NuStats identified all the working blocks for telephone numbers in the study area. For each working exchange/block combination a comprehensive analysis was conducted to determine its efficiency. Telephone companies reserve certain exchange/block combinations strictly for commercial assignments while others may have a mix of business and residential use.

In generating the *listed* sample, NuStats included in the sample frame those exchange/block combinations with a minimum 70% residential listing. However, all exchange/block combinations (including those that have less than 70% residential listings) were used to generate the "unlisted" sample. This assured that mixed-use developments (both commercial and residential use) were not excluded from the sample frame. Using a minimum 70% residential listing cut-off minimized time spent screening out businesses during the recruitment stage.

Unlisted telephone numbers were generated based on the telephone exchanges and blocks identified from the listed sample generation. Telephone numbers were randomly generated from these exchange/block combinations and then compared to all phone numbers listed (business/government and residential) in the 58 counties as identified in the telephone database. Any generated telephone numbers that are also listed within the database were eliminated from the sample frame thus providing assumed unlisted telephone numbers.

The incidence of unlisted telephone numbers for each region is shown in the following table.

Table 3.4
Unlisted Telephone Rate by Region

Region	Percentage of Unlisted Households
Western Slope / Sierra Nevada	39%
AMBAG	36%
MTC	36%
SACOG	37%
SCAG	32%
Butte	41%
Fresno	32%
Kern	35%
Merced	37%
San Diego	36%
San Joaquin	31%
San Luis Obispo	46%
Santa Barbara	36%
Shasta	33%
Stanislaus	33%
Tulare	40%
Rural	36%
Overall	34%

Source: Survey Sampling, Inc.

A recruitment goal of approximately 23,500 households assumes a retrieval rate (that is, percentage of recruited households that will be completed) of 72%. This recruitment sample size and retrieval rate goal produces a minimum of 16,990 completed surveys. The final retrieval rate was achieved, therefore, a total of 17,040 surveys were completed from a base of 24,049 recruited households (71%).

Sample Preparation

The sample was prepared for administration by organizing it into replicates. A replicate is a systematically selected sub-sample of the universe – for this study, it is each county within each region. The main benefit of using replicated samples is that the interviewers do not need to call the entire sample frame in order to ensure proper representation of the study area. When the quota of completed households is accomplished, it is only necessary to attempt to complete households in the current replicate that has been released or opened.

4. Survey Pretest and Final Survey Design

Survey Pretest

4.1 - Background

During the month of August 2000, NuStats conducted a pilot test for the 2000-2001 California Statewide Household Travel Survey. The objectives of the pilot test were to refine the survey procedures, survey materials and the Computer Assisted Telephone Interview (CATI) programs. The pilot test allowed the full evaluation of the survey procedures from the recruitment phase to the data processing phase prior to the implementation of the full study. Specific areas of assessment included whether or not respondents understood each question asked, the ease of use of the travel diary, the timing of the mailing materials, and the general flow of the CATI programs.

The objectives of the pilot test were to answer the following questions:

- Did each item elicit the range of responses expected? Were any responses out of range?

- Did the interviewer need to repeat any item, or did the respondent appear to be confused about the meaning?

- What was the level of non-response? Of actual refusals? How can response rates be maximized for the item? Can alternative procedures be introduced to compensate for refusals and other non-response?

- Was every feedback element from interviewers and monitoring staff fully and adequately addressed?

- Was the questionnaire item or the research procedure executed in the most efficient manner? Were productivity levels in line with the budgeted resources and time production estimates? Could they be done more quickly but with equal or higher quality?

- Were the geocoding match rates in line with contractual requirements? Were the address cues sufficient? How can more specific address information be secured?

- What issues resulted in exception reports? At what stage of the CDF process did they occur? Were all CDF-related timelines met? What adjustments to the data flow process are required?

The pilot test included a sample of 209 completed one-day diary surveys from households located in the Western Slope/Sierra Nevada region (Amador, Calaveras, Tuolumne and Mariposa counties), Fresno County and San Diego County. These three regions were selected because they represent relatively different areas with regard to population. The Western Slope/Sierra Nevada region represents the rural region of the state, San Diego as an urbanized area and Fresno in-between. A total of 57 completes were collected in the Western Slope/Sierra Nevada Region, 52 in Fresno County and 100 in San Diego County.

A total of 309 households were recruited with diary data collected from a total of 209 households. The retrieval rate was 68 percent, which is the number of completed households divided by the number of recruited households. Under normal interviewing protocols, which include additional callbacks to non-completed households, the rate should be as high as 75%. However, due to the short time frame of the pilot test, the standard of 12 attempts to contact each household was not possible, producing a lower than expected completion rate. During the full study, the standard number of attempts will be made to contact each household that will produce a completion rate of at least 75%.

4.2 – Pilot Test Limitations

Because the pilot test has such a small sample size at the regional level, caution must be taken when interpreting the quantitative outcomes or when making inferences about the individual regions in general.

The overall total of 209 completes produces a lower standard error than on a regional basis. However, since the full project is on a statewide level, statewide inferences should not be made based on the three regions from the pilot test since these regions have different demographic and travel behavior characteristics compared to the state as a whole.

4.3 – Sample Design and Generation

A total of 2,517 random telephone numbers were called to produce a total of 309 recruited households. Both the listed and unlisted telephone numbers were generated using random digit dial (RDD) procedures for each of the three regions included in the pilot. Using continually updated databases, which contain listed telephone numbers, NuStats identified all the working exchanges and working blocks for telephone numbers in the pilot study regions. NuStats then generated all the possible combinations of telephone numbers within these working exchanges and blocks. These numbers comprised the sampling frame for the listed telephone numbers.

Unlisted telephone numbers were generated based on the telephone exchanges and blocks identified from the listed sample generation. Telephone numbers are randomly generated from these exchanges and blocks then compared to all phone numbers listed in the three regions. Any generated unlisted telephone numbers that are also listed in the database were eliminated from the sample frame and assumed to be listed. These telephone numbers made up the unlisted portion of the sampling frame.

The sample was then organized into replicates. A replicate is a systematically selected sub-sample of the universe; it randomly includes sample units in the same proportions as are found in the entire sample. The main benefit of using replicated samples is that the interviewers do not need to call the entire sample frame in order to ensure proper representation of the study area. When the quota of completed households is accomplished, it is only necessary to attempt to complete households in the current replicate that has been released or opened. This increases the statistical rigor of the data by ensuring an equal chance of selection across all categories of households.

It is NuStats' intention that the data from the 209 pilot test completes will be included in the final data set if no major changes are made to the survey procedures or materials subsequent to the pilot test.

Table 4.1 shows, by household size and auto ownership, the distribution of recruited households. As shown, the zero-car cells, particularly for larger households, may be difficult to fill in certain areas that do not have high levels of public transit options available. Similar to other markets in which NuStats has conducted a household travel survey, zero-car cells are typically the most problematic due to their low incidence of occurrence and propensity to refuse to participate in surveys.

Table 4.1
Pilot Test: Household Size by Auto Ownership Distribution
Recruited Households
(n=309)

	Household Size				
Vehicles Available	1	2	3	4+	Total
0	12	1	1	0	14
1	55	24	3	9	91
2	15	74	23	27	139
3+	1	26	18	20	65
Totals	83	125	45	56	309

Table 4.2 shows the distribution of retrieved households and the percent of completes to recruited households. As shown in the table, larger households were more difficult to retrieve. Part of the reason for the low completion rate is the short length of time afforded during the pilot test. During the course of the full study, multiple call back attempts will be made during various times of the day and various days of the week.

During the full study, NuStats will closely monitor the overall distribution, with particular attention paid to zero-car cells and focus on problematic cells as needed.

Table 4.2
Pilot Test: Household Size by Auto Ownership Distribution
and Percent Complete of Recruits
Completed Households
(n=209)

	Household Size				
Vehicles Available	1	2	3	4+	Total
0	8 / 67%	1 / 100%	1 / 100%	0 / N/A	10 / 71%
1	45 / 82%	19 / 79%	3 / 100%	5 / 56%	72 / 79%
2	8 / 53%	58 / 78%	10 / 43%	13 / 48%	89 / 64%
3+	0 / 0%	16 / 62%	13 / 72%	9 / 45%	38 / 59%
Totals	61 / 73%	94 / 75%	27 / 60%	27 / 48%	209 / 68%

Other key statistics will also be monitored closely during the course of the full study. These include "Driver Trips Per Vehicles Available," "Driver Trips Per Vehicle in Use," "Driver Trips Per Household," and "Total Trips Per Household." These statistics are summarized in Table 4.3 below for each of the three pilot regions.

Table 4.3
Pilot Test: Key Statistics by Region
Completed Households
(n=209)

	Region		
Key Statistic	Western Slope/ Sierra Nevada	Fresno County	San Diego County
Driver Trips Per Vehicle Available	2.6	3.8	3.2
Driver Trips Per Vehicle In Use	4.1	4.9	4.4
Driver Trips Per Household	5.2	6.6	5.6
Total Trips Per Household	6.8	9.4	7.9

4.4 – Survey Materials

Three significant survey instruments are used for the study:

1. Recruitment questionnaire — Used to collect household and person demographic data and also elicit the household commitment to participate in the survey,

2. Travel diary packet — A packet that included a cover letter and a diary that serves as a memory jogger used by survey respondents to record key travel data so that subsequent reporting during "retrieval" is as accurate as possible, and

3. Travel data retrieval questionnaire — Used to collect travel data.

The recruitment and retrieval questionnaires were programmed in a Computer Assisted Telephone Interviewing (CATI) system. These programs eliminated the need for data entry as the telephone interviewers, themselves, input data as they conduct the interview. In addition, available responses to most questions were categorized to minimize the need for post-coding. Specific parameters were also set for particular questions, so that only valid codes could be entered.

4.5 – Pilot Test Administration

A total of ten interviewers were assigned to conduct each phase of the pilot test including recruitment and retrieval. One field coordinator and a data quality control technician supervised the telephone interviewers. The data quality control technician initially reviewed each completed survey (validation) while the final logic checks were conducted by the data services unit (edit check).

Recruitment

Pilot study recruitment was conducted between August 10-16, 2000. A total of 2,517 telephone numbers were dialed producing 309 recruited households. The average length of the recruitment interview was 15.6 minutes.

Weekday recruitment calling began at 5 p.m. and ended at 9 p.m. Pacific Standard Time (PST). Weekend (Saturday) calling began at noon and ended at 7p.m. PST. For the pilot test, the interviewers were able to complete 1.4 recruitment interviews per hour. This is lower than the anticipated rate of 1.8 per hour. As the interviewers become more familiar and comfortable with the interview script during the full study, higher productivity should be achieved.

Table 4.4 summarizes the outcomes of the recruitment calls.

Table 4.4
Pilot Test: Recruitment Call Outcomes
(n=309)

Call Outcome	Frequency
Eligible Units	
Recruited	309
Refused to participate	71
Partial Completes	29
SUB-TOTAL ELIGIBLE	*409*
Ineligible Units	
Disconnected number	367
Business / Government	89
Computer / fax line	89
SUB-TOTAL INELIGIBLE	*545*
Eligibility Unknown Units	
No answer	474
Call Back	590
Answering machine	457
Busy	42
SUB-TOTAL ELIGIBILITY UNKNOWN UNITS	*1,563*
GRAND TOTAL:	*2,517*

Using Council of American Survey Research Organization guidelines, the response rate is calculated as the number of "completed interviews divided by total eligible sample plus a proportion of eligibility undetermined sample." The portion of the undetermined sample is calculated from the proportion of eligible sample to the total of eligible and ineligible sample. Table 4.5 displays the sample dispositions used to calculate the response rate.

Table 4.5
Pilot Test: Sample Dispositions Used to Calculate Response Rate

Disposition Category	Disposition Outcome	Number	Percent
Eligible	Completed	309	
	Refused	71	
	Partial Complete	29	
Total Eligible Sample		409	43%
Non-Eligible	Business/ Government	367	
	Disconnected	89	
	Computer/Fax	89	
Total Non-Eligible		545	57%
Total Eligible and Non Eligible Sample		954	100%
Not Determined	No Answer	474	
	Answering Machine	457	
	Busy	42	
	Call Back	590	
Total Not Determined		1,563	
Total Sample Dialed		2,517	
Total "Eligible" Sample Dialed (409+41% of eligibility undetermined = 641)		1,050	
Recruitment Response Rate		29%	

The recruitment response rate is 29%, which is relatively low. It is assumed that during the actual study this response rate can be raised to as high as 35% due to NuStats' ability to "complete" the partially completed interviews and through systematic call backs of first refusals and eligibility undetermined. Again, due to the short time frame of the pilot, multiple calls to households were not feasible.

Data Retrieval

Data retrieval was conducted from August 22-31, 2000. Of the 309 recruited households, 209 households completed the entire study (i.e., each member of the household reporting each place visited on their assigned travel date). This is a completion rate of 68%. The average length of the retrieval interview for the one-day diary is 21.2 minutes.

Calls made to retrieve travel data were conducted at a rate of 1.5 completed surveys per hour. This rate is about where it is expected to be for the full study. For the two-day diary, the productivity is expected to be approximately 1.2 per hour.

Table 4.6 on the following page summarizes the final call dispositions during the retrieval phase. Of the 309 recruited households, 11% are callbacks. As in the recruitment phase, additional attempts will be made during the full study reducing the number of callbacks (partial completes, first refusals, no answers).

Table 4.6
Pilot Test: Retrieval Call Outcomes

Call Outcome	Frequency
Eligible Units	
Completed	209
Refused to participate	22
Pending (no answer, call backs, answering machines)	77
SUB-TOTAL ELIGIBLE	*308*
Ineligible Units	
Disconnected/non-working	1
SUB-TOTAL INELIGIBLE UNITS	*1*
GRAND TOTAL:	*309*

Overall Response Rate

In addition to the completion rates, an overall response rate is calculated. The overall response rate can be computed using the following formula:

$$RR = \left(\frac{a_1}{A_1 + (C_1 * ER_1)} \right) * \left(\frac{a_2}{A_2 + (C_2 * ER_2)} \right)$$

Where,

RR is the Overall Response Rate,
a_1 and a_2 are the number of completed surveys for each of the two phases,
A_1 and A_2 are the number of eligible telephone numbers for each of the two phases,
C_1 and C_2 are the number of eligibility unknown for each of the two phases, and
ER_1 and ER_2 are the eligibility rates for each of the two phases.

Using this formula, the Overall Response Rate is 20.0% (0.29 * 0.68). Although response rates are declining in the survey research field, this response rate is about average compared to typical results in previous household travel surveys. NuStats' goal is to obtain a response rate of about 25%. The response rate calculation uses the same formula prescribed by the Council of American Survey Research Organizations (CASRO).

4.6 – Geocoding

Home addresses were geocoded soon after sample generation. Home addresses that did not geocode were investigated and corrected during the recruitment interview. The goal of 100% geocoding of home addresses was achieved.

Work and school addresses for all household members were collected during the recruitment interview. Work and school addresses that did not geocode were investigated and corrected during the reminder calls.

Addresses of trip origins and destinations were geocoded via batch mode within 48 hours of data retrieval. All addresses that did not geocode during the batch mode were geocoded interactively by NuStats' GIS staff. We achieved our geocoding standards for the home and non-work and non-school addresses and fell short with the work and school address match rates. During the full study, call backs to households and other means of address research will be conducted so that the contractual standards are met.

Table 4.7 summarizes the geocoding match rates for the pilot address data.

Table 4.7
Pilot Test: Geocoding Outcomes

Address Type	Unique Locations	# Matched	Match Rate	Contractual Standard
Home	209	209	100%	100%
Work	191	164	89%	97%
School	98	92	94%	97%
Trips	798	741	96%	90%

Note: Match Rate is defined as "Number Matched" divided by "Total Locations."

4.7 – Data Processing

Data inspection was conducted in an on-going manner, from review of frequencies from the CATI program after the first few days of data collection to review of data during processing and editing.

Data checks (noted below) were conducted manually as they were conceived as a system of individual checks that would test data logic and consistency. After this review of the pilot data, the individual checks will be compiled into a master program for future large-scale (global) checks of the data.

Across all Files:
- Range of values for each data item is valid, including values for non-response.

Household File:
- Compare number of persons in household with number of person records in person file for that household.
- Compare number of vehicles in household with number of vehicle records in vehicle file for that household.

Person File:
- Verify that the number of places recorded for each person is at least as many as the number of places the respondent indicates visiting (at start of retrieval interview).
- Check to see if workers went to work on travel days. If not, reason must be provided.
- Check to see if students went to school on travel days. If not, reason must be provided.

Vehicle File:
- Check year of vehicle. Flag anything older than 1960 to verify.
- Check make and model. Flag if blank.

Trip File:
- Verify that household and person records exist for each sample number in the trip file.
- Check the travel times. Arrival at place (n) must be before departure from place (n). Arrival at place (n+1) must be after departure from place (n).
- Place numbers must be sequential and inclusive.
- Check to see if the person returned home at the end of each day. If not, flag as a possible missing trip.
- Verify that each place has address and trip data associated with it.
- Ensure that activities are consistent with reported location.

4.8 – Item Non-response

An objective of the pilot test was to estimate item non-response for data variables. Variables listed in the following tables in this section are the variables in which non-response is typically an issue. Each are categorized into one of the four main data types: Household, Person, Vehicle, and Trip Variables. Achieving high levels of responses to these questions will minimize non-response bias.

Household Variables

Household size, vehicle ownership, and physical home address are three data elements that require a 100% response rate. These variables determine sampling outcomes (household size and vehicles available) and trip origin (home address).

An important variable in which the non-response should be minimized is household income. During the pilot only 8% of all households refused or did not know the household's 1999 income. NuStats was able to achieve a high response rate for two reasons. First, this question can be asked in up to three phases. If a household refuses to answer during the recruitment phase, the question is asked again during the retrieval interview (these households are called by the most experienced interviewers who are specifically trained to handle households that appear to be the least cooperative as determined during the recruitment phase). If the question is not answered during the retrieval phase, it is asked again during the verification phase. At least 10% of all completed households are called back to verify that the interview was actually conducted and to verify certain data elements. Second, respondents are presented a more general income category (above or below a certain amount) first, which puts them at ease; then the more detailed categories are presented.

Table 4.8
Pilot Test: Household Item Non-Response
(n=209)

Variable	Percent Non-Response
Household income	8.1%
Number of other households that share this phone line	1.0%
Lack of phone service in the past 12 months	1.5%
Residential type	0.5%
Household ownership status	0.5%

Person Variables

Item non-response for person variables is extremely low. The variable with the highest level of non-response is age, as expected. Table 4.9 summarizes the item non-response for key person variables.

Table 4.9
Pilot Test: Person Item Non-Response
(n=451)

Variable	Percent Non-Response
Age	2.9%
Driver's license status	0.3%
Relation to respondent	0.2%
Primary economic activity	0.2%
Ethnicity	1.1%
Disability status	0.4%
Highest level of education attained	0.9%
Student status	0.2%
Primary job industry	0.4%
Primary occupation	0.7%
Hours worked at main job	0.7%
Location of primary job	0.2%

Vehicle Variables

Item non-response for the vehicle data items are also extremely low. The variable with the highest level of non-response is vehicle model, as expected. Table 4.10 below summarizes the item non-response for key person variables.

Table 4.10
Pilot Test: Vehicle Item Non-Response
(n=381)

Variable	Percent Non-Response
Year	2.1%
Make	0.3%
Model	4.2%
Body type	0.3%
Fuel type	0.0%
Owned or leased	0.0%
Year acquired	2.1%

Trip Variables

A 100% item response for key trip data elements was achieved. These include primary activity, trip mode, number in traveling party, number of family members on trip, arrival hour, and departure hour. The variable with the highest level of non-response is which vehicle was used on the trip (4.9%). All other variables are less than 0.5%.

Final Survey Design

Overall, the pilot test was successful in terms of the survey methods and the interviewer protocol. All items elicited a high response rate (respondents did not have difficulty in understanding the questions posed or in completing their diary). However, both the NuStats and Caltrans project managers agree that a more comprehensive interviewer training session was needed prior to the start of the full study. Because there was a need for a quick start-up of the project to make up for time lost during the contracting process, only a brief training session was possible. Prior to the start of the full study, a full-day briefing will occur. The briefing allows for practicing the interviewer script, role playing and preparing the interviewers on how to respond to specific questions from respondents. The training time also allows the interviewers to become more familiar and comfortable with the actual wording of the questions.

In this section, suggestions are provided for improvements to the survey process as well as the survey materials. The first section contains recommended changes, by survey task, while the second section contains recommended changes to the survey materials.

