Transportation Agency of Monterey County

2002
TAMC
Household Travel Survey
Final Report
(July 2002)

NuStats

Table of Contents

1. Survey Objectives and Overall Approach

This report documents the design, implementation and results of the Household Travel Survey conducted on behalf of the Transportation Agency of Monterey County (TAMC). The survey served to collect additional samples in Monterey and San Benito Counties in addition to the surveys collected through the 2000-2001 California Statewide Household Travel Survey, sponsored by Caltrans. The TAMC survey was conducted between March 2002 through May 2002.

The household socioeconomic and travel data collected will be used to help refine travel estimates, models, and forecasts throughout the region. The TAMC data set (Monterey and San Benito counties) was combined with the surveys from collected in Monterey, San Benito, Santa Clara and Santa Cruz from the Statewide Survey providing a more robust sample.

Two survey instruments for each the household recruitment and data retrieval stages were used (see Data User's Guide).

The TAMC Household Travel Survey, like the Caltrans Household Travel Survey and all other recent household travel surveys, relied on the willingness of area residents to complete diary records of their daily travel over a 24-hour (weekday) period. 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 day. Beginning the day after the assigned date, attempts to contact households were made to retrieve the diary information.

Specifically for the TAMC survey, a total of 1,000 households were recruited to participate in the study. Of these, 693 households (69%) completed travel diaries, and the information was retrieved from all household members regardless of age. The 693 households represent 1,701 persons, 1,470 vehicles and 5,785 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 distribution, obtained from the 2000 Census data files as provided by TAMC. 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.

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 section by major topic. The sections include:

1. Survey Objectives and Overall Approach
2. Survey Instruments and Materials
3. Sampling Design and Procedures
4. Interviewer Training, Survey Methods, and Quality Control Procedures
5. Survey Data Weighting and Expansion
6. Survey Results
7. Survey Quality Assessment

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 TAMC survey materials were identical to Caltrans with the exception of the sponsoring name and removing language about an incentive used in the Caltrans survey. 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 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 Data User's Guide in Appendix A 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. The retrieval questionnaire is included in the Data User's Guide in Appendix A 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 the two counties and required sample size, sample selection, sample frame generation and sample preparation.

Survey Universe and Sample Size

The universe for the TAMC survey is defined as all households located within Monterey and San Benito counties that had not already been selected to participate in the Caltrans survey.

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. 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 98.5%. 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.

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 two 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 the AMBAG region is 36% (Survey Sampling, Inc.).

A recruitment goal of approximately 1,000 households assumes a retrieval rate (that is, percentage of recruited households that will be completed) of 70%. This recruitment sample size and retrieval rate goal produces a minimum of 686 completed surveys. The final retrieval rate was achieved, therefore, a total of 693 surveys were completed from a base of 1,000 recruited households (69%).

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. 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 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 TAMC Household 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.

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.
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). 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 693 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 remaining 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 90% 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.

5. Survey Data Weighting and Expansion

The TAMC Household Travel Survey data set was combined with the 2000-2001 California Statewide Household Survey data set to provide a more robust sample. The TAMC survey included households located in the counties of Monterey and San Benito. The data from these household were combined with surveys from the counties of Monterey, San Benito, Santa Clara and Santa Cruz from the Statewide Survey.

The following is a summary of the number of completed surveys by county in the final data set.

Table 5.1
Surveyed Households by County

County	Households
Monterey	832
San Benito	291
Santa Clara	256
Santa Cruz	441
Total	**1,820**

1.   Statistical weights were used to adjust the sample proportions to actual proportions as determined by most recent Census data available (Year 2000 Census). The weight variable for the TAMC Household Travel Survey is comprised of two factors, household size distribution and county distribution (as a percentage of the four county region).

The weighting process includes the development of these factors, then multiplying them together to determine the "final" weight. Each weight factor for each data file (Household, Person and Vehicle) includes an expansion factor (which is the actual number of households, persons or vehicles in each county). The Trip File uses the same weight as the Person File.

Household Size Distribution

The calculation of this weight factor is based on the 2000 Census data provided by TAMC. For each county, the survey distribution of the household size is compared to the Census distribution. The weights for each cell are calculated using the following formula:

$$\text{Household Size Distribution Weight} = \frac{\text{Household Size Distribution (Census)}}{\text{Household Size Distribution (Survey)}}$$

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.

The categories used in developing the weight were for household sizes of one person through seven persons. The number of Census households, for each of the seven categories, was divided by the number of sampled households. The resulting weight was applied to the appropriate households in each household size category in the household file.

The process was repeated for the number of persons per county and the number of vehicles per county. The resulting expansions were applied to the corresponding data file.