Survey Process

Recruitment

There was a general sense of eagerness to participate among respondents during the pilot, however, recruitment productivity was somewhat lower (1.4 recruits per hour) than what is expected for the full study. Although pilot test participation was high, relative to the short calling time frame, it is expected that recruitment productivity will increase to about 1.8 completes per hour. Repeated callbacks and refusal conversion attempts will minimize any response bias.

The seven days allowed between the mailing of materials and the assigned travel date appeared to be adequate for the full study. However, because the delivery of mail was lengthened due to the events of September 11, 2001, a larger window between the mailing and assigned travel dates was needed during the Fall 2001 data collection period.

Reminder Call

The reminder calls were also successful during the pilot. Of the 309 households recruited, and reminded, 156 (51%) were actually contacted and said they received their diary packet. Six households (2%) were contacted but did not receive their diary packet. These households were asked to record their trip information if they so chose. The pilot time frame did not allow reschedules; however, during the full study reschedules will be made for these outcomes. One hundred and thirty-four (43%) were answering machines and a reminder message was left that included our 1-888 number if they had any questions. The remaining 12 households (4%) were "no answer/no answering machine."

There were no changes made to the reminder call process prior to the start of the full study.

Retrieval Call

During the pilot, it was suggested that during the address collection, all available variables be collected. For example, the nearest intersecting street and landmark variables should be collected even if an exact address is collected so that in instances where the exact address is not geocodable there are other variables from which to code. This was implemented during the full study. However, interviewers were trained to listen for cues from the respondent and not collect more address data than necessary for geocoding if the respondent indicated any frustration of providing more than the minimum requirement.

The collection of activity start and end times posed some confusion among respondents as well as interviewers. The confusion stems from how to distinguish the difference between trip times and activity times. The activity duration can be calculated based on trip times so the activity start time and end time were eliminated from the interview prior to the start of the full study.

The Caltrans project manager monitored a few surveys in progress during the pilot and was pleased with how the interviewers conducted the survey. However, he suggested that the interviewer stick to the interview script verbatim. This request was covered during the interviewer training session prior to the start of recruitment.

Materials

There were virtually no changes to the survey materials for the full study except for one change to the diary. The item "Travel by car, bus, walk, etc." from activity list in the diary was deleted since it is actually a mode and not an activity.

There were no changes made to the brochure or cover letter prior to the start of the full study.

5. Interviewer Training, Survey Methods, and Quality Control

The purpose of this chapter is to review the interviewer training procedures, methods used to conduct the survey and the quality control procedures.

Interviewer Training

In addition to receiving rigorous and detailed industry-wide training NuStats provides to all interviewers, all interviewers were also required to successfully complete a project-specific training. The project manager and data collection manager designed a project-specific Interviewer Training Manual. All NuStats interviewers are trained using the most up-to-date materials and methods prescribed by the Marketing Research Association (MRA). This training takes place throughout the year and is unrelated to the project specific training that is the focus of the Interviewer Training Manual. The manual covers both general and specific information related to all data collection tasks for the 2000-2001 California Statewide Households Travel Survey.

Despite the fact that NuStats employs the largest, permanent interviewing staff trained in the conduct of household travel surveys in the country, each project is approached as a new project and all nuances explicitly addressed as if for the first time through the use of interviewer training manuals. This ensures that all interviewers understand the importance of the study and the need to collect complete and accurate data. A beginning concept that all interviewers understand at NuStats is the importance of "first contact" with each household. Knowing this "first step" into each household is the largest contributor to the household actually completing the survey.

Beginning with a project overview, including specific study objectives, the manual guides both interviewers and supervisors through all procedural issues including timelines, quotas and survey procedures. The manual also clearly identifies all survey management personnel as well as specific client contact information (each Caltrans district office contact person and the Caltrans project manager).

The manual then moves forward with task-specific training of recruitment interviews, refusal procedures and conversions, diary package contents and procedures, reminder calls and data retrieval methods.

NuStats firmly believes that open and clear communication from the onset of each project strengthens all training and data collection efforts. Therefore, all key personnel associated with the project, both internally and externally, participated in project start-up briefings and training sessions held in NuStats' data collection facilities.

In addition to the project training covering interviewing protocol, more specific "local knowledge" training was provided by the Geocoding Task Manager. This course included a Power Point presentation with specific tables and maps showing political geography, geographic terrain, transportation networks, major trip attractions and unusual or easily mispronounced locations or names.

The third part of the training dealt with the intricacies of the survey instruments themselves, with a separate training session held for each survey instrument involved. Details ranging from termination points and qualifiers for eligible respondents, to a careful review of skip patterns and rotations to group reviews of probing and clarifying techniques as they apply to the questionnaire were covered in great detail by project trainers. Mock interviews were conducted to help the interviewers quickly become familiar with the survey instruments, glossary of terms required for this project, and areas where the respondent might need further explanation.

The fourth training module addressed common questions and how to maximize respondent participation in the survey. After passing a project-specific test, interviewers began work on the project and were monitored frequently and received specialized one-on-one training with supervisors and the project manager. Project

trainers and supervisors regularly debriefed with interviewers and other project staff to keep everyone informed of any improvements to the interviewing process.

NuStats understood that respondents were going to refuse to participate in the survey. Since refusals are unavoidable, NuStats prepared for a certain percentage of refusals to occur during each stage of the interviewing process. The most experienced interviewers were trained to specifically handle refusals. Only those surveyors were allowed to call back respondents who refused to attempt to convince them to participate.

Survey Methods

As detailed in the "Survey Instruments and Materials" chapter, the survey process followed a six-step plan.

1) The recruitment call secured the household's participation in the study and obtained demographic information, which was used to

2) Prepare personalized travel diaries for each household member.

3) The diaries, along with other materials, were then mailed to each member of a household and used during their assigned travel day to record all of their travel.

4) In addition, a reminder call was made to each household one to two days prior to the assigned travel date to confirm receipt of the packet and answer any last minute questions.

5) Following the assigned travel day, retrieval calls were made to obtain the recorded information. Several techniques were employed during the retrieval interview to help ensure that all trips were accounted for. These included a simple question of "did you make any stops along the way" as each new location was reported. Tracking whether any other household members also went on a given trip helped to ensure consistency within the household records as well as providing a method for ensuring that each household member then reported the shared trip. Proxy reporting and diary usage were also tracked for each respondent.

6) Data are edited and processed, and locations geocoded to x/y coordinates

These six steps comprise the most visible tasks involved in the survey process. However, there were seven additional "behind the scenes" steps involved. The survey followed a Continuous Data Flow, or CDF, process. This process was created for use in the New York Transportation Commission/New Jersey Transportation Planning Agency Household Interview Survey and has been continually refined in the past five years. Its most recent application was in the Mid-Ohio Area Household Travel Survey, in which 5,300 households were interviewed in the first quarter of 2000.

The CDF process has 13 essential stages each associated with a key aspect of the sample progression. Within each stage, there are also criteria that specify the standards by which sampled households can move to the next stage of the project.

The following table is provided to document the CDF stages from sample allocation to timely data delivery. The progression criteria are stated in the third column. Two types of reports are used to monitor progress: production reports show movement of the data (how many interviews completed last night, geocoding progress, etc.) and Exception Reports show lack of movement – how many households could not be geocoded and therefore did not move to the next CDF stage? Both are critical to successful completion of the project.

Table 5.1
Continuous Data Flow (CDF) Process

Stage	Day	Stage Description	Progression Criteria
1	1	Generate Sample	None
2	2	Geocode Home Addresses	▪ Geocoded addresses go to Stage 3 ▪ Unmatched (geocode) listed addresses and unlisted addresses go to Stage 3
3	3	Recruitment Interview – Sampled households are contacted to secure participation in the study. Those who agree to participate provide demographic data and are assigned a travel day.	▪ If the interview is completed, goes to Stage 4 ▪ If the interview is not completed, exception report is generated ▪ If interview is not attempted, sample status is updated and sample is scheduled for callback according to sample management rules
4	4	Geocode Habitual Addresses – work and school addresses are geocoded	▪ If address geocodes, goes to Stage 7 ▪ If address does not geocode, exception report generated and also proceeds to Stage 7 but flagged with address information need.
5	10	Diary Placement – A personalized diary packet is prepared and mailed to each recruited household.	▪ If packet is mailed, goes to Stage 6 ▪ If packet is not mailed, exception report generated to indicate reason
6	14	Reminder Call – Recruited households are contacted to confirm receipt of diary packet and remind about upcoming travel day(s).	▪ If household is ready, goes to Stage 7 ▪ If household needs new packet, goes to Stage 5 ▪ If household is rescheduled, can go to Stage 5 or 7 ▪ If household refuses, exception report is generated and assigned to interviewer specializing in refusals
7	15/16	Travel Day – Household members record travel on assigned day(s).	None
8	16/17	Retrieval Interview – The first retrieval call is placed the day following travel or at a respondent-designated time. Contractual freshness standards control the length of time between travel days and data retrieval.	▪ If household provides data according to definition of "complete", goes to Stage 9. ▪ If household provides partial data, exception report is generated and household does not progress. ▪ If household did not record travel data and is rescheduled, can go to Stage 5 or 7. ▪ If household refuses, exception report is generated and assigned to interviewer specializing in refusals.

Table 5.1 (Continued)
Continuous Data Flow (CDF) Process

9	16/17	Field Edits – the night the retrieval interview is completed, work is checked for completeness.	▪	If work meets standards, goes to Stage 10.
			▪	If work does not meet standards, exception report is generated and household is assigned for callback / correction
10	16/17	Data Processing – at the conclusion of each data collection shift, all data are processed and prepared for edit check and geocoding.	▪	If processed data meets completeness standards, goes to Stage 11.
			▪	If processed data does not meet completeness standards, exception report is generated and household is assigned for correction / callback
11	17/18	Geocoding of Trip Ends – all new address information (new or updates to previously collected information) is geocoded through both batch and interactive processes.	▪	If geocoded, goes to Stage 12
			▪	If not geocoded, exception report is generated and household assigned for correction/callback
			▪	Daily reports monitoring hit rates
12	17/18	Data Quality Checks – all data is subjected to visual inspection and edit check program to ensure quality standards and data specifications are met.	▪	If passes, goes to Stage 13
			▪	If fails, exception report is generated and household assigned for correction/callback
			▪	Daily reports monitoring pass rates
13	22/23	Process complete – data ready for delivery.	▪	If process complete, data flagged for delivery and process ends.
			▪	If process not complete and time thresholds crossed, exception report is produced and data specialist addresses household to ensure data movement.

Following the data collection procedures, the addresses collected during recruitment and retrieval were geocoded to x/y coordinates. More detail about this stage follows.

Geocoding

All locations were geocoded using Arc View 3.1 using the June 2000 Geographic Data Technology's (GDT) Street Centerline Coverage Files (provided by Caltrans) and the Thomas Brothers Coverage Files provided by the Southern California Association of Governments (SCAG). Home addresses were geocoded for listed households soon after sample generation. Home addresses that did not geocode were investigated and corrected during the recruitment interview. Each of the 17,040 household addresses was geocoded (100% match rate).

Work and school addresses for all household members collected during the retrieval interview were also geocoded. Addresses that did not geocode were researched through the Internet or through callback to the respondent. Ninety percent of the all work and school addresses traveled to are geocoded.

Addresses of non-home, work or school trips were also geocoded. Of the 50,336 trip addresses (non-home, work or school), 94% were geocoded.

Geocoding occurred at three distinct stages in the survey after sample generation, recruitment, and retrieval.

▪ All listed sample (home addresses) will be geocoded immediately after sample generation.

▪ All home addresses not yet geocoded and habitual work and school addresses will be geocoded subsequent to the recruitment interview. Possible problem addresses will be identified and resolved as much as possible by using other location reference data such as the Caltrans geocoding

database, Zip2.com, and various map/reference lists (hard copy or electronic). Remaining problem habitual work and school addresses will be clarified during the retrieval interview.

- Trip end addresses will be geocoded within 72 hours of the retrieval interview. The same process will be used to identify and resolve the problem addresses.

Each stage will be unique, but all four will share the same underlying principles. The following section describes the generic geocoding process.

Prior to geocoding, electronic geographic coverage files will be prepared in Environmental Systems Research Institute's (ESRI) ArcView GIS Software. This step will include joining county street coverage files together, setting the properties for matching, and indexing the files.

The basic geocoding process consists of four steps. These steps will be performed each time addresses are matched to the geographic coverage files. The following outline describes these steps generally.

1. **Prepare geocoding file.** As addresses are submitted for geocoding, a table of address information will be created in dBase format with a field containing concatenated address data. This table will be sorted in various formats to make global changes. Global changes include correcting misspelled place names, misspelled city names, and correcting any other global address problems. The file will then be imported into ArcView for geocoding. Additional information such as traffic analysis zones are loaded in ArcView.

2. **Geocoding.** Batch and/or interactive geocoding will be performed on all addresses in the files. This will include all three address types (home, habitual school and work, and trips). A batch run is an automated process, and interactive sessions will be used to geocode addresses one at a time. In this process, the sensitivity measure will not be set to less than "85% corrective." Addresses will receive a status (AV_STATUS) variable of "M" for matched, "O" for out of area, "U" for unmatched. Other variables will be added as needed, and will be specified in the data matrix.

3. **Attach coordinates.** After addresses are geocoded, ArcView will calculate and pull longitude and latitude coordinates for the matched cases in decimal degrees to five decimal places. Any additional information such as TAZ and geocoded zip code and city were also added. Then, the geocoded file is saved and exported to a tabular data file that will be used to update a master data file.

4. **Address research.** The unmatched cases will require further research efforts to obtain the needed AV_STATUS of "M" (matched), and manual address research efforts will be performed. Addresses that are not matched will be researched and checked against a large array of materials, including:

   - Internet Directories
     - ✓ Zip2.com (online directory of all U.S. phone books containing schools, restaurants, shops, and other place names)
   - Electronic Directories
     - ✓ Street Atlas USA 7.0 (DeLorme)
     - ✓ Select Phone 2000 (ProCD)
   - Maps, Atlases, Gazetteers, and Street Finders
     - ✓ County and City Street Guides and Maps
     - ✓ Thomas Brothers Atlases
   - Telephone Directories
     - ✓ All available telephone directories

Steps 2 through 4 will be repeated until the desired percentage of addresses are geocoded.

NuStats geocoded 100% of the Household addresses. NuStats geocoded a minimum of 95% of all Work and School trip locations. NuStats geocoded at least 90% of all non-home, non-work activity trip locations to year 2000 Census tracts. NuStats geocoded 90% of all non-home, non-work activity trip locations to

latitude/longitude coordinates except when the coverage is missing a street or street segment in the street files.

While blank addresses will be excluded, NuStats conducted several immediate checks on the data during recruitment and retrieval interviews so that the surveyors were not allowed to enter a blank address field. The surveyor is at least required to type in "don't know/ refused" in order for CATI to accept the response. NuStats, with five years of continuous geocoding experience in ArcView, has standard quality control checks. These detailed quality control checks were developed to ensure the most accurate data possible. Details of the geocoding quality control procedures are provided in the next section following the data quality control procedures.

Quality Control

The collected data were subjected to a rigorous edit check program, which performed automated global quality control checks of the data. These checks included both within file checks (intrafile) for consistency, as well as cross-file checks (interfile) for logic and compatibility. For example, the edit check program confirmed that all responses were appropriate (e.g., if a household reported not having a car in the household data file, all vehicle variables should be blank in the vehicle data file). Additionally, if the household reported having four vehicles, the program checked to confirm that there were four vehicle records.

The following are the general descriptions of the edit check programs implemented for each of the data files. A complete list of edits follows the general descriptions.

Across all Files:

- Range of values for each data item is valid, including values for non-response (logic: responses cannot be outside range).

Household File:

- Compare number of persons in household with number of person records in person file for that household.
- Compare number of vehicles in household with number of vehicle records in vehicle file for that household.

Person File:

- Check to see if the number of persons indicated in the household file matches number of person records.
- Check to see if persons traveled on travel days. If not, reason must be provided.
- If person is not licensed, check to make sure there are no trips in which they were a driver.

Vehicle File:

- Check year of vehicle. Verify if year is 1960 or earlier.
- Check make and model. Flag if blank.

Trip File:

- Verify that each person has at least one place per day.
- Verify that household and person records exist for each sample number in the trip file.
- Check the travel times. Arrival at place (n) must be after departure from place (n-1). Arrival at place (n+1) must be after departure from place (n).
- Place numbers must be sequential and inclusive.
- Check to see if the person returned home at the end of each day. If not, flag as potential missing trip.

- Verify that each place has address and trip data associated with it.
- Ensure that activities are consistent with reported location.

Specific edit checks are provided in the following table.

Table 5.2
Edit Checks

Household size does not equal total person file record count
Total vehicle count does not equal total vehicle file record count
Trip summary does not equal total trip file record count
Income is missing or refused (verify from all stages)
Arrival time before departure time or times missing
No driver reported for an auto trip
Student Information not complete
Worker information not complete
Transit information not complete
Parking information not complete
Unemployed person reports work activity
Loop trip – trip began and ended at the same place
Occupation not given or out of range
Household address not geocoded
Household members traveled together but data are inconsistent
Odometer reading is excessive
Vehicle year <1960
Auto driver is underage
Auto driver is unlicensed or unknown
Place numbers not sequential and inclusive
Person didn't return home at the end of day – check for missing trip

Geocoding Quality Control

Quality control procedures to check the accuracy of the geocoding were conducted. The main procedure involved sorting geocoded locations by county, then displaying all geocoded points for a particular county using the county coverage file. Any points falling outside the county boundaries were verified and re-geocoded if necessary. The final data file contains a geocoding quality control variable that identifies the action taken on a particular record, the quality control check performed and/or the outcome of the check.

Specifically, the quality controls included:

- Blank records are flagged in CATI and are not sent for geocoding until a proper address has been collected.

- A random selection of 5% of the geocoded address file is reviewed in detail to ensure proper placement of the overall latitude/longitude points. This entails using ArcView and displaying the points on the street layer and comparing the points with DeLorme.

- Daily tracking reports on the status of the overall geocoding including the quantity, quality and match rates.

- Since a cross-street geocode does not reference a zone (zip code or city) in ArcView, all cross-street geocodes are queried and analyzed to ensure proper placement of the geocodes.

- After completing a geocoded file, the geocoded zip code, and geocoded city is attached to the file. This is used to determine the four codes used in the "quality control" flag field.

- A "quality control" flag field is added to the geocoded file for attaching a code to each record in the geocoded file. These include:

  1 = given zip code matches geocoded zip code and given city matches geocoded city

  2 = given zip code matches geocoded zip code

  3 = given city matches geocoded city

  4 = Visual confirmation required - these records are imported back into ArcView and manually displayed, queried off by city, and thoroughly reviewed to ensure accurate geocodes.

A second electronic check on the data involved calculating travel speeds and comparing them against mode-specific standards. A rate of speed is calculated and compared to a predetermined range of speeds deemed appropriate for each mode of travel. For multi-modal trips, a hierarchy of modes is often established and the rate of travel is subjected to the standards of the dominant mode.

This check was performed on all trip records for which both the origin and destination were successfully geocoded. The x/y coordinates were used to calculate a trip distance, while the reported travel times were used to calculate travel time.

The process was as follows:

1) *Create File.* A file is created from the trip file that places the origin, destination, and travel data in one record (as opposed to two records in the trip file).

2) *Calculate Distance.* The distance formula is used and thus the following variables are added to the speed check file. The following is used to calculate distances when coordinates are given in degrees of latitude and longitude:
$$d = \sqrt{(x_o - x_d)^2 \cos^2[(y_o + y_d)/2] + (y_o - y_d)^2}$$

where

x_o = longitude of origin
x_d = longitude of destination
y_o = latitude of origin
y_d = latitude of destination

The x and y coordinates are translated into decimal degrees before running this process and thus this formula yields a distance in decimal degrees. This is then converted to miles by multiplying the decimal degree distance by 69.1105 (factor that changes decimal degrees to miles on the curvature of the Earth).

3) *Calculate Travel Time.* The trip duration (expressed in minutes) is divided by 60 to get trip time in hours.

4) *Calculate Speed.* Miles are divided by hours to calculate the travel speed.

5) *Compare Calculated Speed to Mode Thresholds.* The calculated speed is then compared to "reasonable" speed thresholds. Those trips with speeds within the bounds are acceptable, those outside are flagged for a check on time rounding. The proposed thresholds for this project include:

  Auto trips ....0 to 70 mph

  Bus.............................0 to 35 mph

  School Bus ....0 to 45 mph

  Bicycle ....0 to 15 mph

  Walk...........................0 to 10 mph

6) *Determine the Effect of Time Rounding on Trips with Speed Violations.* Given the variations in reporting time as compared to the preciseness of calculated distance, a large proportion of speed violations actually result from respondents misreporting time. As such, the next step in the process is to vary the trip duration by up to 15 minutes to determine if that slight rounding would result in the speed becoming reasonable for the reported mode. Any trip records with speed violations that cannot be attributed to time rounding are flagged for visual inspection.

Visual Inspection. The remaining cases are then checked for these characteristics – respondent reporting incorrect mode, incorrect trip times, or reporting traveling to the same place consecutively (same shopping center or business center) thus, creating a distance of 0.

Once the data has passed all checks, the file is ready for a final pass prior to delivery. The purpose of this final pass is to eliminate duplicate geocodes for a single location, and ensure proper spelling of place names and cities.

6. Survey Data Weighting and Expansion

Calculation of Weights

The 2000-2001 California Statewide Household Travel Survey required predetermined sample sizes, by region, to meet statistical reliability requirements within each region. However, this requirement results in non-proportional sampling when analyzing at the statewide level.

Statistical weights were used to adjust the sample proportions to actual proportions as determined by most recent Census data available (actual Year 2000 Census or 2000 Supplemental Survey). The weight variable for the 2000-2001 California Statewide Household Travel Survey is comprised of five factors in which their product adjusts the survey data for the following:

1. Vehicle availability by own/rent status

2. County weight (proportional to households statewide)

3. Household size distribution

4. Income distribution

5. Normalization of weight factors

The weighting process includes the development of these factors, then multiplying them together to determine the "final" weight. While the household data file contains each of the five factors for documentation purposes, only the final weights (final expansion factors), as discussed at the end of this chapter should be applied when analyzing the data. There are three different final expansion factors including Weekday, Weekend, and Seven Day (Weekday and Weekend).

1. Vehicle availability by own/rent status

The calculation of this weight factor (FACTOR1) is based on the 2000 Census Supplementary Survey and actual survey data. For each region, the survey distribution of vehicle availability and own/rent status is compared to the Census distribution for each cell cross-classification. Tables 6.1 and 6.2 summarize the weights for each region based on Census data and survey data, respectively. Table 6.3 displays the weights for each region. The weights for each cell are calculated using the following formula:

$$\text{Vehicle Availability by Own/Rent Weight} = \frac{\text{Vehicle Availability by Own/Rent Weight (Census)}}{\text{Vehicle Availability by Own/Rent Weight (Survey)}}$$

For example, at the statewide level 44.8% of all households have two vehicles in owner occupied households. However, 46.4% of the survey sample falls into this cross-classification cell. A weight factor of 0.96552 is calculated as 44.8% (Census) divided by 46.4% (Survey). As in all calculations of weight factors, a weight of less than 1.00000 indicates more samples were collected compared to Census parameters. Weights greater than 1.00000 indicate fewer samples were collected in the household travel survey compared to the Census data.

Table 6.1
Vehicle Availability by Own/Rent Cross-Classification - Census

Vehicles	Owner Occupied Housing Units				Renter Occupied Housing Units			
	0	1	2	3+	0	1	2	3+
California	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
Region								
Western Slope/Sierra Nevada*	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
AMBAG	3.6%	20.2%	44.7%	31.5%	15.4%	43.4%	30.8%	10.4%
MTC	3.0%	24.3%	43.2%	29.4%	17.7%	43.1%	27.8%	11.4%
SACOG	2.2%	28.5%	45.9%	23.4%	12.6%	50.8%	29.4%	7.3%
SCAG	3.0%	24.4%	45.3%	27.4%	15.4%	45.9%	30.7%	8.0%
Butte*	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
Fresno	2.8%	29.5%	42.9%	24.8%	20.3%	50.4%	24.0%	5.3%
Kern	3.1%	28.3%	44.5%	24.2%	15.1%	48.4%	28.5%	8.0%
Merced*	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
San Diego	2.4%	25.7%	46.1%	25.8%	10.5%	45.6%	34.7%	9.2%
San Joaquin	2.6%	20.9%	47.2%	29.4%	12.4%	41.2%	35.2%	11.1%
San Luis Obispo*	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
Santa Barbara	3.8%	23.1%	44.4%	28.8%	5.9%	44.9%	35.0%	14.2%
Shasta	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%
Stanislaus	3.3%	18.7%	49.1%	28.8%	11.5%	40.0%	37.4%	11.1%
Tulare	3.5%	26.2%	46.1%	24.2%	15.4%	48.9%	28.3%	7.4%
Rural*	3.0%	24.5%	44.8%	27.7%	15.6%	45.4%	30.1%	9.0%

Source: 2000 Supplementary Survey, U.S. Bureau of the Census.
*Census data were not available for these regions at the time of writing this report; therefore, statewide percentages were used.

Table 6.2
Vehicle Availability by Own/Rent Cross-Classification - Survey

Vehicles	Owner Occupied Housing Units				Renter Occupied Housing Units			
	0	1	2	3+	0	1	2	3+
California	1.3%	24.0%	46.6%	28.1%	9.6%	46.3%	33.8%	10.2%
Region								
Western Slope/Sierra Nevada	1.1%	21.1%	45.1%	32.7%	8.0%	42.3%	38.1%	11.7%
AMBAG	1.2%	23.7%	45.8%	29.3%	7.7%	44.1%	33.8%	14.4%
MTC	1.4%	22.6%	48.8%	27.2%	13.2%	47.0%	32.1%	7.8%
SACOG	1.0%	24.9%	48.8%	25.3%	10.6%	47.6%	32.7%	9.1%
SCAG	1.6%	25.3%	47.0%	26.2%	10.2%	48.7%	32.1%	9.0%
Butte	1.2%	22.0%	45.3%	31.5%	9.5%	45.2%	35.9%	9.5%
Fresno	2.4%	19.0%	61.9%	16.7%	N/A	75.0%	25.0%	N/A
Kern	1.4%	26.6%	44.1%	27.9%	11.8%	51.6%	23.7%	12.9%
Merced	2.5%	24.3%	48.0%	25.1%	10.8%	48.4%	32.3%	8.6%
San Diego	1.3%	29.6%	48.5%	20.6%	5.6%	49.8%	33.8%	10.8%
San Joaquin	1.0%	25.9%	48.6%	24.5%	13.5%	44.9%	33.7%	7.9%
San Luis Obispo	.6%	22.9%	49.7%	26.8%	7.0%	39.5%	39.5%	14.0%
Santa Barbara	1.2%	26.2%	46.3%	26.2%	8.6%	45.2%	34.9%	11.3%
Shasta	1.1%	25.4%	39.0%	34.5%	12.9%	43.5%	33.9%	9.7%
Stanislaus	2.1%	29.1%	44.7%	24.1%	7.3%	46.3%	31.7%	14.6%
Tulare	2.1%	22.8%	47.4%	27.7%	8.9%	51.1%	31.1%	8.9%
Rural	1.1%	28.2%	43.3%	27.4%	12.0%	43.7%	34.5%	9.9%

Table 6.3
Vehicle Availability by Own/Rent Cross-Classification Weight Factor

Vehicles	Owner Occupied Housing Units				Renter Occupied Housing Units			
	0	1	2	3+	0	1	2	3+
California	2.00	0.96	0.97	1.05	1.61	0.94	0.93	0.92
Region								
Western Slope/Sierra Nevada	2.70	1.16	.99	.85	1.96	1.07	.79	.77
AMBAG	3.12	.85	.98	1.07	2.01	.98	.91	.72
MTC	2.09	1.07	.89	1.08	1.34	.92	.87	1.47
SACOG	2.12	1.15	.94	.93	1.19	1.07	.90	.81
SCAG	1.92	.97	.96	1.05	1.51	.94	.96	.89
Butte	2.69	.87	1.03	1.01	1.30	1.04	.87	.91
Fresno	1.18	1.55	.69	1.49	N/A	.67	.96	N/A
Kern	2.25	1.07	1.01	.87	1.28	.94	1.20	.62
Merced	1.18	1.01	.93	1.10	1.45	.94	.93	1.04
San Diego	1.87	.87	.95	1.25	1.88	.92	1.03	.85
San Joaquin	2.55	.81	.97	1.20	.92	.92	1.04	1.41
San Luis Obispo	4.92	1.07	.90	1.03	2.22	1.15	.76	.64
Santa Barbara	3.10	.88	.96	1.10	.69	.99	1.00	1.26
Shasta	2.66	.96	1.15	.80	1.21	1.04	.89	.93
Stanislaus	1.55	.64	1.10	1.19	1.57	.86	1.18	1.76
Tulare	1.65	1.15	.97	.87	1.73	.96	.91	.83
Rural	2.51	1.11	.99	.88	1.65	1.00	.84	.95

2. County Weight

The calculation of this factor is a fairly straightforward process. First, the natural or proportionate distribution of the households, by county, is considered. The proportional distribution for each county is calculated as the number of Year 2000 Census households for the county divided by the total number of households in the State.

The weight factor is then calculated for each county by dividing the number of proportional sample size by the distribution of samples as calculated by the first weight factor (vehicle availability by own/rent status). For example, in Fresno County the final weekday sample size was 575 as shown in Table 6.7. One of the eight weight factors for Fresno in Table 6.3 was applied to each household in Fresno County, depending on the number of vehicles available and the own/rent status. The resulting number of weighted samples in Fresno County, when this first weight is applied, is shown in Table 6.4 (n=620).

Table 6.4
County Weight Factors

Region	County	Year 2000 Households	Proportional Sample Size	Factor 1 Weighted Sample Size (Weekday Travel)	FACTOR
Western Slope/	Amador	12,759	11	130	0.147097
Sierra Nevada	Calaveras	16,469	14	168	0.146923
	Mariposa	6,613	6	90	0.110125
	Tuolumne	21,004	18	301	0.104584
	Region Totals	56,845	47	689	
AMBAG	Monterey	121,236	101	279	0.651266
	San Benito	15,885	13	159	0.149734
	Santa Cruz	91,139	76	445	0.306955
	Region Totals	228,260	191	1,572	
MTC	Alameda	523,366	437	375	2.091728
	Contra Costa	344,129	287	187	2.758104
	Marin	100,650	84	125	1.206799
	Napa	45,402	38	97	0.701511
	San Francisco	329,700	275	164	3.013049
	San Mateo	254,103	212	162	2.350855
	Santa Clara	565,863	472	256	3.312855
	Solano	130,403	109	154	1.269106
	Sonoma	172,403	144	132	1.957501
	Region Totals	2,466,019	2,059	704	
SACOG	El Dorado	58,939	49	104	0.849377
	Placer	93,382	78	157	0.891445
	Sacramento	453,602	379	507	1.340905
	Sutter	27,033	23	93	0.435655
	Yolo	59,375	50	96	0.926965
	Yuba	20,535	17	36	0.854916
	Region Totals	712,866	595	732	
SCAG	Imperial	39,384	33	507	0.116424
	Los Angeles	3,133,774	2,616	590	7.960617
	Orange	935,287	781	584	2.400287
	Riverside	506,218	423	575	1.319474
	San Bernardino	528,594	441	523	1.514787
	Ventura	243,234	203	639	0.570498
	Region Totals	5,386,491	4,496	3,418	
Butte	Butte	79,566	66	555	0.214865
Fresno	Fresno	252,940	211	620	0.611444
Kern	Kern	208,652	174	584	0.535477
Merced	Merced	63,815	53	507	0.188645
San Diego	San Diego	994,677	830	1,208	1.234088
San Joaquin	San Joaquin	181,629	152	583	0.466926
San Luis Obispo	San Luis Obispo	92,739	77	654	0.212528
Santa Barbara	Santa Barbara	136,622	114	823	0.248801
Shasta	Shasta	63,426	53	522	0.182108
Stanislaus	Stanislaus	145,146	121	545	0.399153
Tulare	Tulare	110,385	92	544	0.304118
Rural	All Others	322,792	269	2,460	0.196661
	TOTALS	11,502,870			

3. *Household Size Distribution Weight*

Household travel surveys typically under represent larger households. The total household burden in completing a diary coupled with a much greater effort to interview as many household members as possible is obviously much greater in larger households than smaller. To correct for this under representation, a household size weight was applied to the data set.

The sample data were compared to 2000 Census data for each of the 58 counties and a weight factor was created. The household size distribution comparison between the two data sources were made using categories of 1-person, 2-person, 3-person and 4 or more persons for each of the 58 counties. The final weight factor produces a mean household size of 2.8.

The following page displays the household size distribution weights (1-person, 2-person, 3-person and 4+-person household percentages) by region.

Table 6.5
Household Size Weight Factors

County	Census Total	Sample Total	1-person Census	Census%	1-Person Sample	Sample %	Factor	2-person Census	Census %	2-person Sample	Sample %	Factor	3-person Census	Census %	3-person Sample	Sample %	Factor	4+-person Census	Census %	4+-person Sample	Sample %	Factor
Alameda	523,366	374	136,066	26%	114	30%	0.852982	157,553	30%	161	43%	0.696894	87,461	17%	52	14%	1.222692	142,286	27%	47	13%	2.148511
Alpine	483	23	134	28%	8	35%	0.805000	170	35%	8	35%	1.006250	79	16%	4	17%	0.920000	100	21%	3	13%	1.610000
Amador	12,759	128	3,055	24%	32	25%	0.960000	5,567	44%	60	47%	0.938667	1,745	14%	17	13%	1.054118	2,392	19%	19	15%	1.280000
Butte	79,566	546	21,636	27%	156	29%	0.945000	28,753	36%	254	47%	0.773858	11,958	15%	74	14%	1.106757	17,219	22%	62	11%	1.937419
Calaveras	16,469	167	3,831	23%	56	34%	0.685893	7,073	43%	74	44%	0.970405	2,296	14%	22	13%	1.062727	3,269	20%	15	9%	2.226667
Colusa	6,097	99	1,308	21%	30	30%	0.693000	1,725	28%	34	34%	0.815294	976	16%	13	13%	1.218462	2,088	34%	22	22%	1.530000
Contra Costa	344,129	186	78,759	23%	50	27%	0.855600	110,311	32%	81	44%	0.734815	57,935	17%	27	15%	1.171111	97,124	28%	28	15%	1.860000
Del Norte	9,170	124	2,322	25%	41	33%	0.756098	3,236	35%	51	41%	0.850980	1,385	15%	23	19%	0.808696	2,227	24%	9	7%	3.306667
El Dorado	58,939	103	11,859	20%	26	25%	0.792308	22,387	38%	46	45%	0.850870	9,646	16%	14	14%	1.177143	15,047	26%	17	17%	1.575294
Fresno	252,940	616	52,100	21%	143	23%	0.904615	70,177	28%	239	39%	0.721674	41,391	16%	101	16%	0.975842	89,272	35%	133	22%	1.621053
Glenn	9,172	140	2,022	22%	41	29%	0.751220	2,981	33%	53	38%	0.871698	1,350	15%	23	16%	0.913043	2,819	31%	23	16%	1.886957
Humboldt	51,238	245	14,826	29%	76	31%	0.934868	18,308	36%	92	38%	0.958696	8,088	16%	36	15%	1.088889	10,016	20%	41	17%	1.195122
Imperial	39,384	500	6,724	17%	79	16%	1.075949	9,067	23%	156	31%	0.737179	6,660	17%	96	19%	0.885417	16,933	43%	169	34%	1.272189
Inyo	7,703	215	2,416	31%	62	29%	1.075000	2,861	37%	100	47%	0.795500	945	12%	28	13%	0.921429	1,481	19%	25	12%	1.634000
Kern	208,652	574	42,379	20%	157	27%	0.731210	59,384	28%	219	38%	0.733881	34,284	16%	80	14%	1.148000	72,605	35%	118	21%	1.702542
Kings	34,418	133	5,843	17%	28	21%	0.807500	9,207	27%	55	41%	0.652909	6,111	18%	22	17%	1.088182	13,257	39%	28	21%	1.852500
Lake	23,974	241	6,954	29%	67	28%	1.043134	9,123	38%	108	45%	0.847963	3,213	13%	31	13%	1.010645	4,684	20%	35	15%	1.377143
Lassen	9,625	71	2,354	24%	19	27%	0.896842	3,335	35%	26	37%	0.955769	1,513	16%	14	20%	0.811429	2,423	25%	12	17%	1.479167
Los Angeles	3,133,774	586	771,854	25%	215	37%	0.681395	820,368	26%	201	34%	0.758010	494,369	16%	83	14%	1.129639	1,047,183	33%	87	15%	2.222759
Madera	36,155	259	5,975	17%	47	18%	0.936809	11,441	32%	119	46%	0.696471	5,667	15%	39	15%	1.062564	13,072	36%	54	21%	1.726667
Marin	100,650	125	30,041	30%	39	31%	0.961538	36,452	36%	59	47%	0.762712	15,371	15%	16	13%	1.171875	18,786	19%	11	9%	2.159091
Mariposa	6,613	90	1,755	27%	25	28%	0.972000	2,741	41%	45	50%	0.820000	865	13%	8	9%	1.462500	1,252	19%	12	13%	1.425000
Mendocino	33,266	193	8,983	27%	61	32%	0.854262	11,475	34%	87	45%	0.754253	5,208	16%	20	10%	1.544000	7,600	23%	25	13%	1.775600
Merced	63,815	498	11,318	18%	110	22%	0.814909	16,958	27%	202	41%	0.665644	10,349	16%	88	18%	0.905455	25,190	39%	98	20%	1.981837
Modoc	3,784	61	1,064	28%	19	31%	0.898947	1,468	39%	25	41%	0.951600	494	13%	4	7%	1.982500	758	20%	13	21%	0.938462
Mono	5,137	17	1,366	27%	5	29%	0.918000	1,971	38%	7	41%	0.922857	738	14%	2	12%	1.190000	1,062	21%	3	18%	1.190000
Monterey	121,236	271	25,748	21%	63	23%	0.903333	34,953	29%	108	40%	0.727685	18,612	15%	42	15%	0.967857	41,923	35%	58	21%	1.635345
Napa	45,402	96	11,733	26%	27	28%	0.924444	15,693	35%	42	44%	0.800000	6,696	15%	10	10%	1.440000	11,280	25%	17	18%	1.411765
Nevada	36,894	113	8,429	23%	32	28%	0.812188	15,173	41%	55	49%	0.842364	5,660	15%	17	15%	0.997059	7,632	21%	9	8%	2.636667
Orange	935,287	578	197,650	21%	158	27%	0.768228	278,979	30%	230	40%	0.753913	151,748	16%	80	14%	1.156000	306,910	33%	110	19%	1.734000
Placer	93,382	156	19,860	21%	43	28%	0.761860	33,757	36%	52	33%	1.080000	15,536	17%	29	19%	0.914483	24,229	26%	32	21%	1.267500
Plumas	9,000	73	2,478	28%	25	34%	0.817600	3,853	43%	30	41%	1.046333	1,124	12%	9	12%	0.973333	1,545	17%	9	12%	1.378889
Riverside	506,218	573	104,557	21%	121	21%	0.994463	153,900	30%	215	38%	0.799535	76,334	15%	86	15%	0.999419	171,427	34%	151	26%	1.290199
Sacramento	453,602	500	120,985	27%	135	27%	1.000000	143,307	32%	207	41%	0.772947	73,210	16%	74	15%	1.081081	116,100	26%	84	17%	1.547619
San Benito	15,885	155	2,245	14%	28	18%	0.775000	4,310	27%	64	41%	0.653096	2,750	17%	29	19%	0.908621	6,580	41%	34	22%	1.869118
San Bernardino	528,594	516	97,482	18%	108	21%	0.860000	141,509	27%	181	35%	0.769724	89,839	17%	76	15%	1.154211	199,764	38%	151	29%	1.298543
San Diego	994,677	1,187	240,756	24%	377	32%	0.755650	317,981	32%	490	41%	0.775184	160,306	16%	154	13%	1.233247	275,634	28%	166	14%	2.002169
San Francisco	329,700	162	127,376	39%	72	44%	0.877500	102,564	31%	65	40%	0.772615	41,725	13%	16	10%	1.316250	58,035	18%	9	6%	3.240000
San Joaquin	181,629	577	37,650	21%	151	26%	0.802450	51,672	28%	214	37%	0.754953	29,895	16%	74	13%	1.247568	62,412	34%	138	24%	1.421594
San Luis Obispo	92,739	161	24,081	26%	43	27%	0.973488	34,220	37%	75	47%	0.794267	13,804	15%	19	12%	1.271053	20,634	22%	24	15%	1.475833
San Mateo	254,103	648	62,626	25%	157	24%	1.031847	80,949	32%	277	43%	0.748592	41,161	16%	92	14%	1.126957	69,367	27%	122	19%	1.434098
Santa Barbara	136,622	817	33,210	24%	238	29%	0.823864	43,966	32%	327	40%	0.799511	20,298	15%	112	14%	1.094196	39,148	29%	140	17%	1.692357
Santa Clara	565,863	256	121,109	21%	68	27%	0.790588	171,848	30%	108	42%	0.711111	95,999	17%	41	16%	1.061463	176,907	31%	39	15%	2.034872
Santa Cruz	91,139	441	22,905	25%	125	28%	0.882000	30,166	33%	185	42%	0.786649	14,517	16%	63	14%	1.120000	23,551	26%	68	15%	1.686176
Shasta	63,426	511	15,650	25%	128	25%	0.998047	23,353	37%	219	43%	0.863333	10,002	16%	80	16%	1.022000	14,421	23%	84	16%	1.399167
Sierra	1,520	19	441	29%	4	21%	1.377500	587	39%	8	42%	0.926250	220	14%	3	16%	0.886667	272	18%	4	21%	0.855000
Siskiyou	18,556	182	5,310	29%	48	26%	1.099583	7,206	39%	73	40%	0.972329	2,593	14%	27	15%	0.943704	3,447	19%	34	19%	1.017059
Solano	130,403	151	25,525	20%	41	27%	0.736585	38,952	30%	69	46%	0.656522	23,734	18%	18	12%	1.510000	42,192	32%	23	15%	2.100870