The following tables summarize the Census and surveyed number of households, persons and vehicles and expansion factors for each corresponding data file.[1]

Table 5.2
Key 2000 Census Parameters by County – Household Size

County	Household Size						
	1	**2**	**3**	**4**	**5**	**6**	**7+**
Monterey							
Census	226	307	121	104	43	23	8
Survey	25,748	34,953	18,612	17,212	10,613	5,826	8,272
Factor	113.929204	113.853420	153.818182	165.500000	246.813953	253.304348	1034.000000
San Benito							
Census	60	117	49	38	17	9	1
Survey	2,245	4,310	2,750	2,982	1873	887	838
Factor	37.416667	36.837607	56.122449	78.473684	110.17647	98.555556	838.000000
Santa Clara							
Census	68	108	41	23	11	4	1
Survey	121,109	171,848	95,999	89,161	43,383	20,775	23,588
Factor	1781.014706	1591.185185	2341.439024	3876.565217	3943.909091	5193.750000	23588.000000
Santa Cruz							
Census	125	185	63	50	14	3	1
Survey	22,905	30,166	14,517	12,266	5,609	2,575	3,101
Factor	183.240000	163.059459	230.428571	245.320000	400.642857	858.333333	3101.000000

Table 5.3
Key 2000 Census Parameters by County - Persons[2]

County	Census Persons	Survey Persons	Expansion Factor
Monterey	380,786	2,032	187.394685
San Benito	52,727	739	71.349120
Santa Clara	1,652,871	585	2825.420513
Santa Cruz	246,574	979	251.863126
Total	**2,332,958**		

---

[1] Source: Transportation Agency of Monterey County (TAMC) and US Census Bureau.
[2] Persons in occupied housing units.

Table 5.4
Key 2000 Census Parameters by County - Vehicles

County	Census Vehicles[3]	Survey Vehicles	Expansion Factor
Monterey	234,276	1,798	130.298109
San Benito	35,974	705	51.026950
Santa Clara	1,179,292	541	2179.837338
Santa Cruz	185,298	938	197.545842
Total	**1,700,609**		

---

[3] Source: Transportation Agency of Monterey County (TAMC) and US Census Bureau.

6. Survey Results

Survey Results from Combined CalTrans and TAMC Surveys

This section contains the summary tables for weighted and expanded data and is based on unlinked trips. The results include all households in the dataset from the 2000-2001 CalTrans Household Travel Survey and the 2002 TAMC Household Travel Survey. The counties represented are Monterey, San Benito, Santa Clara, and Santa Cruz. For the 90 CalTrans households that had a two-day travel period, only the weekday travel data is represented.

The unweighted TAMC survey results are reported as summary tables in the Appendix D.

To run other results, refer to the Data User's Guide on how to select subsets of households. All trip-level results presented in this section and throughout the main report are based on unlinked trips.

HOUSEHOLD CHARACTERISTICS

TABLE 1
HOUSEHOLD SIZE DISTRIBUTION

HOUSEHOLD SIZE	FREQUENCY	PERCENT
One Person	172,007	22%
Two Persons	241,227	30%
Three Persons	131,878	17%
Four+ Persons	248,961	31%
Total	**794,123**	**100%**

Household sizes for this area of California are large. The overall average household size for this area is 2.9, and nearly eighty percent (78%) of households have more than one member. Nearly half (49%) of all households own two vehicles, while only one percent of households do not own a vehicle at all. The average vehicle ownership for these counties combined is 1.9.

TABLE 2
HOUSEHOLD VEHICLES

HOUSEHOLD VEHICLES	FREQUENCY	PERCENT
Zero Vehicles	9,985	1%
One Vehicle	190,064	24%
Two Vehicles	391,553	49%
Three+ Vehicles	202,521	26%
Total	**794,123**	**100%**

TABLE 3
HOUSEHOLD WORKERS

WORKERS	FREQUENCY	PERCENT
No Workers	140,577	18%
One Worker	338,072	43%
Two Workers	238,099	30%
Three + Workers	77,375	9%
Total	**794,123**	**100%**

Table 3 above displays the number of household workers. Majority of households (43%) have only one worker in the household. Nearly twenty percent (18%) of households do not have an employed person in the household. The average number of workers per household for this area is 1.35. There are half as many students as workers per household, with the average being 0.72 students per household. Nearly six in ten households (58%) do not have a person attending school, while less than ten percent of households (7%) contain more than two students. Table 4 below shows the distribution of household students.