Table 6.5 (Continued)
Household Size Weight Factors

County	Census Total	Sample Total	1-person Census	Census%	1-Person Sample	Sample %	Factor	2-person Census	Census %	2-person Sample	Sample %	Factor	3-person Census	Census %	3-person Sample	Sample %	Factor	4+-person Census	Census %	4+-person Sample	Sample %	Factor
Sonoma	172,403	132	44,340	26%	34	26%	1.009412	58,827	34%	55	42%	0.816000	27,250	16%	25	19%	0.844800	41,986	24%	18	14%	1.760000
Stanislaus	145,146	536	28,211	19%	110	21%	0.925818	41,630	29%	212	40%	0.733208	24,544	17%	85	16%	1.072000	50,761	35%	129	24%	1.454264
Sutter	27,033	92	5,732	21%	19	21%	1.016842	8,529	32%	47	51%	0.626383	4,309	16%	12	13%	1.226667	8,463	31%	14	15%	2.037143
Tehama	21,013	166	5,045	24%	42	25%	0.948571	7,507	36%	66	40%	0.905455	3,146	15%	26	16%	0.957692	5,315	25%	32	19%	1.296875
Trinity	5,587	63	1,648	29%	15	24%	1.218000	2,267	41%	29	46%	0.890690	717	13%	5	8%	1.638000	955	17%	14	22%	0.765000
Tulare	110,385	536	18,913	17%	109	20%	0.835963	29,601	27%	232	43%	0.623793	18,033	16%	80	15%	1.072000	43,838	40%	115	21%	1.864348
Tuolumne	21,004	296	5,453	26%	75	25%	1.026133	8,754	42%	152	51%	0.817895	2,929	14%	30	10%	1.381333	3,868	18%	39	13%	1.366154
Ventura	243,234	631	45,931	19%	163	26%	0.735521	72,814	30%	243	39%	0.779012	41,481	17%	91	14%	1.178791	83,008	34%	134	21%	1.601045
Yolo	59,375	95	13,829	23%	27	28%	0.809259	18,883	32%	38	40%	0.800000	10,184	17%	16	17%	1.009375	16,479	28%	14	15%	1.900000
Yuba	20,535	36	4,456	22%	10	28%	0.792000	6,494	32%	14	39%	0.822857	3,515	17%	1	3%	6.120000	6,070	30%	11	31%	0.981818
	11,502,870	17,040	2,708,308		4,502			3,408,296		6,944			1,841,968		2,459			3,544,298		3,135		

4. Income Distribution

Representativeness of the data are of paramount importance when projecting to the universe. A key demographic variable that affects other variables including trip rates, vehicle ownership, and others is household income. Statewide survey data were compared to the 2000 Census Supplementary Survey and weight factor for income categories were calculated as shown in Table 6.6.

Table 6.6
Income Weight Factors

Income	Census %	Survey %	Factor
Less than $10,000	8.2%	4.9%	1.664952
$10,000 to $24,999	18.5%	16.3%	1.135626
$25,000 to $34,999	11.4%	15.1%	0.755159
$35,000 to $49,999	14.8%	16.0%	0.929707
$50,000 to $74,999	18.8%	22.8%	0.821818
$75,000 to $99,999	11.2%	12.5%	0.897453
$100,000 to $149,999	10.4%	8.5%	1.216445
$150,000 or more	6.8%	3.9%	1.720828

5. Normalization of Weights

If the final weight is calculated using only Factors 1 through 4, the weighted data would represent more or fewer households than the 17,040 households actually contained in the data set. To account for this and still maintain the relative contribution to the data set of each household after weighting, all households are given a normalization factor.

Once each case has received a value for each of the six factors, the final weight can be calculated. The final weight is the product of FACTORS 1-5. The final weight is applied in the analysis of the data through the standard weighting conventions of the statistical analysis software package used.

Data Expansion Factors

The final electronic data set includes three expansion variables that expand the data to the universe (either at the county, regional or statewide level). The expansion factors include:

1. Weekday – (Monday through Friday – one-day households and Friday for the Friday/Saturday households and Monday for the Sunday/Monday households)

2. Weekend – (Saturday or Sunday) – See Data User's Guide on how to select two-day households

3. Seven-Day – (Weekday and Weekend households)

Each expansion factor is calculated by dividing the Census 2000 estimated number of households by the number of households surveyed in each region and multiplying each weight by the Final Weight Variable. The final expansion factors take into account all weight factors including probability of selection, disproportionate sampling by owner/renter status, county distribution and episodic telephone ownership. The final expansion factors also include the Trip Underreporting Correction Factor as discussed in Chapter 7 (GPS). The final weights (including the expansion factors) can be found in the final data files.

An update to the actual Year 2000 should be made when data become available.

The following table summarizes the Weekday expansion factors by region. This expansion factor is applied to households that traveled Monday, Tuesday, Wednesday, Thursday or Friday.

Table 6.7
Weekday Expansion Factor by Region

Region	Year 2000 Households	Sample Size	Weekday Expansion Factor
Western Slope/ Sierra Nevada	56,845	669	85
AMBAG	228,260	816	280
MTC	2,466,019	1,378	1,790
SACOG	712,866	908	785
SCAG	5,386,491	3,354	1,606
Butte	79,566	529	478
Fresno	252,940	575	363
Kern	208,652	529	121
Merced	63,815	481	2,068
San Diego	994,677	1,117	163
San Joaquin	181,629	524	177
San Luis Obispo	92,739	612	223
Santa Barbara	136,622	779	81
Shasta	63,426	499	291
Stanislaus	145,146	498	222
Tulare	110,385	503	642
Rural	322,792	2,309	34
TOTALS	11,502,870	16,080	715

Weekend Expansion Factor

The following table summarizes the Weekend expansion factors by region. Households that were assigned 48-hour travel, had either a Saturday (Friday/Saturday assignment) or Sunday (Sunday/Monday assignment) included. The weekend expansion factor is applied to those trips records that occurred on Saturday or Sunday only. The criteria for selecting these subsets of households is detailed in the Data User's Guide.

Table 6.8
Weekend Expansion Factor by Region

Region	Year 2000 Households	Sample Size	Weekend Expansion Factor
Western Slope/ Sierra Nevada	56,845	12	4,737
AMBAG	228,260	51	4,476
MTC	2,466,019	265	9,306
SACOG	712,866	74	9,633
SCAG	5,386,491	30	179,550
Butte	79,566	128	622
Fresno	252,940	17	14,879
Kern	208,652	41	5,089
Merced	63,815	45	1,418
San Diego	994,677	17	58,510
San Joaquin	181,629	70	2,595
San Luis Obispo	92,739	53	1,750
Santa Barbara	136,622	36	3,795
Shasta	63,426	38	1,669
Stanislaus	145,146	12	12,096
Tulare	110,385	38	2,905
Rural	322,792	33	9,782
TOTALS	11,502,870	960	11,982

Seven-Day Expansion Factor

The following table summarizes the Seven-Day expansion factors by region. This expansion factor is applied to all households.

Table 6.9
Seven-Day Expansion Factor by Region

Region	Year 2000 Households	Sample Size	Seven-Day Expansion Factor
Western Slope/ Sierra Nevada	56,845	681	83
AMBAG	228,260	867	263
MTC	2,466,019	1,643	1,501
SACOG	712,866	982	726
SCAG	5,386,491	3,384	1,592
Butte	79,566	546	146
Fresno	252,940	616	411
Kern	208,652	574	364
Merced	63,815	498	128
San Diego	994,677	1187	838
San Joaquin	181,629	577	315
San Luis Obispo	92,739	648	143
Santa Barbara	136,622	817	167
Shasta	63,426	511	124
Stanislaus	145,146	536	271
Tulare	110,385	536	206
Rural	322,792	2437	132
TOTALS	11,502,870	17,040	675

7. Global Positioning System (GPS) and Adjustment Factors for Trip Underreporting

Background

Trip underreporting has long been a problem in household travel surveys due to the self-reporting nature of traditional survey method, which consists of mail out diaries with computer-assisted telephone interview (CATI) retrieval. Memory decay, failure to understand or to follow survey instructions, unwillingness to report full details of travel, and simple carelessness have all contributed to the incomplete collection of travel data in these self-reporting surveys. To fully understand the nature of these effects and their contribution to underreporting, it is necessary to collect independent data on observed trips.

In the late 1990's, several pilot studies were conducted to investigate the use of Global Positioning System (GPS) technology as a supplement in the collection of personal travel data. These pilot studies confirmed the feasibility of applying GPS technology to improve both the accuracy and completeness of travel data. In 1998, the first real deployment of GPS equipment in a household travel survey occurred in Austin; however, challenges with GPS data accuracy related to the U.S. government's intentional GPS signal degradation (known as Selective Availability) at the time made it difficult to assess the benefit of collecting GPS data concurrently with travel diary data. On May 1, 2000, the Federal Government announced the immediate termination of Selective Availability – which improved, literally overnight, the positional accuracy of raw GPS data from a 50-100 meter range down to 5-10 meters.   This dramatic improvement in GPS positional accuracy made the use of GPS in household travel surveys more desirable, while the continuously declining costs associated with GPS equipment made the application of this technology more feasible.

In February 2001, the GPS component of the California Statewide Household Travel Survey commenced – the first large-scale GPS-enhanced travel survey of its kind.

In addition to this effort, a concurrent GPS study was being conducted in southern California by the Southern California Association of Governments (SCAG). The initial part of this chapter provides a detailed review of the purpose, scope, process, and results of only the Caltrans study since the SCAG GPS component was conducted independent of this study. However, the overall trip under reporting correction factor and its application include the results from both the Caltrans and SCAG GPS efforts. The subsequent analysis of the SCAG GPS data followed the same methodology as for the Caltrans GPS study.

Purpose

In this GPS study, the primary goal of the GPS component is the derivation of trip rate correction factors for under reporting CATI-reported travel data. Households recruited into the GPS-enhanced travel study was provided with both paper diaries and in-vehicle GPS data loggers. The data recorded on the paper diaries was collected by CATI-retrieval methods and then compared with the processed GPS data to identify under-reported trips and other reporting discrepancies.

For the initial analysis, separate geographic regions within the state of California were selected for GPS deployment, with the participating households controlled for demographic. Data were collected during four spring months and then again during one fall month, resulting in a sample size of 517 recruited households.

The set of household characteristics evaluated within this study include the size of household, the number of vehicles, the number of children, the proportion of adults in the household with either a full-time job and/or school attendance, the presence of only one adult in household, and the presence of multi-activity adults. Travel and trip characteristics examined include trip lengths, trip times of day, and trip purposes. Once the trip rate correction factors were developed for this sample, an evaluation of the applicability of these rates to other households with the same demographic characteristics statewide was made. Finally, this approach provided significant insight into the suspected determinants of trip underreporting.

Equipment Description

For this GPS-enhanced travel study, the GeoStats GeoLogger was the GPS data logger used. The GeoLogger is a rugged yet simple GPS data logging device (see Figure 7.1 below) that has been deployed in household travel surveys and travel time studies within the US and in other countries. The GeoLogger is very easy to install – the respondent only needs to plug the power connector into the cigarette lighter socket within the vehicle and to place the combination GPS receiver/antenna on the roof of the vehicle that attaches via a magnetic mount.

Figure 7.1: The GeoStats GeoLogger

This device can log at either one-second or five-second frequencies, can log all valid GPS points or only those valid points for which the speed is greater than 1 MPH (to screen out non-movement events), and is available in 1 MB, 2 MB, and 4 MB versions. For the purpose of this study, the 1 MB units were used, the logging logic was set at one second logging frequencies, and points were not to be logged when speed measured was less than or equal to 1 MPH.

The standard GPS data stream elements recorded by the GeoLogger include date, time, latitude, longitude, speed, heading, altitude, number of satellites, and horizontal dilution of precision (HDOP, a measure of positional accuracy). These elements are stored in the logger in standard NMEA units and are converted into user-specified units and formats upon download.

Geography and Sample Size

The initial deployment goal for the California Statewide Travel Survey GPS Study was 500 households that were to be recruited from within the 16,990 households participating in the statewide household travel survey. Early on, it was realized that 500 deployments would most likely produce between 300 and 400 complete households, factoring in household travel survey dropouts, GPS study dropouts, and vehicle cigarette lighter power failures. Given such a small sample size (as compared to the statewide sample) and to allow for deployment efficiencies, a sampling plan was initially developed for three geographic regions – San Diego, Sacramento, and Alameda counties. An additional five counties in the Southern California Association of Governments (SCAG) region were also surveyed through the SCAG Household Travel Survey.

This focused plan was designed to allow for an in-depth analysis of trip reporting behaviors while controlling for household, person, and travel characteristics evident in the initial three regions.   Weekday samples were the focus of the GPS study; weekend GPS sampling was not scheduled due to the small weekend sample size for the household travel survey.

Figure 7.2: GPS-sampled Counties

Data Collection Methods

In order to perform a trip rate comparison between GPS-measured trips and CATI-reported trips, data were required from both the Caltrans and SCAG region survey GPS sources. After being recruited into the GPS portion of the study, the GeoLoggers were delivered to each participating household with instructions to install the devices in each household vehicle prior to the start of the first trip of the survey date. The installation instructions were very simple – the respondent only needed to plug the power connector into the cigarette lighter and to place the GPS receiver/antenna on the roof of the vehicle.

Each GeoLogger was programmed to store date, time, position, and speed information for each second that the vehicle was in motion. Once the GeoLoggers were retrieved after the study dates, the deployment firm downloaded the GPS data from each device and transferred the GPS files to the GeoStats office in Atlanta. The GPS files were then logged according to the study region and unique household identification number.

Although the recruitment call did ask for confirmation of working cigarette lighters in each household vehicle, it became evident upon retrieval of the equipment and downloading of the data loggers that it was probable that some of the cigarette lighter sockets were not working. However, an 'empty' data logger could also indicate that the respondent chose not to install the GeoLogger into a particular vehicle or that no household members traveled in that vehicle on the travel day. The true reason for total GPS data loss for a particular vehicle would become clearer once the household reported travel data, and data were examined later in the process. Attempts to retrieve the household travel survey information occurred concurrently via CATI methods. Once these data were collected, additional data verification and location geocoding procedures occurred prior to the delivery of the CATI trip files for the GPS households to the GeoStats office. At this point in time, both data streams were ready for processing.

Household Recruitment

In order to maximize the number of GPS recruits, NuStats developed a CATI program tailored specifically for GPS recruitment. The sample frame in this recruitment program was targeted to households living within the initial three GPS sample counties – Alameda, San Diego and Sacramento. Every household with working vehicles recruited in this program, was asked to participate in the GPS study.

The program contained the same questions as the regular diary recruitment program, along with scripting and questions relevant for GPS. These questions ascertained which vehicles were to be used by the household and whether they had working cigarette lighters or 12-volt outlets (needed to power the GeoLoggers). The program also collected information to assist in delivering the GPS units such as a designated contact person, specific delivery address with directions, and suggested times for delivery.

NuStats was provided a production plan that was designed to exactly match the deployment schedule developed by GeoStats. The schedule allowed for eight GPS units to be deployed per county per day. Recruiting efforts had to be monitored meticulously because of the unusual nature of the quotas (household GPS vehicles rather than simply households). Assuming a maximum of three vehicles per household, the eight units were roughly equivalent to three or four households per county or approximately twelve recruits per day. Actual daily recruiting results per travel date varied widely, from as low as six to as many as 18.

The day following recruitment, NuStats e-mailed the pertinent recruitment interview information directly to the deployment team each weekday.

Equipment Deployment

A deployment firm was contracted to deliver and pick up the GeoStats GeoLoggers (the passive in-vehicle GPS data loggers, see Figure 7.1) to and from each household recruited into the GPS study. A GeoLogger was provided for up to three vehicles in each household and was delivered to each household one to two days prior to the assigned travel day. In addition, the date of delivery of the GeoLoggers was also scheduled during the recruitment call.

Once the assigned travel day passed, the equipment was picked up within one or two days; the date and time for pickup was scheduled by the deployment field staff during equipment delivery. Each household was also asked to fill out the standard paper travel diary provided for the household travel survey and to report the household travel information via traditional CATI-retrieval methods.

Nustats transmitted newly recruited participants via a daily email. The emails were broken out by the designated travel dates. Upon receipt, the contents of the emails were logged by the deployment firm and assigned to the appropriate deployment team member based on geography. Although the scheduling of the delivery day was established during recruitment call, the exact time and location of deployment was handled by the local deployment field staff, which allowed them to more efficiently coordinate the new deployments with the pickups already scheduled from previous days' participants.

Although the deployment team initially arranged equipment pickups to occur at the participant's place of work, it became evident that this option was problematic. People either forgot the equipment at home or were not available to transfer the equipment to the person retrieving the GeoLogger. Consequently, it was more time and cost efficient to schedule home pickups, even if it meant waiting two or three days before retrieving the equipment. Once the GeoLoggers were retrieved from the participating households, the field staff would then download and confirm the GPS data, reset the GeoLoggers for a new deployment, and transfer the data files to the GeoStats office in Atlanta.

NuStats Diary Collection

The GPS households were expected to participate in the regular diary portion of the survey exactly as any other Caltrans travel survey participant. The households recruited from the GPS recruitment program were called for diary data retrieval starting the day following their travel day in the regular diary CATI retrieval program.

Data collected during the retrieval telephone call includes all trips made on the designated travel day for all persons in the household. Specific information about each trip is recorded. The information includes the origin and destination addresses, travel duration, travel mode, including specified household vehicle if applicable, and the purpose of each trip.

GeoStats Data Processing

Figure 7.3 depicts the key steps developed to process both the GPS-measured data and the CATI-retrieved data. The GPS second-by-second data, once received, are first converted into GIS-compatible formats and then reviewed for potentially bad or poor data points. Then, a program is run on the GPS data stream to identify potential trip ends based on time intervals between consecutively logged points. For this study, 120 seconds was defined as the appropriate dwell time between GPS-recorded trips. Next, each potential trip is evaluated within an interactive GIS-based application to allow the project analyst to identify both missing and false trip ends. Once this step is complete, the updated GPS-based trip file for a given household vehicle is ready for comparison.

When the electronic CATI-retrieved household trip files arrive at the GeoStats office, the first processing step is to convert the files into a vehicle-based format to provide a standard unit against which GPS trips can be compared. Once this conversion is complete, both the GPS vehicle-based trip files and the CATI-retrieved vehicle-based trip files are ready for comparison. A program is run using the two files as input – comparing the individual trip records within each based on departure times and producing several output files containing the individual trip-level results of the comparison process as well as the aggregate vehicle-level results per household. The comparison results are then examined in tabular form, with questionable items reviewed and modified if necessary. Once the comparisons are finalized, missed trips are flagged and trip rate correction factors are calculated.

Figure 7.3: Data Processing Flow Chart
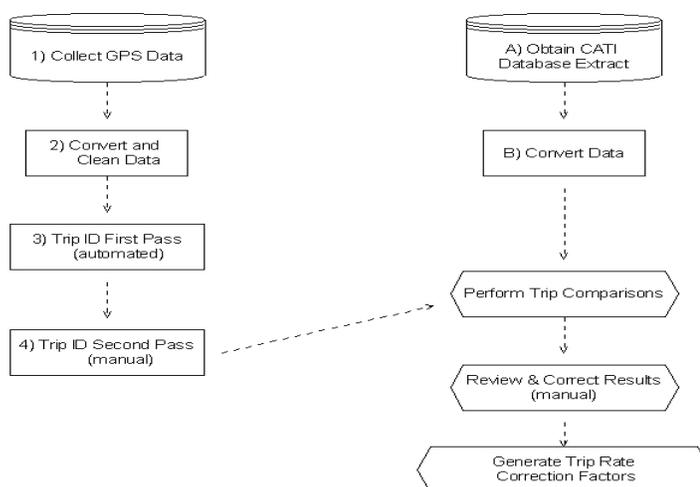
Initial Analysis

The initial (three county analysis) GPS study was conducted over a 20-week period, beginning on February 12, 2001, stopping during the summer months, and finishing on October 3, 2001. Initially, 112 GeoLoggers

were used during the course of this study – none of which were lost or damaged. A breakdown of the intial three county GPS and CATI data collection results can be seen in Table 1.

In the spring phase, 434 initial three county households were recruited and 233 households completed both the CATI and GPS data collection. In the fall, 83 households were recruited and 59 households completed both collection efforts. In total, there are 292 households out of 517 households (57%) that are considered to be complete households and are included in this GPS analysis. Of the 292 complete households, 7 households did not travel at all on their assigned travel day and thus did not collect any GPS data.

Of the 517 households that were originally recruited for the GPS portion of the study, 225 did not complete some or all parts of the study, or the data collected by them included data anomalies that could not be resolved. These households have been categorized as partials (43 households) or refusals (182 households).

A partial household is classified as a household that has one or more vehicles with useable CATI and GPS data, but other vehicles in the household that either refused the GPS unit, forgot to use it, had a broken cigarette lighter, experienced GPS equipment malfunction or traveled with their GPS unit on a day different then their scheduled travel day. Although these data might be used in a trip rate analysis, they have not been included within the scope of analysis covered by this report.

A refusal household is one for which no comparative analysis can be made due to either a lack of useable GPS data or useable CATI data for all vehicles in the household. Of the 182 households in this category:

- 8 used their GPS units on a different day than when they recorded their diary data
- 10 experienced broken cigarette lighters or other equipment malfunctions
- 48 either refused the GPS units upon delivery or forgot to install them
- 85 were never reached for CATI retrieval although GPS data was collected
- 27 refused the GPS equipment and never reported their CATI data
- 4 had data irregularities that could not be resolved

Table 7.1: Breakdown of GPS Households

	Completes	Partials	Refusals	Totals	% Total
Alameda	88	10	50	148	28.6%
Sacramento	93	21	60	174	33.7%
San Diego	111	12	72	195	37.7%
Totals	292	43	182	517	100%
% Total	56.5%	8.3%	35.2%	100%	

Household Characteristics

The recruited GPS households were categorized into four non-mutually exclusive types according to household demographic information. The four types are:

Type 1: Presence of school age children (ages 3-18) in the household

Number of Children	Recruited	Retrieved	Percent Retrieved
0	408	307	75%
1	91	69	76%
2	41	26	63%
3	19	13	68%
4	4	2	50%
Total	563	417	74%

Type 2: Proportion of adults with a full-time job and/or attendance in a post-secondary school

Proportion	Recruited	Retrieved	Percent Retrieved
< 50%	171	124	73%
50 – 99%	141	102	72%
100%	249	191	77%
Total	561	417	74%

Type 3: Presence of multi-activity adults – at least one adult who works and attends school, has two jobs or attends more than one school

Presence of multi-activity adult	Recruited	Retrieved	Percent Retrieved
Yes	93	66	71%
No	472	351	74%
Total	**568**	**417**	**73%**

Type 4: Single adults living alone, with or without children present

Single Adult	Recruited	Retrieved	Percent Retrieved
Under 30	14	12	86%
Over 30	145	117	81%
Total	**159**	**129**	**81%**

Trip Comparisons - Initial Three County Analysis

Once the GPS and CATI data from the 292 initial three county complete households in the California GPS Study were received, a total of 2566 GPS trips were identified based on a minimum 2-minute stop or dwell time; further processing of the GPS trip data within the interactive GIS application revealed another 45 stops with a duration of less than two minutes that occurred off of the vehicle's main travel path. A 2-minute minimum was established based on previous GPS studies that revealed this threshold to be a good starting value. Stops with durations less than two minutes tend to be associated with traffic signals and congestion, and stops with durations greater than two minutes screen out other typical short-duration stops, including drop-offs and pick-ups. The CATI trip file generated a total of 2128 trips for the same households and vehicles.

Next, the CATI-reported trips for each vehicle in each household were compared with the GPS-derived trip data. This comparison was performed automatically via a program designed to compare individual trip records in each vehicle file using only the departure time as the significant variable for matching. A match rate of 58.5% was attained for all GPS measured trips using a plus or minus 12.5-minute departure time buffer as the only match criteria.   Using the CATI reported trips as the baseline, 71.8% of the CATI trips were automatically matched to a GPS trip using this process.

The results of the matching process for each vehicle were then reviewed and analyzed, with matching corrections made based on the type of discrepancy found. A final match rate of 75.9% for all GPS measured trips was achieved upon completion of the review process, along with the identification of nine GPS stops that were not likely to be true stops but rather traffic-related delays. These were removed as part of the GPS trip totals. Using the CATI reported trips as the baseline, 93.1% of the CATI trips were successfully matched to GPS trips upon completion of this review.

Matching results and discrepancies fell into the following categories:

1) Trips detected and reported (i.e., a correct match)
2) Trips detected but reported with poor accuracy in trip start time, resulting in either:
   a) an unmatched trip, or
   b) a mismatched trip
3) Trips detected but not reported (i.e., underreported or missed trips)
4) Trips not detected but reported

Category 1: Exact Matches

For these households, all reported trip start times fell within plus or minus 12.5 minutes of the recorded GPS departure times for each vehicle. Therefore, it is very likely that these respondents were very committed to recording and reporting their trip times very accurately. Further analysis of these matched trips revealed that the geographic locations of the individual trip ends were also sufficiently close and thus confirm the match assignment.

Of all trips made by the 523 vehicles being analyzed as part of the completed households, 123 had perfect matches of GPS-detected and CATI-reported trips without any corrections needed. Once the manual review process was complete, another 63 vehicles were perfectly matched. In addition to these, 104 vehicles had no trips recorded or reported on their travel day, which is also considered an exact match.

Category 2: Trips Detected but Reported with Poor Trip Start Time

In these cases, it seemed apparent that the respondents failed to record, report, or remember one or more trip departure times correctly. As a result, the matching process would either result in an unmatched trip (where the CATI-reported and GPS-measured trip start times were more than 12.5 minutes apart) or in a mismatched trip (were a sequence of CATI and GPS trips were offset significantly, yet the matching algorithm caused a match on incorrect trip pairs). For both types of matching discrepancies caused by poor reporting of trip times, matching corrections were applied based on trip end locations.

Category 3: Trip Detected but Not Reported

As suspected and targeted as the study objective, there were many trips detected by the GPS data logger that were not reported by the respondent. These 'missed trips' were initially tagged as either single links missed within a trip chain, multiple links missed within a trip chain, or as complete round-trips missing all links in tour (which could be two or more).

In this initial analysis, a total of 2128 trips were reported via CATI; after the comparisons were complete, a total of 483 trips were detected within the GPS data that were not reported by respondents.

Category 4: Trip Not Detected but Reported

In a few cases, trips were reported that had no corresponding GPS data. These occurrences fell into three categories:

1) First trips of the day
2) Last trips of the day
3) Trips that are not first or last trips of the day

Of the 292 complete households, 63 households in the initial three county analysis were identified as having a least one missed GPS trip. A total of 147 missing GPS trips were identified and classified as one of three possible missed GPS trip types – initial trip(s) of day, mid-day trips, or end of day trips. It is suspected that a vast majority of start-of-day missed GPS trips are due to a lack of power to the GeoLogger resulting from delayed installation. In addition, it is highly likely that the end-of-day missed GPS trips are the result of GeoLogger removal followed by unplanned, end-of-day errands. To explain missing mid-day GPS trips, further research is underway. These could have resulted from a misreporting of vehicle identification in the CATI retrieval script, a loss of signal in the middle of the day, or if the respondent removed the device for some reason (i.e., to use the cigarette lighter to power a cell phone or to light a cigarette).

Initial Trip Comparison

A summary of the results of the trip comparison and review processes for the 292 complete households, broken down by county, can be seen in Table 7.2. This table shows the number of complete households (# hh) for each county, the number of vehicles instrumented in these households (# veh), the total number of GPS-identified trips after the review process for all instrumented vehicles (# GPS trips); the total number of CATI-retrieved trips associated with all household vehicles (# CATI trips); the number of missed trips detected (# Missed Trips), for which a baseline measure has been calculated as simply the difference between the total number of GPS-detected trips and the total number of CATI-reported trips, and the percentage of missed CATI trips.

Table 7.2: Summary of Trip Comparisons

County	# hh	# veh	# GPS trips	# CATI trips	# missed trips (baseline)	% missed (baseline)	# missed GPS trips	# missed trips (adjusted)	% missed (adjusted)
Alameda	88	152	711	605	106	17.5%	28	134	22.1%
Sacramento	93	171	854	635	219	34.5%	45	264	41.6%
San Diego	111	200	1046	888	158	17.8%	28	186	20.9%
Totals	**292**	**523**	**2611**	**2128**	**483**	**22.7%**	**101**	**584**	**27.4%**

As mentioned earlier in the analysis of trip comparisons, there were also missing GPS trips that clearly should have been logged, but for some reason (such as operator error) were not captured. The number of undetected or missed GPS trips by county (for start-of-day and end-of-day missed GPS trips only, assuming

that the GeoLogger was installed after the first trip or removed prior to the last trip of the travel day) is also shown in Table 7.2, along with the adjusted total missed trips and adjusted percentage of missed CATI trips.

Initially, there were 2,611 GPS trips detected and 2,128 trips reported across the 523 vehicles instrumented in this study. This 483 trip reported shortfall is equivalent to 22.7% of the reported trips. When the missed GPS trip count is included, the shortfall increases to 584 trips or a 27.4% underreporting rate. Finally, for these initial 292 households (Sacramento, Alameda and San Diego County households), the average household trip rate for all vehicle trips made on the assigned travel day and measured by GPS for Sacramento, Alameda, and San Diego counties is 8.9, and increases to 9.3 trips if the missed GPS trips are included; the equivalent CATI-reported trip rate is 7.3.

When comparing the initial results for the three counties, the discrepancy in reporting accuracy is significant between Sacramento (41.6% adjusted underreporting) and the other two counties. It is important to detail the differences in reporting levels across the study based on region, household and trip characteristics. It will also be important to look at the proxy reporting percentages across each of these three counties since proxy reporting often results in a lack of accurately reported data.

Caltrans and SCAG Combined GPS Results

Trip rate correction factors will be developed at several reporting levels, including all household vehicle trips and all household trips. A statistical examination of underreporting trip behaviors will be conducted based on household, travel, and trip characteristics. The set of household characteristics to be evaluated include the size of household, the number of vehicles, the number of children, the proportion of adults in the household with either a full-time job and/or school attendance, the presence of only one adult in household, and the presence of multi-activity adults. Travel and trip characteristics to be examined include trip chaining behaviors, trip lengths, trip times of day, and trip purposes.

Once the trip rate correction factors are developed for this sample, it is expected that the rates can then be applied to other households with the same demographic characteristics statewide. This approach should also provide much insight into the suspected determinants of trip underreporting.

For purposes of this report and the survey data set, an overall weekday driver trip correction factor is being calculated and applied to weekday driver trip rates. The GPS driver trip rates are compared to the 100% sample from the CATI data and not just the GPS households. A more accurate estimate of trip under reporting is determined from all samples since GPS households are aware that their trips are being measured by the GPS units and therefore may tend to be more diligent about providing an accurately completed diary.

The Southern California Association of Governments (SCAG) Household Travel Survey also utilized GPS technology in trip under reporting detection and correction. The results from the both Caltrans and SCAG are used in determining the overall trip under reporting correction factor that is being applied to the Caltrans data set. Households in all SCAG counties, with the exception of Imperial, were included in the SCAG GPS survey. The combined results for all eight counties are shown in the following table. Caution must be taken when analyzing results at the county level due to the limited reliability of small sample sizes such as in Ventura, Orange and Riverside counties. It is the overall correction factor across all counties that is applied to the weekday driver trips.

Table 7.3: GPS Driver Trip Summary by County (Caltrans and SCAG Efforts)

County	Households	GPS Driver Trips	GPS Driver Trips Per Household	CATI Driver Trips Per Household	Trip Under reporting Correction Factor
Alameda	88	711	8.1	5.7	42.1%
Sacramento	93	854	9.2	5.0	84.0%
San Diego	111	1,046	9.4	5.5	31.1%
Los Angeles	51	384	7.0	4.0	75.0%
San Bernardino	158	1,083	6.6	5.0	32.0%
Orange	21	201	9.1	5.5	65.5%
Riverside	24	285	11.9	5.1	133.3%
Ventura	7	32	4.6	5.6	-17.8%
Totals	**565**	**4,596**	**8.2**	**5.0**	**64.7%**

The trip under reporting correction factor of 64.7% is applied to weekday driver trip rates (i.e., vehicular trip rates), including person and total trip rates.

In addition to the GPS-related Driver Trips Per Household results, other useful GPS-related statistics include driver trips per vehicle available and driver trips per vehicle in use. These are summarized in the table below.

Table 7.3b: GPS Driver Trip Summary by County

County	Vehicles Available	Vehicles In Use	GPS Driver Trips	GPS Driver Trips Per Vehicle Available	CATI Driver Trips Per Vehicle Available	GPS Driver Trips Per Vehicle In Use	CATI Driver Trips Per Vehicle In Use
Alameda	166	119	711	4.3	3.2	6.0	4.3
Sacramento	186	132	854	4.6	2.8	6.5	3.5
San Diego	224	170	1,046	4.7	2.7	6.2	4.0
Los Angeles	82	70	384	4.7	2.5	5.5	3.7
San Bernardino	188	178	1,083	5.8	2.7	6.1	3.7
Orange	40	34	201	5.0	2.9	5.9	4.1
Riverside	48	41	285	5.9	2.6	7.0	4.1
Ventura	10	10	32	3.2	2.8	3.2	4.2
Totals	**926**	**754**	**4,596**	**4.9**	**2.7**	**6.2**	**4.0**

8. Survey Results

Background

The chapter contains the summary tables for weighted and expanded data and is based on unlinked trips. The results represent all households in the dataset (i.e., seven-day results, or both weekday only and weekend only traveling households combined, as discussed in the data expansion section of Chapter 6). To run weekday only or weekend only results, refer to Chapter 6 or the Data User's Guide on how to select these subsets of households.

Caution must be taken when comparing results in this report to the 1991 California Statewide Travel Survey results for two reasons:

3) The 1991 Survey results were *not* adjusted for trip under reporting and

4) The 1991 Survey utilized one weight factor – own/rent status by vehicles available while the 2000-2001 Survey results were weighted by four weight factors – own/rent status by vehicle availability, households by county distribution, household size and income.

All trip-level results presented in this section and throughout the main report are based on unlinked trips.

Table 8.1
Persons per Household by Housing Unit Type and Region

Region	Single Housing Unit		Multiple Housing Unit		Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean
Western Slope/Sierra Nevada	49,798	2.5	5,422	1.8	55,220	2.5
AMBAG	156,928	2.9	65,552	2.2	222,480	2.7
MTC	1,281,764	2.9	639,680	2.0	1,921,445	2.6
SACOG	524,425	3.1	192,730	2.0	717,155	2.9
SCAG	3,739,328	3.1	1,900,564	2.7	5,639,892	3.0
Rural	275,568	2.7	47,332	2.0	322,900	2.6
Butte	62,196	2.5	16,681	2.2	78,877	2.5
Fresno	173,617	3.2	84,716	2.4	258,333	2.9
Kern	168,247	3.1	43,707	2.3	211,953	2.9
Merced	53,058	3.5	16,848	2.9	69,906	3.4
San Diego	611,652	2.8	364,133	2.2	975,786	2.6
San Joaquin	129,405	3.1	48,753	2.4	178,159	2.9
San Luis Obispo	286,205	2.8	87,120	2.2	373,325	2.7
Santa Barbara	95,991	3.0	40,219	2.2	136,209	2.8
Shasta	53,499	2.8	11,519	2.4	65,018	2.7
Stanislaus	115,125	3.1	29,349	2.3	144,473	2.9
Tulare	105,112	3.9	26,628	4.1	131,739	4.0
Statewide	7,881,917	3.0	3,620,953	2.4	11,502,870	2.8

Table 8.2
Vehicles per Household by Housing Unit Type and Region

Region	Single Housing Unit		Multiple Housing Units		Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean
Western Slope/Sierra Nevada	49,798	2.3	5,422	1.1	55,220	2.1
AMBAG	156,928	2.2	65,552	1.3	222,480	2.0
MTC	1,281,764	2.3	639,680	1.3	1,921,445	1.9
SACOG	524,425	2.1	192,730	1.1	717,155	1.9
SCAG	3,739,328	2.1	1,900,564	1.3	5,639,892	1.8
Rural	275,568	2.1	47,332	1.1	322,900	1.9
Butte	62,196	2.0	16,681	1.3	78,877	1.9
Fresno	173,617	2.0	84,716	1.1	258,333	1.7
Kern	168,247	1.9	43,707	1.1	211,953	1.8
Merced	53,058	2.2	16,848	1.1	69,906	1.9
San Diego	611,652	2.3	364,133	1.4	975,786	2.0
San Joaquin	129,405	2.2	48,753	1.2	178,159	1.9
San Luis Obispo	286,205	2.2	87,120	1.4	373,325	2.0
Santa Barbara	95,991	2.3	40,219	1.4	136,209	2.0
Shasta	53,499	2.2	11,519	1.3	65,018	2.0
Stanislaus	115,125	2.0	29,349	1.1	144,473	1.8
Tulare	105,112	2.0	26,628	0.7	131,739	1.8
Statewide	7,881,917	2.2	3,620,953	1.3	11,502,870	1.9

Table 8.3
Weekday Total Trips per Household by Housing Unit Type and Region

Region	Single Housing Unit		Multiple Housing Units		Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean
Western Slope/Sierra Nevada	49,798	10.6	5,422	8.0	55,220	10.3
AMBAG	156,928	13.7	65,552	8.0	222,480	12.0
MTC	1,281,764	14.6	639,680	9.3	1,921,445	12.9
SACOG	524,425	13.3	192,730	7.7	717,155	11.8
SCAG	3,739,328	12.8	1,900,564	9.3	5,639,892	11.6
Rural	275,568	12.6	47,332	8.8	322,900	12.1
Butte	62,196	12.7	16,681	9.5	78,877	12.0
Fresno	173,617	11.0	84,716	7.7	258,333	9.9
Kern	168,247	12.0	43,707	7.0	211,953	11.0
Merced	53,058	14.3	16,848	9.3	69,906	13.1
San Diego	611,652	13.2	364,133	8.7	975,786	11.5
San Joaquin	129,405	12.5	48,753	6.9	178,159	10.9
San Luis Obispo	286,205	13.3	87,120	10.0	373,325	12.6
Santa Barbara	95,991	15.5	40,219	8.1	136,209	13.3
Shasta	53,499	11.6	11,519	9.9	65,018	11.3
Stanislaus	115,125	11.0	29,349	7.1	144,473	10.2
Tulare	105,112	16.5	26,628	9.2	131,739	15.0
Statewide	7,881,917	13.2	3,620,953	9.0	11,502,870	11.9

Table 8.4
Weekday Driver Trips per Household by Housing Unit Type, Trip Type, and Region