TABLE 4
HOUSEHOLD STUDENTS

STUDENTS	FREQUENCY	PERCENT
No Students	465,132	58%
One Student	156,326	20%
Two Students	120,209	15%
Three + Students	52,456	7%
Total	**794,123**	**100%**

TABLE 5
HOUSEHOLD INCOME

INCOME	FREQUENCY	PERCENT
Below $10,000	14,051	2%
$10,000 to $24,999	52,657	8%
$25,000 to $34,999	67,206	10%
$35,000 to $49,999	63,281	9%
$50,000 to $74,999	140,130	20%
$75,000 to $99,999	128,132	18%
$100,000 to $149,999	141,672	20%
More than $150,000	92,684	13%
Total	**699,813**	**100%**

Base: Among those responding

The household income distribution is shown in Table 5 above. Household income for this area is relatively high, with over seventy percent (71%) of households reporting an annual income at least $50,000. Only ten percent

of households in this area make less than $25,000 per year. As expected with higher income households, most (71%) live in an unattached single family home. Another nine percent of households live in a duplex, condominium, or townhouse.

TABLE 6
RESIDENCE TYPE

RESIDENCE TYPE	FREQUENCY	PERCENT
Unattached Single Family Home	565,234	71%
Duplex	32,816	3%
Apartment	108,181	14%
Condominium or Townhouse	60,841	6%
Mobile Home or Trailer	16,487	5%
Group Quarters	1,558	1%
Total	**785,116**	**100%**

Base: Among those responding

PERSON CHARACTERISTICS

TABLE 7
GENDER

GENDER	FREQUENCY	PERCENT
Male	1,178,523	51%
Female	1,133,827	49%
DK/RF	20,608	1%
Total	**2,332,958**	**100%**

Base: Among persons responding. May not add to 100% due to rounding

There was an even distribution of male and female respondents who completed the survey. Table 8 displays the age distribution of household members. Over twenty percent (22%) of household members are under the age of 20. The most household members (51%) fell between the ages of 25 to 54 years old. Only five percent of household members are 75 years or older.

TABLE 8
AGE

AGE	FREQUENCY	PERCENT
Under 5 years	127,938	6%
5 years to 9 years old	129,024	6%
10 years to 14 years old	103218	5%
15 years to 19 years old	112,723	5%
20 years to 24 years old	94,714	4%
25 years to 34 years old	372,427	17%
35 years to 44 years old	377,321	17%
45 years to 54 years old	379,876	17%
55 years to 64 years old	273,136	12%
65 years to 74 years old	147,713	7%
75 years and older	118,327	5%
Total	**2,236,418**	**100%**

Base: Among those responding. May not add to 100% due to rounding.

Over seven in ten (72%) respondents are "White/non-Hispanic", while another thirteen percent of respondents classify themselves as "Hispanic." Only fifteen percent of respondents fall outside these two categories.

TABLE 9
ETHNICITY

ETHNICITY	FREQUENCY	PERCENT
White / non-Hispanic	1,598,608	72%
Hispanic	290,917	13%
African American	21,602	1%
Asian / Pacific Islander	166,851	8%
Native American	8,882	0%
European (non-specific)	1,383	0%
Other	80,671	4%
Mixed	66,497	3%
Total	**2,235,411**	**100%**

Base: Among those responding. May not add to 100% due to rounding.

TABLE 10
DRIVER'S LICENSE

DRIVER'S LICENSE	FREQUENCY	PERCENT
Yes	1,732,052	89%
No	159,485	8%
Don't Know / Refused	50,413	3%
Total	**1,941,950**	**100%**

Base: Among those responding. May not add to 100% due to rounding.

Nearly nine in ten (89%) of persons over the age of fifteen have a driver's license. Over half (55%) of household members are employed either full or part-time, while another six percent are self-employed. Fifteen percent of household members are retired.

TABLE 11
PRIMARY ACTIVITY

ACTIVITY	FREQUENCY	PERCENT
Not currently working for pay	321,193	16%
Employed full or part-time	1,084,332	55%
Self-employed	116,148	6%
Regular volunteer work (treat as worker)	10,176	1%
Retired	299,877	15%
Full-time homemaker	110,218	6%
Child not in school/infant	3711	0%
Other	2081	0%
Don't know / Refused	250,43	1%
Total	**1701**	**100%**

Base: Respondents ages 15 and over. May not add to 100% due to rounding.

More than seven in ten (72%) respondents have education to at least a high school diploma. Nearly one-fourth (24%) of all household members has accomplished an undergraduate/Bachelor's degree. Only three percent have no type of formal education at all.

TABLE 12
HIGHEST LEVEL OF EDUCATION FINISHED

EDUCATION LEVEL	FREQUENCY	PERCENT
No school completed	74,635	3%
Preschool/nursery school	65,571	3%
K-4th grade	135,251	6%
5th-8th grade	110,061	5%
9th-12th grade (no diploma)	151,292	7%
High school graduate	238,025	10%
Some college, no degree	290,705	13%
Associates degree	108,479	5%
Undergraduate/Bachelor's degree	557,900	24%
Some graduate school/ no degree	56,334	2%
Master's degree	302,082	13%
Professional school degree	33,466	1%
Doctorate or higher degree	85,371	4%
DK/RF	123,786	5%
Total	**2,332,958**	**100%**

Base: Among persons responding. May not add to 100% due to rounding.