Region	Home-Other			Other-Other			Work-Other			Home-Work			Home-Shopping			Total		
	Single Housing Unit	Multiple Housing Units	Total Housing Units	Single Housing Unit	Multiple Housing Units	Total Housing Units	Single Housing Unit	Multiple Housing Units	Total Housing Units	Single Housing Unit	Multiple Housing Unit	Total Housing Units	Single Housing Unit	Multiple Housing Unit	Total Housing Units	Single Housing Unit	Multiple Housing Unit	Total Housing Units
Western Slope / Sierra Nevada	3.09	2.12	3.00	2.16	1.57	2.10	0.76	0.16	0.70	1.54	1.10	1.50	0.65	0.35	0.62	8.21	5.32	7.93
AMBAG	3.93	1.94	3.34	2.22	0.87	1.82	1.26	0.42	1.01	2.18	1.34	1.93	0.86	0.40	0.73	10.45	4.97	8.83
MTC	4.36	2.16	3.63	2.55	1.20	2.10	1.28	0.69	1.09	1.99	1.36	1.78	0.93	0.43	0.76	11.11	5.85	9.36
SACOG	3.65	1.96	3.20	1.98	0.97	1.71	1.03	0.64	0.93	1.97	1.36	1.80	0.89	0.61	0.81	9.52	5.54	8.45
SCAG	4.51	2.73	3.91	1.85	0.96	1.55	0.83	0.55	0.74	1.98	1.50	1.82	0.09	0.05	0.08	9.27	5.78	8.09
Rural	3.37	2.18	3.19	2.32	1.30	2.17	1.03	0.74	0.99	1.75	0.92	1.63	0.69	0.61	0.68	9.16	5.74	8.66
Butte	3.48	2.53	3.28	2.56	1.26	2.29	0.96	0.61	0.89	1.56	1.23	1.49	0.88	0.73	0.85	9.45	6.37	8.80
Fresno	2.92	2.17	2.68	1.09	0.70	0.96	0.75	0.55	0.68	2.31	1.16	1.93	0.66	0.28	0.54	7.72	4.87	6.79
Kern	3.36	1.49	2.98	1.60	0.93	1.47	0.90	0.28	0.77	2.28	1.21	2.06	0.56	0.32	0.51	8.72	4.23	7.79
Merced	3.80	2.12	3.40	2.11	0.89	1.82	1.04	0.71	0.96	2.40	1.52	2.18	0.72	0.29	0.61	10.07	5.52	8.98
San Diego	4.31	2.59	3.67	2.41	1.22	1.96	1.20	0.75	1.03	1.79	1.37	1.63	0.77	0.57	0.70	10.48	6.49	8.99
San Joaquin	3.66	1.30	3.01	1.89	0.70	1.57	0.89	0.50	0.78	2.23	1.43	2.01	0.76	0.35	0.65	9.43	4.27	8.02
San Luis Obispo	3.99	3.15	3.79	2.26	1.25	2.02	0.85	0.76	0.83	1.68	1.36	1.61	0.84	0.63	0.79	9.61	7.16	9.04
Santa Barbara	4.79	2.07	3.99	2.27	1.07	1.92	1.10	0.62	0.95	2.02	1.58	1.89	0.89	0.43	0.75	11.06	5.77	9.50
Shasta	3.25	2.84	3.18	2.20	1.58	2.09	0.90	0.59	0.84	1.68	1.73	1.69	0.74	0.69	0.73	8.77	7.43	8.53
Stanislaus	3.05	2.30	2.90	1.41	0.59	1.24	0.84	0.18	0.71	2.34	1.19	2.11	0.47	0.31	0.44	8.10	4.58	7.39
Tulare	3.85	2.01	3.48	2.61	0.85	2.26	0.82	0.35	0.72	1.74	0.65	1.52	1.08	0.21	0.90	10.10	4.07	8.88
Statewide	4.21	2.49	3.67	2.05	1.03	1.73	0.97	0.60	0.85	1.97	1.42	1.80	0.48	0.25	0.41	9.68	5.79	8.46

Table 8.5
Weekday Driver Trips per Vehicle by Housing Unit Type and Region

Region	Single Housing Unit		Multiple Housing Units		Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean
Western Slope/Sierra Nevada	406,131	3.6	28,563	4.7	434,694	3.6
AMBAG	1,626,748	4.7	323,679	3.8	1,950,427	4.5
MTC	14,108,552	4.9	3,697,505	4.7	17,806,057	4.8
SACOG	4,960,615	4.5	1,064,165	4.9	6,024,779	4.5
SCAG	34,405,628	4.4	10,937,394	4.6	45,343,022	4.5
Rural	2,528,491	4.5	271,085	5.4	2,799,576	4.6
Butte	583,264	4.7	105,275	4.8	688,539	4.7
Fresno	1,329,431	3.9	408,767	4.6	1,738,198	4.0
Kern	1,456,364	4.5	183258	3.8	1,639,622	4.4
Merced	530,028	4.6	92,808	5.1	622,836	4.7
San Diego	6,323,446	4.5	2,355,045	4.7	8,678,491	4.5
San Joaquin	1,205,456	4.3	207,671	3.7	1,413,128	4.2
San Luis Obispo	2,728,162	4.5	617,703	5.2	3,345,865	4.6
Santa Barbara	1,052,767	4.8	231,280	4.0	1,284,047	4.7
Shasta	467,757	4.1	84,738	5.6	552,495	4.3
Stanislaus	921,990	4.0	133,138	4.1	1,055,128	4.0
Tulare	1,055,483	5.0	107,727	5.5	1,163,210	5.0
Statewide	75,690,312	4.5	20,849,803	4.6	96,540,116	4.6

Table 8.6
Weekday Person Trips per Household by Housing Unit Type and Vehicle Ownership

Region	0 Vehicles — Single Housing Unit		0 Vehicles — Multiple Housing Units		0 Vehicles — Total Housing Units		1 Vehicle — Single Housing Unit		1 Vehicle — Multiple Housing Units		1 Vehicle — Total Housing Units		2 Vehicles — Single Housing Unit		2 Vehicles — Multiple Housing Units		2 Vehicles — Total Housing Units		3 or More Vehicles — Single Housing Unit		3 or More Vehicles — Multiple Housing Unit		3 or More Vehicles — Total Housing Units		Total — Single Housing Unit		Total — Multiple Housing Unit		Total — Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean
Western Slope / Sierra Nevada	303	0.3	673	0.6	975	0.4	52,434	5.3	27,970	8.8	80,404	6.1	224,073	10.3	5,346	9.8	229,419	10.3	234,723	14.3	4,870	9.7	239,594	14.2	511,534	10.4	38,859	7.2	550,392	10.1
AMBAG	0	0.0	11,133	1.0	11,133	0.8	248,154	7.3	188,658	6.5	436,812	7.0	947,346	14.0	169,987	8.9	1,117,334	12.9	858,801	17.0	47,626	8.5	906,427	16.1	2,054,301	13.2	417,405	6.4	2,471,706	11.2
MTC	177,192	2.8	330,361	2.5	507,553	2.6	1,435,079	6.5	1,744,290	6.4	3,179,369	6.5	7,678,159	14.5	2,163,206	11.9	9,841,365	13.8	8,046,662	17.6	593,693	12.3	8,640,354	17.1	17,337,092	13.7	4,831,550	7.6	22,168,642	11.6
SACOG	30,451	1.6	101,087	2.9	131,538	2.4	1,166,645	9.4	731,365	7.3	1,898,010	8.5	2,888,274	13.1	501,275	9.9	3,389,549	12.5	2,576,312	16.4	74,711	13.1	2,651,024	16.3	6,661,683	12.8	1,408,438	7.4	8,070,121	11.4
SCAG	495,242	2.5	746,205	2.2	1,241,447	2.3	6,856,461	7.8	7,289,327	8.5	14,145,788	8.1	19,620,967	12.9	5,999,459	11.1	25,620,427	12.4	17,805,158	16.2	1,303,429	9.9	19,108,587	15.6	44,777,829	12.1	15,338,420	8.1	60,116,249	10.8
Rural	29,988	2.2	49,742	4.5	79,730	3.2	552,560	7.9	177,328	7.6	729,888	7.8	1,531,003	13.2	130,032	11.4	1,661,035	13.1	1,182,198	16.1	25,520	21.5	1,207,718	16.1	3,295,749	12.1	382,622	8.2	3,678,371	11.5
Butte	4,097	1.3	1,807	0.8	5,904	1.1	142,718	8.2	54,973	6.9	197,691	7.8	332,506	13.3	72,644	15.2	405,150	13.6	265,992	16.7	8,904	6.4	274,895	15.9	745,313	12.1	138,327	8.4	883,640	11.3
Fresno	28,877	3.3	36,430	1.7	65,307	2.1	268,293	6.2	261,760	6.6	530,053	6.4	864,315	11.2	233,824	13.1	1,098,139	11.6	596,818	13.9	57,052	12.5	653,870	13.8	1,758,303	10.2	589,066	7.0	2,347,369	9.2
Kern	19,290	1.8	18,342	2.4	37,632	2.0	396,138	8.4	160,257	6.7	556,395	7.8	865,928	12.9	64,820	6.3	930,748	12.0	643,397	15.5	8,143	6.7	651,540	15.3	1,924,753	11.5	251,562	5.8	2,176,315	10.4
Merced	1,781	0.9	2,764	0.8	4,545	0.8	89,329	8.6	80,703	8.6	170,032	8.6	433,763	16.3	36,451	15.5	470,214	16.3	203,089	15.0	23,750	18.2	226,839	15.3	727,962	13.9	143,669	8.6	871,631	12.6
San Diego	11,300	0.9	89,788	1.6	101,087	1.5	750,986	6.9	1,116,589	7.5	1,867,575	7.3	3,556,881	13.3	1,340,115	10.5	4,896,997	12.4	3,417,424	15.7	389,720	13.0	3,807,144	15.3	7,736,592	12.8	2,936,211	8.1	10,672,803	11.0
San Joaquin	2,996	0.6	29,598	2.9	32,593	2.1	213,795	7.7	117,165	5.1	330,960	6.5	708,056	13.0	111,275	8.9	819,331	12.2	616,433	15.2	31,832	11.1	648,264	14.9	1,541,279	12.0	289,870	6.0	1,831,149	10.4
San Luis Obispo	13,657	1.3	25,518	2.8	39,175	2.0	429,089	6.4	324,988	7.7	754,077	6.9	1,712,859	14.2	394,874	13.7	2,107,733	14.1	1,398,088	16.4	65,582	11.3	1,463,671	16.1	3,553,693	12.5	810,963	9.4	4,364,655	11.8
Santa Barbara	3,967	1.5	10,234	1.9	14,201	1.8	103,849	6.3	117,376	6.2	221,224	6.2	689,817	15.9	111,689	10.0	801,506	14.6	564,172	17.4	49,536	11.8	613,708	16.8	1,361,805	14.3	288,834	7.2	1,650,640	12.2
Shasta	2,415	1.1	5,304	2.5	7,719	1.8	77,778	5.8	47,801	8.9	125,579	6.7	232,312	11.0	37,332	12.7	269,644	11.2	289,347	17.6	15,227	15.6	304,574	17.5	601,851	11.4	105,664	9.3	707,515	11.0
Stanislaus	6,045	0.9	10,537	2.1	16,581	1.4	191,062	6.7	97,441	6.2	288,503	6.6	542,448	11.6	67,262	8.2	609,711	11.1	456,085	14.1	2,884	11.9	458,970	14.1	1,195,640	10.5	178,124	6.1	1,373,765	9.6
Tulare	16,291	3.8	33,106	2.6	49,397	2.9	297,152	10.6	97,065	10.0	394,217	10.4	649,041	16.1	21,768	8.3	670,810	15.6	703,359	22.5	21,902	14.1	725,261	22.1	1,665,842	16.0	173,841	6.6	1,839,684	14.1
Statewide	843,891	2.3	1,502,627	2.2	2,346,518	2.3	13,271,522	7.6	12,635,055	7.7	25,906,577	7.7	43,477,750	13.3	11,461,360	11.1	54,939,110	12.8	39,858,058	16.5	2,724,382	11.0	42,582,440	16.0	98,255,038	12.5	28,557,047	7.9	126,812,085	11.0

Table 8.7
Weekday Driver Trips per Household by Housing Unit Type and Vehicle Ownership

Region	0 Vehicles — Single Housing Unit Count	Mean	Multiple Housing Units Count	Mean	Total Housing Units Count	Mean	1 Vehicle — Single Housing Unit Count	Mean	Multiple Housing Units Count	Mean	Total Housing Units Count	Mean	2 Vehicles — Single Housing Unit Count	Mean	Multiple Housing Units Count	Mean	Total Housing Units Count	Mean	3 or More Vehicles — Single Housing Unit Count	Mean	Multiple Housing Unit Count	Mean	Total Housing Units Count	Mean	Total — Single Housing Unit Count	Mean	Multiple Housing Unit Count	Mean	Total Housing Units Count	Mean
Western Slope / Sierra Nevada	0	0.0	0	0.0	0	0.0	44,794	4.5	20,171	6.3	64,965	5.0	171,423	7.9	4,846	8.9	176,269	7.9	189,914	11.6	3,547	7.1	193,461	11.4	406,131	8.2	28,563	5.3	434,694	7.9
AMBAG	0	0.0	1,459	0.1	1,459	0.1	187,165	5.5	142,215	4.9	329,380	5.2	733,708	10.8	135,936	7.1	869,644	10.0	705,876	13.9	44,069	7.9	749,945	13.3	1,626,748	10.5	323,679	5.0	1,950,427	8.9
MTC	27,419	0.4	69,373	0.5	96,791	0.5	1,164,845	5.3	1,428,839	5.3	2,593,684	5.3	6,060,874	11.4	1,706,781	9.4	7,767,655	10.9	6,855,414	15.0	492,513	10.2	7,347,927	14.6	14,108,552	11.1	3,697,505	5.8	17,806,057	9.4
SACOG	0	0.0	0	0.0	0	0.0	596,011	4.8	575,203	5.8	1,171,214	5.3	2,231,087	10.1	424,991	8.4	2,656,078	9.8	2,133,517	13.6	63,971	11.2	2,197,488	13.5	4,960,615	9.5	1,064,165	5.6	6,024,779	8.5
SCAG	120,098	0.6	17,077	0.0	137,174	0.3	5,147,772	5.8	5,232,804	6.1	10,380,575	6.0	14,621,050	9.6	4,664,543	8.6	19,285,593	9.3	14,516,709	13.2	1,022,971	7.8	15,539,680	12.7	34,405,628	9.3	10,937,394	5.8	45,343,022	8.1
Rural	3,504	0.3	2,620	0.2	6,125	0.2	418,528	6.0	139,154	6.0	557,681	6.0	1,147,239	9.9	106,927	9.4	1,254,166	9.9	959,221	13.0	22,384	18.8	981,605	13.1	2,528,491	9.3	271,085	5.8	2,799,576	8.8
Butte	0	0.0	548	0.2	548	0.1	105,142	6.0	43,467	5.4	148,608	5.8	260,497	10.4	52,357	11.0	312,853	10.5	217,625	13.6	8,904	6.4	226,529	13.1	583,264	9.5	105,275	6.4	688,539	8.8
Fresno	9,317	1.1	0	0.0	9,317	0.3	195,938	4.5	197,762	5.0	393,700	4.7	645,538	8.4	162,771	9.2	808,309	8.5	478,637	11.2	48,234	10.6	526,871	11.1	1,329,431	7.7	408,767	4.9	1,738,198	6.8
Kern	0	0.0	2,831	0.4	2,831	0.2	259,223	5.5	123,325	5.1	382,549	5.4	671,975	10.0	48,958	4.7	720,933	9.3	525,167	12.7	8,143	6.7	533,309	12.5	1,456,364	8.7	183,258	4.2	1,639,622	7.8
Merced	1,179	0.6	0	0.0	1,179	0.2	63,934	6.2	49,151	5.2	113,086	5.7	301,179	11.3	24,944	10.6	326,123	11.3	163,736	12.1	18,713	14.3	182,449	12.3	530,028	10.1	92,808	5.6	622,836	9.0
San Diego	0	0.0	14,736	0.3	14,736	0.2	597,939	5.5	889,324	6.0	1,487,263	5.8	2,853,185	10.7	1,097,223	8.6	3,950,408	10.0	2,872,322	13.2	353,762	11.8	3,226,085	13.0	6,323,446	10.4	360,701	6.5	8,678,491	9.0
San Joaquin	0	0.0	0	0.0	0	0.0	142,968	5.2	101,062	4.4	244,030	4.8	567,092	10.4	79,638	6.4	646,730	9.7	495,396	12.2	26,972	9.4	522,367	12.0	1,205,456	9.4	207,671	4.3	1,413,128	8.0
San Luis Obispo	0	0.0	3,323	0.4	3,323	0.2	364,462	5.5	259,050	6.1	623,512	5.7	1,272,768	10.5	299,183	10.4	1,571,951	10.5	1,090,931	12.8	56,147	9.7	1,147,078	12.6	2,728,162	9.6	617,703	7.2	3,345,865	9.0
Santa Barbara	0	0.0	2,451	0.5	2,451	0.3	86,483	5.2	97,066	5.1	183,549	5.2	514,289	11.8	88,971	7.9	603,260	11.0	451,995	13.9	42,793	10.2	494,787	13.5	1,052,767	11.1	231,280	5.8	1,284,047	9.5
Shasta	0	0.0	0	0.0	0	0.0	64,788	4.9	37,803	7.1	102,591	5.5	174,557	8.3	31,708	10.8	206,266	8.6	228,412	13.9	15,227	15.6	243,639	14.0	467,757	8.8	84,738	7.4	552,495	8.6
Stanislaus	0	0.0	1,944	0.4	1,944	0.2	136,171	4.8	75,490	4.8	211,661	4.8	411,890	8.8	53,305	6.5	465,195	8.5	373,929	11.6	2,399	9.9	376,328	11.6	921,990	8.1	133,138	4.6	1,055,128	7.4
Tulare	3,048	0.7	0	0.0	3,048	0.2	195,936	7.0	71,000	7.3	266,937	7.1	439,630	10.9	17,355	6.6	456,985	10.6	416,868	13.3	19,371	12.5	436,239	13.3	1,055,483	10.1	107,727	4.1	1,163,210	8.9
Statewide	164,565	0.4	116,363	0.2	280,927	0.3	9,772,099	5.6	9,482,885	5.8	19,254,984	5.7	33,077,980	10.1	9,000,436	8.7	42,078,417	9.8	32,675,668	13.5	2,250,120	9.1	34,925,787	13.1	76,236,269	9.7	20,972,655	5.8	97,108,923	8.5

Table 8.8
Weekday Driver Trips per Vehicle by Housing Unit Type and Vehicle Ownership - By Region

Region	1 Vehicle						2 Vehicles						3 or More Vehicles						Total					
	Single Housing Unit		Multiple Housing Units		Total Housing Units		Single Housing Unit		Multiple Housing Units		Total Housing Units		Single Housing Unit		Multiple Housing Units		Total Housing Units		Single Housing Unit		Multiple Housing Units		Total Housing Units	
	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean
Western Slope / Sierra Nevada	44,794	4.5	20,171	6.3	64,965	5.0	171,423	3.9	4,846	4.5	176,269	3.9	189,914	3.2	3,547	1.9	193,461	3.2	406,131	3.6	28,563	4.7	434,694	3.6
AMBAG	187,165	5.5	142,215	4.9	329,380	5.2	733,708	5.4	135,936	3.6	869,644	5.0	705,876	4.0	44,069	2.4	749,945	3.8	1,626,748	4.7	323,679	3.8	1,950,427	4.5
MTC	1,164,845	5.3	1,428,839	5.3	2,593,684	5.3	6,060,874	5.7	1,706,781	4.7	7,767,655	5.5	6,855,414	4.3	492,513	3.1	7,347,927	4.1	14,108,552	4.9	3,697,505	4.7	17,806,057	4.8
SACOG	596,011	4.8	575,203	5.8	1,171,214	5.3	2,231,087	5.1	424,991	4.2	2,656,078	4.9	2,133,517	3.9	63,971	3.6	2,197,488	3.9	4,960,615	4.5	1,064,165	4.9	6,024,779	4.5
SCAG	5,147,772	5.8	5,232,804	6.1	10,380,575	6.0	14,621,050	4.8	4,664,543	4.3	19,285,593	4.7	14,516,709	3.8	1,022,971	2.4	15,539,680	3.6	34,405,628	4.4	10,937,394	4.6	45,343,022	4.5
Rural	418,528	6.0	139,154	6.0	557,681	6.0	11,47,239	5.0	106,927	4.7	1,254,166	4.9	959,221	3.7	22,384	6.0	981,605	3.8	2,528,491	4.5	271,085	5.4	2,799,576	4.6
Butte	105,142	6.0	43,467	5.4	148,608	5.8	260,497	5.2	52,357	5.5	312,853	5.3	217,625	3.9	8,904	2.1	226,529	3.7	583,264	4.7	105,275	4.8	688,539	4.7
Fresno	195,938	4.5	197,762	5.0	393,700	4.7	645,538	4.2	162,771	4.6	808,309	4.3	478,637	3.3	48,234	3.4	526,871	3.3	1,329,431	3.9	408,767	4.6	1,738,198	4.0
Kern	259,223	5.5	123,325	5.1	382,549	5.4	671,975	5.0	48,958	2.4	720,933	4.7	525,167	3.7	8,143	2.2	533,309	3.7	1,456,364	4.5	183258	3.8	1,639,622	4.4
Merced	63,934	6.2	49,151	5.2	113,086	5.7	301,179	5.7	24,944	5.3	326,123	5.6	163,736	3.2	18,713	4.8	182,449	3.3	530,028	4.6	92,808	5.1	622,836	4.7
San Diego	597,939	5.5	889,324	6.0	1,487,263	5.8	2,853,185	5.3	1,097,223	4.3	3,950,408	5.0	2,872,322	3.7	353,762	3.6	3,226,085	3.7	6,323,446	4.5	2,355,045	4.7	8,678,491	4.5
San Joaquin	142,968	5.2	101,062	4.4	244,030	4.8	567,092	5.2	79,638	3.2	646,730	4.8	495,396	3.5	26,972	3.0	522,367	3.5	1,205,456	4.3	207,671	3.7	1,413,128	4.2
San Luis Obispo	364,462	5.5	259,050	6.1	623,512	5.7	1,272,768	5.3	299,183	5.2	1,571,951	5.3	1,090,931	3.6	56,147	2.9	1,147,078	3.5	2,728,162	4.5	617,703	5.2	3,345,865	4.6
Santa Barbara	86,483	5.2	97,066	5.1	183,549	5.2	514,289	5.9	88,971	4.0	603,260	5.5	451,995	4.0	42,793	2.7	494,787	3.8	1,052,767	4.8	231,280	4.0	1,284,047	4.7
Shasta	64,788	4.9	37,803	7.1	102,591	5.5	174,557	4.1	31,708	5.4	206,266	4.3	228,412	3.9	15,227	4.0	243,639	3.9	467,757	4.1	84,738	5.6	552,495	4.3
Stanislaus	136,171	4.8	75,490	4.8	211,661	4.8	411,890	4.4	53,305	3.3	465,195	4.2	373,929	3.5	2,399	3.3	376,328	3.5	921,990	4.0	133,138	4.1	1,055,128	4.0
Tulare	195,936	7.0	71,000	7.3	266,937	7.1	439,630	5.4	17,355	3.3	456,985	5.3	416,868	4.0	19,371	4.2	436,239	4.0	1,055,483	5.0	107,727	5.5	1,163,210	5.0
Statewide	9,772,099	5.6	9,482,885	5.8	19,254,984	5.7	33,077,980	5.1	9,000,436	4.4	42,078,417	4.9	32,675,668	3.9	2,250,120	2.8	34,925,787	3.8	75,690,312	4.5	20,849,803	4.6	96,540,116	4.6

Travel Mode	Western Slope/ Sierra Nevada	AMBAG	MTC	SACOG	SCAG	Rural	Butte	Fresno	Kern	Merced	San Diego	San Joaquin	San Luis Obispo	Santa Barbara	Shasta	Stanislaus	Tulare	Statewide
Drove	265,803	1,193,506	10,862,270	3,749,604	27,611,718	1,705,289	435,743	1,075,396	1,001,379	377,978	5,287,445	867,485	2,048,284	785,047	339,143	646,694	708,311	58,961,095
Percent	77%	73%	73%	72%	70%	72%	70%	69%	71%	69%	78%	73%	72%	72%	76%	73%	59%	71%
Passenger	60,234	283,711	2,149,971	954,534	7,902,030	465,142	142,166	311,484	262,637	120,498	1,095,502	211,452	544,353	196,938	84,141	163,122	374,528	15,322,444
Percent	17%	17%	14%	18%	20%	20%	23%	20%	19%	22%	16%	18%	19%	18%	19%	18%	31%	19%
Local bus	204	8,996	191,420	253,679	352,343	7,220	4,132	20,018	14,353	812	51,007	8,366	14,527	8,449	2,641	2,654	6,649	947,471
Percent	0%	1%	1%	5%	1%	0%	1%	1%	1%	0%	1%	1%	1%	1%	1%	0%	1%	1%
Express bus	0	950	18,563	2,067	16,306	89	313	2,848	220	72	3,155	414	0	1,101	0	0	1,237	47,334
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Light Rail	0	595	68,316	7,763	1,475	197	0	260	0	0	8,980	397	0	0	96	0	0	88,079
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Metro Blue Line	0	0	0	0	11,273	0	0	0	0	0	0	0	0	0	0	0	0	11,273
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Metro Green Line	0	0	2,344	0	6,367	0	1,031	0	0	0	0	0	0	0	0	0	0	9,741
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Metro Red Line	0	753	8,785	743	15,634	0	0	0	0	32	3,111	2,200	0	116	0	0	118	31,492
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
BART	0	0	132,013	271	0	0	0	0	0	561	0	0	0	0	0	136	0	132,981
Percent	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Heavy Rail	0	1,856	43,321	0	21,229	73	131	0	415	0	1,989	727	0	0	0	373	0	70,116
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Dial-a-Ride/Paratransit	0	0	0	673	17,407	3,339	0	0	184	124	3,083	981	478	451	0	0	486	27,206
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
School Bus	10,478	22,358	49,777	62,536	631,078	62,869	2,021	36,933	47,820	29,191	42,788	32,815	61,225	17,734	8,253	27,180	25,765	1,170,820
Percent	3%	1%	0%	1%	2%	3%	0%	2%	3%	5%	1%	3%	2%	2%	2%	3%	2%	1%
Taxi/Shuttle Bus/Limo	0	1,250	34,868	1,512	50,749	1,473	108	0	1,949	0	11,774	883	1,837	1,257	257	797	3,698	112,412
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Motorcycle/Moped	0	321	3,242	1,349	18,796	1,760	0	0	0	0	0	529	405	1,298	362	0	0	28,061
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Bicycle	681	30,373	104,622	30,865	171,335	12,560	10,325	9,777	4,404	1,067	22,273	3,155	30,785	13,563	2,087	3,188	15,101	466,162
Percent	0%	2%	1%	1%	0%	1%	2%	1%	0%	0%	0%	0%	1%	1%	0%	0%	1%	1%
Walk	7,584	77,766	1,230,062	152,698	2,595,705	92,197	25,641	98,990	68,078	18,325	241,313	55,141	130,131	68,894	9,209	46,757	52,593	4,971,084
Percent	2%	5%	8%	3%	7%	4%	4%	6%	5%	3%	4%	5%	5%	6%	2%	5%	4%	6%
Greyhound/Trailways	0	168	11,112	0	19,099	0	1,568	0	0	0	1,260	0	0	122	0	0	1,947	35,276
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Airplane-Commercial	31	516	7,335	1,288	12,583	289	213	324	0	52	5,230	213	4,416	81	83	0	232	32,885
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Airplane-Private	0	0	0	1,685	0	85	0	0	0	0	0	0	0	242	0	0	0	2,012
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Other	306	2,136	16,979	3,043	7288	1,051	0	0	0	531	2,564	843	472	1,443	0	0	517	37,172
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
DK/RF	139	0	8,234	0	95,926	1,787	0	536	966	2,498	1,905	333	1,678	948	428	315	1,030	116,723
Percent	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Total	345,461	1,625,255	14,943,234	5,224,310	39,558,341	2,355,421	623,390	1,556,567	1,402,406	551,740	6,783,379	1,185,933	2,838,593	1,097,682	446,701	891,214	1,192,213	82,621,838

Table 8.10
Weekday Vehicle Occupancy by Region and Statewide
(Persons per vehicle)

Region	All Trips	Home to Work Trips (24 Hour)
Western Slope/Sierra Nevada	1.5	1.1
AMBAG	1.4	1.0
MTC	1.4	1.1
SACOG	1.4	1.0
SCAG	1.5	1.1
Rural	1.5	1.1
Butte	1.7	1.2
Fresno	1.5	1.1
Kern	1.5	1.1
Merced	1.5	1.1
San Diego	1.5	1.0
San Joaquin	1.5	1.1
San Luis Obispo	1.5	1.0
Santa Barbara	1.5	1.1
Shasta	1.4	1.0
Stanislaus	1.4	1.0
Tulare	1.9	1.1
Statewide	1.5	1.1

All Trips	Western Slope/Sierra Nevada	AMBAG	MTC	SACOG	SCAG	Rural	Butte	Fresno	Kern	Merced	San Diego	San Joaquin	San Luis Obispo	Santa Barbara	Shasta	Stanislaus	Tulare	Statewide
0.0 - 4.9	47,046	219,693	2,011,773	679,494	4,603,707	473,154	79,278	164,741	139,900	91,345	779,079	160,422	605,179	162,571	60,712	110,679	129,898	10,518,671
Percent	8%	8%	8%	8%	7%	12%	8%	6%	6%	10%	7%	8%	13%	9%	8%	8%	7%	8%
5.0 - 9.9	112,846	590,229	5,563,954	1,730,840	13,604,364	1,037,240	259,914	560,410	533,344	255,720	2,133,147	406,687	1,252,682	451,971	168,559	315,156	383,993	29,361,054
Percent	20%	22%	23%	20%	21%	27%	25%	22%	23%	28%	19%	21%	27%	25%	23%	21%	20%	22%
10.0 - 14.9	106,970	524,890	4,988,644	1,555,369	11,895,931	736,013	245,903	480,477	429,459	184,537	2,154,513	383,710	919,666	382,638	155,069	271,567	480,556	25,895,911
Percent	19%	20%	20%	18%	18%	19%	24%	19%	20%	19%	20%	20%	20%	21%	21%	19%	24%	19%
15.0 - 15.9	96,866	444,416	3,767,578	1,392,722	11,409,064	563,534	211,593	470,199	465,682	140,691	2,105,266	337,727	590,904	362,419	145,180	242,430	374,485	23,120,756
Percent	17%	17%	15%	16%	18%	15%	21%	18%	20%	15%	19%	17%	13%	20%	20%	17%	19%	17%
20.0 - 24.9	48,615	196,022	1,964,940	1,000,723	5,782,710	228,821	50,101	255,631	201,755	51,168	1,063,178	177,886	315,065	131,183	63,540	111,692	127,675	11,770,704
Percent	9%	7%	8%	12%	9%	6%	5%	10%	9%	6%	10%	9%	7%	7%	9%	8%	7%	9%
25.0 - 29.9	23,285	92,312	1,208,039	390,418	2,517,745	111,390	22,798	143,450	111,754	29,534	576,476	60,779	173,494	47,607	30,208	76,429	52,222	5,667,940
Percent	4%	3%	5%	5%	4%	3%	2%	6%	5%	3%	5%	3%	4%	3%	4%	5%	3%	4%
30.0 - 34.9	46,855	216,341	1,904,768	827,514	6,664,958	292,644	71,444	266,001	212,561	52,524	1,013,320	164,160	357,668	119,774	45,145	166,098	159,791	12,581,565
Percent	8%	8%	8%	10%	10%	8%	7%	10%	9%	6%	9%	8%	8%	7%	6%	11%	8%	9%
35.0 - 39.9	9,646	44,092	465,849	141,056	1,094,312	55,935	18,771	33,569	31,814	13,231	266,033	21,887	76,807	21,071	10,236	23,857	22,516	2,350,681
Percent	2%	2%	2%	2%	2%	1%	2%	1%	1%	1%	2%	1%	2%	1%	1%	2%	1%	2%
40.0 - 44.9	6,716	51,453	347,269	137,832	1,149,667	49,864	9,944	29,109	21,706	8,494	234,350	26,023	65,847	20,058	6,665	26,303	18,953	2,210,252
Percent	1%	2%	1%	2%	2%	1%	1%	1%	1%	1%	2%	1%	1%	1%	1%	2%	1%	2%
45.0 - 49.9	16,598	76,353	688,057	214,508	1,613,054	67,697	14,108	50,176	57,769	20,381	231,720	48,375	91,948	28,499	12,011	32,896	91,055	3,355,205
Percent	3%	3%	3%	2%	2%	2%	1%	2%	3%	2%	2%	2%	2%	2%	2%	2%	5%	2%
50.0 - 54.9	4,781	26,783	178,368	58,484	463,225	29,619	4,269	9,688	8,760	7,089	78,472	14,393	12,840	12,089	7,176	17,956	17,372	951,363
Percent	1%	1%	1%	1%	1%	1%	0%	0%	0%	1%	1%	1%	0%	1%	1%	1%	1%	1%
55.0 - 59.9	2,235	11,036	123,872	22,813	247,665	20,187	1,820	9,919	6,218	1,674	39,443	10,675	20,918	3,959	1,256	2,773	2,222	528,683
Percent	0%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%
60.0 - 64.9	15,501	56,749	471,289	151,583	1,478,391	69,285	8,983	44,574	29,826	21,441	167,024	35,327	48,258	25,262	6,170	18,276	54,261	2,702,199
Percent	3%	2%	2%	2%	2%	2%	1%	2%	1%	2%	1%	2%	1%	1%	1%	1%	3%	2%
65.0 - 69.9	2,369	17,591	77,028	17,014	158,117	12,068	2,721	2,731	3,605	1,168	29,980	7,987	8,189	2,658	2,544	5,116	9,887	360,772
Percent	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%
70.0 - 74.9	1,761	15,503	81,410	27,849	285,505	14,203	2,382	3,379	3,023	2,812	33,156	9,263	4,869	4,417	2,385	2,299	3,969	498,186
Percent	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
75.0 - 79.9	4,808	20,699	154,365	65,160	399,667	17,301	3,907	4,642	6,685	4,451	37,421	14,972	27,734	4,062	3,965	6,923	10,247	787,008
Percent	1%	1%	1%	1%	1%	0%	0%	0%	0%	0%	0%	1%	1%	0%	1%	0%	1%	1%
80.0 +	22,076	72,634	612,815	191,060	1,781,627	99,891	18,402	34,971	44,963	22,320	225,805	71,802	103,095	27,644	14,897	37,380	24,473	3,405,856
Percent	4%	3%	2%	2%	3%	3%	2%	1%	2%	2%	2%	4%	2%	2%	2%	3%	1%	3%
Mean	22.2	20.1	20.1	19.9	20.6	18.6	17.2	18.8	18.7	18.0	19.5	21.2	17.5	16.7	18.1	19.6	18.9	20.0
Median	15	14	13	15	15	10	10	15	15	10	15	15	10	10	13	15	14	15

Home to Work Trips	Western Slope/ Sierra Nevada	AMBAG	MTC	SACOG	SCAG	Rural	Butte	Fresno	Kern	Merced	San Diego	San Joaquin	San Luis Obispo	Santa Barbara	Shasta	Stanislaus	Tulare	Statewide
0.0 - 4.9	3,824	10,209	61,482	44,009	258,574	45,219	3,564	14,151	15,079	11,382	26,612	13,855	21,257	9,337	2,879	9,819	9,851	561,103
Percent	4%	2%	2%	3%	2%	8%	3%	3%	3%	7%	2%	4%	3%	3%	2%	3%	4%	2%
5.0 - 9.9	11,319	67,368	347,616	163,895	1,076,874	114,560	27,409	53,931	65,790	25,627	105,407	41,083	97,252	40,957	14,277	43,561	39,261	2,336,189
Percent	13%	14%	9%	12%	9%	20%	20%	10%	13%	16%	6%	11%	15%	14%	12%	13%	15%	10%
10.0 - 14.9	12,176	64,568	524,695	162,978	1,365,841	103,260	30,399	88,560	78,800	29,296	163,141	50,796	95,768	59,795	19,645	46,157	34,932	2,930,807
Percent	14%	14%	13%	11%	12%	18%	22%	16%	16%	18%	10%	13%	14%	21%	17%	14%	13%	13%
15.0 - 15.9	15,365	69,724	591,268	220,646	1,837,344	100,440	34,196	106,880	112,323	31,199	315,097	61,783	103,004	59,933	35,176	58535	47,490	3,800,403
Percent	18%	15%	15%	16%	16%	18%	25%	19%	23%	19%	18%	16%	15%	21%	30%	18%	18%	16%
20.0 - 24.9	10,067	47,628	426,580	142,381	1,299,666	39,714	5,287	79,696	55,620	14,907	210,253	42,576	67,619	33,594	14,964	31,042	22,370	2,543,963
Percent	11%	10%	11%	10%	11%	7%	4%	14%	11%	9%	12%	11%	10%	12%	13%	10%	9%	11%
25.0 - 29.9	4,745	19,446	262,171	93,667	574,490	19,875	2,853	52,752	19,026	4,517	118,891	16,964	38,219	11,053	7,273	23,820	8,340	1,278,101
Percent	5%	4%	7%	7%	5%	3%	2%	10%	4%	3%	7%	4%	6%	4%	6%	7%	3%	6%
30.0 - 34.9	10,862	70,153	676,210	294,599	2,311,434	71,052	16,086	100,887	78,396	19,093	340,117	63,272	138,069	33,407	11,898	57,231	32,853	4,325,619
Percent	12%	15%	17%	21%	20%	12%	12%	18%	16%	12%	20%	17%	21%	12%	10%	18%	13%	19%
35.0 - 39.9	2,048	9,681	175,929	55,499	342,988	12,515	3,639	9,396	11,536	7,104	79,928	3,876	16,901	5,935	2,728	6,154	2,779	748,636
Percent	2%	2%	4%	4%	3%	2%	3%	2%	2%	4%	5%	1%	3%	2%	2%	2%	1%	3%
40.0 - 44.9	1,523	10,998	99,485	39,291	325,147	6,288	2,211	10,633	4,460	786	75,653	4,996	17,708	5,389	1,158	9,859	3,170	618,756
Percent	2%	2%	3%	3%	3%	1%	2%	2%	1%	0%	4%	1%	3%	2%	1%	3%	1%	3%
45.0 - 49.9	4,671	26,631	243,476	75,948	608,204	16,871	2,650	11,910	27,996	3,539	86,965	16,807	22,592	7,247	1,692	9,362	24,307	1,190,868
Percent	5%	6%	6%	5%	5%	3%	2%	2%	6%	2%	5%	4%	3%	3%	1%	3%	9%	5%
50.0 - 54.9	1,168	8,030	77,802	21,015	176,673	4,428	981	3,662	3,463	2,606	22,816	5,532	2,636	2,198	363	5,344	14,362	353,079
Percent	1%	2%	2%	1%	2%	1%	1%	1%	1%	2%	1%	1%	0%	1%	0%	2%	5%	2%
55.0 - 59.9	716	6,233	45,450	4,901	105,241	3,457	0	1,353	1,738	351	17,452	4,485	6,039	1,033	97	183	791	199,520
Percent	1%	1%	1%	0%	1%	1%	0%	0%	0%	0%	1%	1%	1%	0%	0%	0%	0%	1%
60.0 - 64.9	4,721	18,562	166,970	59,329	577,401	15,342	2,866	11,099	7,959	2,995	62,214	12,714	20,126	10,134	945	10,223	15,608	999,211
Percent	5%	4%	4%	4%	5%	3%	2%	2%	2%	2%	4%	3%	3%	4%	1%	3%	6%	4%
65.0 - 69.9	357	4,974	20,271	3,871	62,236	1,570	595	1,276	681	176	10,966	2,521	1,638	561	0	476	1,548	113,717
Percent	0%	1%	1%	0%	1%	0%	0%	0%	0%	0%	1%	1%	0%	0%	0%	0%	1%	0%
70.0 - 74.9	210	7,761	22,040	2,697	66,165	2,487	166	1,391	1,761	197	8,432	1,037	610	667	219	910	518	117,267
Percent	0%	2%	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%
75.0 - 79.9	900	5,709	63,118	16,583	155,850	4,273	0	304	1,377	1,447	7,442	6,367	9,941	1,805	943	2,519	2,191	280,769
Percent	1%	1%	2%	1%	1%	1%	0%	0%	0%	1%	0%	2%	1%	1%	1%	1%	1%	1%
80.0 +	2,971	20,135	140,691	19,169	448,794	8,785	2,510	6,721	3,645	8,683	52,988	30,393	9,307	1,932	1,627	7,804	2,252	768,406
Percent	3%	4%	4%	1%	4%	2%	2%	1%	1%	5%	3%	8%	1%	1%	1%	2%	1%	3%
Mean	25.9	27.7	28.8	25.4	28.9	20.2	19.6	21.9	21.3	23.3	28.1	30.7	23.9	20.3	20.3	24.5	26.9	27.6
Median	20	20	25	20	24	15	15	20	15	15	25	20	20	15	15	20	20	20

Region	Drive Alone	Carpool / Vanpool	Public Transit	Other	Total
Western Slope/Sierra Nevada	50,430	2,370	0	2,210	55,010
Percent	92%	4%	%	4%	100%
AMBAG	263,420	20,830	1,580	11,330	297,170
Percent	89%	7%	1%	4%	100%
MTC	2,083,890	152,400	168,050	191,270	2,595,610
Percent	80%	6%	6%	7%	100%
SACOG	809,970	61,690	13,140	11,620	896,430
Percent	90%	7%	1%	1%	100%
SCAG	6,214,920	835,600	152,200	367,310	7,570,030
Percent	82%	11%	2%	5%	100%
Rural	319,360	21,420	1,430	22,600	364,810
Percent	88%	6%	0%	6%	100%
Butte	73,620	6,110	1,200	6,850	87,780
Percent	84%	7%	1%	8%	100%
Fresno	305,740	40,670	0	10,390	356,790
Percent	86%	11%	0%	3%	100%
Kern	268,710	41,670	360	5,050	315,800
Percent	85%	13%	%	2%	100%
Merced	92,410	9,150	1,290	1,260	104,110
Percent	89%	9%	1%	1%	100%
San Diego	963,940	61,670	29,150	25,930	1,080,700
Percent	89%	6%	3%	2%	100%
San Joaquin	218,900	11,910	810	6,410	238,030
Percent	92%	5%	%	3%	100%
San Luis Obispo	364,990	42,670	0	24,870	432,540
Percent	84%	10%	0%	6%	100%
Santa Barbara	155,540	8,760	1,830	17,960	184,090
Percent	84%	5%	1%	10%	100%
Shasta	67,030	2,490	0	2,750	72,280
Percent	93%	3%	0%	4%	100%
Stanislaus	183,760	13,660	0	6,690	204,110
Percent	90%	7%	0%	3%	100%
Tulare	121,400	52,390	780	9,510	184,080
Percent	66%	28%	0%	5%	100%
Statewide	12,558,050	1,385,490	371,810	724,010	15,039,370
Percent	84%	9%	2%	5%	100%

Table 8.13
Survey Household Income Distribution by Region and Statewide

Region	Income Strata									Median Income
	<$10,000	$10,000-$24,999	$25,000-$34,999	$35,000-$49,999	$50,000-$74,999	$75,000-$99,999	$100,000-$149,999	>$150,000	DK/RF	
Western Slope/Sierra Nevada	4,290	9,920	5,990	7,900	10,880	4,150	3,830	1,790	6,270	$53,494
Percent	8%	18%	11%	14%	20%	8%	7%	3%	11%	
AMBAG	15,320	26,980	20,170	26,810	36,210	25,090	20,320	16,990	34,800	$56,434
Percent	7%	12%	9%	12%	16%	11%	9%	8%	16%	
MTC	86,900	200,870	127,410	182,880	305,170	255,360	274,250	251,110	238,210	$57,013
Percent	5%	11%	7%	10%	16%	13%	14%	13%	12%	
SACOG	51,420	134,160	77,090	80,990	136,480	69,510	66,200	36,160	71,580	$43,533
Percent	7%	19%	11%	11%	19%	10%	9%	5%	10%	
SCAG	580,640	1,081,720	558,220	589,730	838,360	468,810	523,900	350,200	641,140	$31,242
Percent	10%	19%	10%	11%	15%	8%	9%	6%	11%	
Rural	38,070	82,920	39,840	44,840	44,820	19,690	12,630	5,130	34,870	$21,839
Percent	12%	26%	12%	14%	14%	6%	4%	2%	11%	
Butte	9,870	20,950	10,970	12,450	11,090	5,750	3,360	160	6,340	$39,768
Percent	12%	26%	14%	15%	14%	7%	4%	0%	8%	
Fresno	39,970	57,300	26,440	36,310	38,700	15,670	10,210	5,520	30,180	$26,659
Percent	15%	22%	10%	14%	15%	6%	4%	2%	12%	
Kern	22,690	50,630	24,440	24,460	30,960	16,790	12,330	6,420	24,180	$30,132
Percent	11%	24%	12%	12%	15%	8%	6%	3%	11%	
Merced	8,230	17,440	7,710	10,790	10,420	3,780	2,190	2,500	6,440	$29,005
Percent	12%	25%	11%	16%	15%	5%	3%	4%	9%	
San Diego	41,750	125,130	95,520	119,680	155,030	125,420	105,080	58,770	144,600	$60,032
Percent	4%	13%	10%	12%	16%	13%	11%	6%	15%	
San Joaquin	14,350	27,670	19,310	24,960	29,050	19,090	14,430	9,500	21,210	$37,776
Percent	8%	15%	11%	14%	16%	11%	8%	5%	12%	
San Luis Obispo	28,250	61,460	46,710	68,860	58,070	31,040	24,140	13,970	39,650	$50,166
Percent	8%	17%	13%	19%	16%	8%	7%	4%	11%	
Santa Barbara	9,640	18,860	13,890	14,360	20,840	18,350	11,250	10,530	18,450	$55,040
Percent	7%	14%	10%	11%	15%	14%	8%	8%	14%	
Shasta	6,560	17,720	8,160	10,770	9,040	4,490	2,730	1,260	4,600	$28,256
Percent	10%	27%	13%	17%	14%	7%	4%	2%	7%	
Stanislaus	13,200	29,280	18,450	22,160	24,620	13,590	9,850	3,130	10,580	$44,813
Percent	9%	20%	13%	15%	17%	9%	7%	2%	7%	
Tulare	13,550	40,830	12,710	19,560	14,370	6,690	6,300	2,640	14,570	$32,765
Percent	10%	31%	10%	15%	11%	5%	5%	2%	11%	
Statewide	984,710	2,003,840	1,113,010	1,297,490	1,774,100	1,103,270	1,103,020	775,770	1,347,670	$54,946
Percent	9%	17%	10%	11%	15%	10%	10%	7%	12%	

Table 8.14
Key Weekend Statistics[9]

Travel Data (Unlinked Trips)	Average Weekend
Total Household Trips[10]	40,489,452
Household Person Trips[11]	38,408,151
Household Driver Trips[12]	25,919,195
Mean Trips Per Household	3.5
Person Trips Per Household	3.3
Person Trips Per Person (All ages)	1.1
Person Trips Per Person Five+ Years of Age[13]	1.2
Driver Trips Per Household	2.3
Driver Trips Per Vehicle Available	1.2
Driver Trips Per Vehicle In Use on Travel Day	1.9
Vehicle Occupancy	
All Trips (24 hours)	1.7
All Trips (7-9 a.m.)	1.5
Home to Work Trips (24 hours)	1.1
Home to Work Trips (7-9 a.m.)	1.1
Mean Travel Time (Respondent Reported)	
Trip Length (All Trips)	20.7
Weekday Trip Length (Home to Work Trips)	22.2
Travel Mode Distribution	**Weekend Percent**
Vehicle Driver Trips	64.0%
Vehicle Passenger Trips	29.5%
Public Transportation Trips	1.1%
Bicycle Trips	1.0%
School Bus Trips	0.2%
Walk Trips	4.0%
All Others	0.2%

---

[9] Weekend trips are not adjusted for trip under reporting.
[10] Total trips include all household trips by all modes of travel.
[11] Person trips include all household trips except walk, bicycle, airplane, and "other" mode trips.
[12] Driver trips include household automobile, pickup, RV, van, motorcycle, and truck driver trips.
[13] Only includes trips made by persons five years of age or older.

9. Survey Quality Assessment

Background

In addition to the Quality Control procedures as discussed in Chapter 5, data were also compared to available 2000 Census data and the 1995 Nationwide Personal Transportation Survey (NPTS now known as the National Household Travel Survey). The purpose for comparing key survey results with other independently collected data is to validate the results – Are the survey results within a reasonable estimate? and to use available data for data weighting and expansion as discussed in Chapter 6.

Although the most useful comparisons are to either California statewide or county level data, comparing results to national data are also useful. For example, knowing that the mean household size in California is larger than the national average, by comparing the survey results with national results is a good validation tool. In other words, the actual numbers are not as important as much as knowing that the survey results must be higher.

Several key survey statistics (weighted and expanded) were compared to both 2000 Census Supplementary Survey data and the 1995 NPTS data for which there was a direct comparison. Some variables were directly comparable while others required recoding for comparison. Household and person data were compared. ($\sqrt{}$ indicates comparability)

Table 9.1
Variable Comparability

Household File Comparison

Travel Survey Variable	2000CSS	1995 NPTS (Nation-wide)
Household Size	$\sqrt{}$	$\sqrt{}$
Vehicles	$\sqrt{}$	$\sqrt{}$
Dwelling type	$\sqrt{}$	Not Available
Owner Status	$\sqrt{}$	Not Available
Income	$\sqrt{}^{14}$	$\sqrt{}$
Language	$\sqrt{}^{9}$	Not Available

Person File Comparison

Travel Survey Variable	2000CSS	1995 NPTS
Gender	$\sqrt{}$	$\sqrt{}$
Age	$\sqrt{}^{9}$	Not Available
Relationship	$\sqrt{}^{9}$	Not Available
Disability	$\sqrt{}^{9}$	Not Available
Student Status	$\sqrt{}$	Not Available
Level Attending	$\sqrt{}^{9}$	Not Available
Employment Status	$\sqrt{}^{9}$	Not Available
Employer Type	$\sqrt{}^{9}$	Not Available
Industry	$\sqrt{}$	Not Available
Occupation	$\sqrt{}$	Not Available
Mode to Work	$\sqrt{}$	Not Available
Ethnicity	$\sqrt{}^{9}$	Not Available

$\sqrt{}$ = comparable
[14] Recategorizing of data required to allow comparison

The following are the comparisons for each variable listed above.

Table 9.2
Data Comparison

Travel Survey Variable	Survey	2000CSS	1995 NPTS (Nationwide)
Household Size	2.83	2.87	2.63
Vehicles	1.86	Not Available	1.78
Gender:			
Male	48.9%	49.8%	48.7%
Female	50.7%	50.2%	51.3%
Dwelling type:			
Single	67.9%	67.5%	
Multiple	31.2%	32.4%	
Owner Status:			
Own	54.8%	56.9%	
Rent	44.2%	43.1%	
Income:			
Less than $10,000	8.6%	8.2%	
$10,000 to $24,999	17.4%	18.4%	
$25,000 to $34,999	9.7%	11.4%	
$35,000 to $49,999	11.3%	14.8%	
$50,000 to $74,999	15.4%	17.7%	
$75,000 to $99,999	9.6%	11.2%	
$100,000 to $149,999	9.6	10.3%	
$150,000 or more	6.7%	6.6%	
Age (median)	33.0	33.3	
Relationship:			
Respondent	35.4%	39.1%	
Spouse	20.7%	19.8%	
Child	30.4%	30.0%	
Other relatives	6.1%	6.0%	
Nonrelatives	2.4%	5.0%	
Other	0.3%	2.0%	
Disability	6.3%	15.7%	
Industry:			
Agriculture, forestry, fishing and hunting, and mining	3.8%	2.1%	
Construction	6.1%	6.3%	
Manufacturing	5.1%	13.0%	
Wholesale trade	0.9%	4.1%	
Retail trade	8.0	11.2%	
Transportation and warehousing, and utilities	4.9%	4.6%	
Information	2.5%	3.9%	
Finance, insurance, real estate and rental and leasing	5.7%	6.9%	
Professional, scientific, management, administrative and waste management services	16.0%	11.6%	
Educational, health and social services	21.5%	17.5%	
Arts, entertainment, recreation, accommodation and food services	7.8%	8.7%	
Other services (except public administration)	1.0%	5.2%	
Public Administration	6.3%	4.5%	

Component and Overall Response Rates

Another key element in assessing the quality of the data is the response rate. The response rate is the percentage of the total contacts completes the survey (provide travel data). Overall response rate is one guide to the representation of the sample respondents. The sampling plan is but a means to an end, because it is the response of the actual sample that matters. The responses of those who completed the survey comprise the data set, and an acceptable response rate is critical. Low response rates can introduce response bias – particular types of households that didn't participate may not be adequately represented in the survey.

The 2000-2001 California Statewide Household Travel Survey used a multistage survey process (i.e., household recruitment and household retrieval). In this case, a rate is calculated for each survey stage — called the component response rate, then the overall response rate is determined by multiplying the rates together.

Recruitment Component Response Rate

From the sampling frame, NuStats randomly selected over 200,000 telephone numbers for inclusion in the study. This is the basis from which households were identified and recruited for inclusion in the study.

As shown in Table 9.3, a total of 73,201 telephone numbers was found to be ineligible for the study (disconnects, business/government, and computer/fax). Dividing the sum of the total eligible and ineligible units (73,201+52,694=125,895) by the total eligible units (52,694) is the eligibility rate for the recruitment phase. The rate is 42%. This rate was used to determine the number of eligibility unknown numbers to allocate to the response rate calculation (42% of the total eligibility unknown of 76,731 – no answer, busy, answering machine, call backs – is 32,227).

The recruitment response rate is calculated using the following formula:

Recruitment Response Rate = Recruits / (Recruits+Refusals+42% of Eligibility Unknown Units)
Recruitment Response Rate = 23,666 / 23,666+29,028+32,227 = 23,666 / 84,921
Recruitment Response Rate = 28%

Table 9.3
Recruitment Call Outcomes

Call Outcome	Frequency
ELIGIBLE UNITS	
Recruited	23,666
Refused to participate	29,028
SUB-TOTAL ELIGIBLE	52,694
INELIGIBLE UNITS	
Disconnected/non-working	44,864
Business/Government	11,781
Facsimile	9,005
Over Quota	7,551
SUB-TOTAL INELIGIBLE UNITS	73,201
ELIGIBILITY UNKNOWN UNITS	
No answer	22,026
Call Back	36,551
Answering machine	15,990
Busy	2,164
SUB-TOTAL ELIGIBILITY UNKNOWN UNITS	76,731
GRAND TOTAL:	*202,626*

Retrieval Component Response Rate

The retrieval rate is then calculated using the same formula as the recruitment survey component rate. Of the total 23,666 recruited households, virtually all are eligible since the vast majority had been contacted previously (the only ineligible households are those in which the phone was not in service at the time contact was attempted. The retrieval component response rate is therefore the number of completed surveys divided by the total sample (i.e., all recruits) or 72%.

Table 9.4
Retrieval Call Outcomes

Call Outcome	Frequency
ELIGIBLE UNITS	
Completed	17,040
Refused to participate	1,257
Pending (no answer, call backs, answering machines)	5,166
SUB-TOTAL ELIGIBLE	23,463
INELIGIBLE UNITS	
Disconnected/non-working	203
SUB-TOTAL INELIGIBLE UNITS	203
GRAND TOTAL:	*23,666*

Overall Response Rate

In addition to the component rates, an overall response rate is calculated. The overall response rate can be computed using the following formula:

$$RR = \left(\frac{a_1}{A_1 + (C_1 * ER_1)} \right) * \left(\frac{a_2}{A_2 + (C_2 * ER_2)} \right)$$

Where,

RR is the Overall Response Rate,
a_1 and a_2 and a_3 are the number of completed surveys for each of the two phases,
A_1 and A_2 are the number of eligible telephone numbers for each of the two phases,
C_1 and C_2 are the number of eligibility unknown for each of the two phases (note that in the retrieval phase all households are determined eligible and known since each was already recruited), and
ER_1 and ER_2 are the eligibility rates for each of the three phases.

Using this formula, the Overall Response Rate is 20% (0.28 * 0.72). The response rate calculation uses the same formula prescribed by the Council of American Survey Research Organizations (CASRO).

In previous household travel surveys conducted across the US, the response rates varied between 20 and 24 percent using the same calculation method (accounting for a portion of the eligibility unknown in the denominator of the formula). For this survey, the response rate is about average.

10. Survey Limitations

As with any survey dataset, there are limitations on its use. End users must understand the methods used to collect the data, the accuracy when analyzing at the subsample level, and various biases. Each of these affect the reliability level when generalizing to the universe.

To ensure adequate representation, NuStats followed prescribed industry procedures during data collection. Each household was given an equal chance of being selected through sufficient variation in call attempts. NuStats interviewers called at various times of the day and various days of the week during the recruitment of households. This procedure is used to ensure that all types of households are contacted from those that are active and spend a large amount time away from home (multiple jobs or attend work and school or large families with school-age children who participate in extra curricular activities) to retired households in which members do not travel often. Despite strict adherence to call back attempts and minimization of refusals, there may be specific types of households that are under represented in the survey ("hard refusals"). Although not part of the scope of this survey, it might prove useful to conduct analysis to determine what type of households might be under represented.

Another limitation in using the data is the issue of trip under reporting. Although the GPS task addresses this issue (Chapter 7), the trip under-reporting only applies to vehicular trips. The theory is that trips of short duration is the type of trip typically under-reported, but other modes of travel in addition to vehicular modes can be used in making short trips. Other modes include walk and bicycle. Until more advances can be made in person-based GPS tracking systems or other advances in calculating a more accurate estimate particularly in non-vehicular trips, trip under reporting will continue to be underestimated by the very nature of the use of a self-reported survey.

As with any survey, data users must always be aware of the accuracy level particularly when analyzing small sample data (or subsets of the full dataset). Although large samples were collected for each region (minimum 500 producing a standard error of +/- 4.5 percentage points at the 95 percent confidence level), caution must be taken when attempting to analyze at smaller geographies. In addition, the final data weights were developed at the region level and not at the county level (except for those counties that are in and of themselves a region).

As discussed in Chapter 3, the survey produced a sample size of 17,040 randomly selected households with an overall reliability of $\pm$ 0.8 percentage points at the 95% confidence level with respect to household level attributes at the Statewide level of analyses. Table 11.2 summarizes the standard error for various sample sizes as in Chapter 3.

Table 10.1
Standard Error Rates at the 95% Confidence Level

Sample Size	Standard Error
500	$\pm$ 4.5
750	$\pm$ 3.7
1,000	$\pm$ 3.2
2,500	$\pm$ 2.0
3,000	$\pm$ 1.8
17,000	$\pm$ 0.8

11. Recommendations for Future Survey Improvements

Background

The 2000-2001 California Statewide Household Travel Survey is one of the largest in the U.S. to ever be undertaken (over 17,000 households). In 1995, the New York Metropolitan Transportation Commission (NYMTC) conducted a survey of over 12,000 households in the New York and New Jersey region. Currently, the Southern California Association of Governments is completing a survey of over 20,000 households. It is recommended that future large scale surveys allocate funds for the use of multimodal methods of collecting data and for more comprehensive analysis. These are discussed below.

Data Collection Techniques

Respondent participation continues to decline over the years due to time required to participate, the personal nature and level of detail of the questions being asked, and the distrust of the caller or the sponsor of the survey. Opting out of participating has also become easier for respondents. The use of screening tools are increasingly being used such as answering machines, Caller I.D. and the recent introduction of a "phone zapper" – an inexpensive machine that sends a message back to the caller's database and automatically purges the respondents phone number. In the future it will become much more difficult to obtain a representative sample.

To address the issue of respondents being too busy to participate, the next survey should maximize the various opportunities for respondents to provide data. One method is the use of web-based data collection. This will allow respondents to complete the survey at their convenience.

To inform potential respondents, the next survey should fully utilize the District Offices around the State in publicizing the survey. Each District Office representative should be fully briefed in person and provided a media kit that can be distributed to local media outlets, law enforcement and elected officials. Although a packet of information was sent to each District Office prior to the start of the survey, a more formal in-person presentation should be made.

By the time the next California Statewide Household Travel Survey is conducted, significant advances in Global Positioning System (GPS) technology will be achieved. The development of highly efficient and effective person-based GPS systems is anticipated and if available should be utilized in the next survey to accurately collect sufficient samples for estimating trip under reporting for all modes of travel in addition to vehicular modes and trip purpose. If the person-based GPS system is utilized, it may be used to collect data in place of the diary method and will require less time of the repondent to participate.

If a person-based GPS system is not fully developed before the next survey, other techniques in using GPS should be implemented such as a retrospective GPS component. The technique follows the basic procedures in collecting GPS data but also involves recontacting respondents to verify data and to "fill in" unreported trips as identified between the GPS and diary collected data. More detail such as trip purpose, activity, occupancy, and others for the missed trips can be captured using this method.

Data Analysis

Because of declining respondent participation, it is important to adequately explore the extent non-response bias affects results. To date, research has been conducted to identify the types of households likely to not participate in travel surveys (large households, zero-car, high income) but the application of a adjustment factor has not been developed and applied to datasets. Non-participating households (both during recruitment and retrival) should be contacted for follow-up.