TABLE 13
STUDENT STATUS

STUDENT STATUS	FREQUENCY	PERCENT
Yes	501,212	22%
No	1,173,542	76%
DK/RF	58,204	2%
Total	**1701**	**100%**

Base: Among persons responding. May not add to 100% due to rounding.

Over twenty percent (22%) of all household members are currently enrolled in school. Of those enrolled in school, nearly thirty percent (28%) are enrolled in either a college or university. Over half (56%) of students are enrolled in kindergarten through 12th grade. Table 14 on the following page displays the school types for those enrolled.

TABLE 14
SCHOOL TYPE

SCHOOL	FREQUENCY	PERCENT
Daycare/Preschool	32,601	7%
K-6th grade	164,587	33%
7th-12th grade	114,146	23%
College/University	142,510	28%
Trade/Vocational	14,041	3%
Post Graduate	7,879	2%
Other (specify)	24,729	5%
DK/RF	2,217	0%
Total	**502,711**	**100%**

Base: Among persons responding. May not add to 100% due to rounding.

TRIP CHARACTERISTICS

The following tables display trip characteristics. The average number of weekday trips per person is 3.5, while the average number of weekday trips per household is 9.5. Household size and the number of vehicles per household directly affect the average number of trips per household, as seen in Tables 16 and 17.

TABLE 15
TRIP RATES

TYPE	RATES
Person	3.53
Household	9.46

Base: 2,332,958 Persons /794,123 Households.

TABLE 16
HOUSEHOLD TRIP RATES BY HOUSEHOLD SIZE

HOUSEHOLD SIZE	RATES
One Person	3.43
Two Persons	7.05
Three Persons	11.31
Four+ Persons	14.98
Total	**9.46**

Base: 794,123 Households.

TABLE 17
HOUSEHOLD TRIP RATES BY HOUSEHOLD VEHICLES

NUMBER OF HH VEHICLES	RATES
Zero	2.77
One	5.10
Two	9.90
Three +	13.04
Total	**9.46**

Base: 794,123 Households.

TABLE 18
TRAVEL MODE

MODE	FREQUENCY	PERCENT
Drove	5,734,258	70%
Passenger in car / truck / van	1,652,396	20%
Public Transit	22,347	1%
School Bus	54,483	1%
Taxi/Shuttle/Limo	1,821	0%
Motorcycle/Moped	143	0%
Bicycle	82,217	1%
Walk	584,437	7%
Other	12,890	0%
Don't know / Refused	1,758	0%
Total	**9,042,72 0**	**100%**

Base: 8,239,410 Trips. May not equal 100% due to rounding

Table 18 shows the mode distribution. Nine in ten trips are made using a personal automobile, while another seven percent of trips are made on foot. Only one percent of trips in this area are made using public transit. Table 19 displays the mean travel time by travel mode distribution. The average length of a driving trip is just under twenty minutes. School bus trips last an average of forty minutes, and express bus trips are closer to fifty minutes. The average length of all trips is nineteen minutes.

TABLE 19
MEAN TRAVEL TIME (IN MINUTES) BY TRAVEL MODE

MODE	MEAN
Drove	19.41
Passenger in car / truck / van	17.93
Express Bus	49.10
School Bus	40.24
Taxi/Shuttle/Limo	20.50
Bicycle	13.87
Walk	14.27
Total	**19.02**

Base: 8,239,410 Trips. May not equal 100% due to rounding.

TABLE 20
PRIMARY TRIP PURPOSE

TRIP PURPOSE	FREQUENCY	PERCENT
Working at home (related to main or second job)	149,066	1.4%
Eating / preparing meals at home	1,024,232	9.7%
Watching TV / Videos at home	287,830	2.7%
Shopping by phone/ TV/Internet at home	10,183	0.1%
Exercising at home	21,767	0.2%
Other activities at home	3,510,533	33.3%
Travel by car, bus walk, etc.	44,889	0.4%
Wait for / get on a vehicle	17,433	0.2%
Leave / park a vehicle	42,762	0.4%
Boarding activities for airplane, rail, intercity bus	36,175	0.3%
Getting off airplane, rail, intercity bus	339,582	3.2%
Pick up someone or get picked up	427,258	4.1%
Drop off someone or get dropped off	1,095,674	10.4%
Work (includes regularly scheduled volunteer work)	339,172	3.2%
Work-related (sales calls, meetings, errands, etc.)	287,763	2.7%
School (Preschool, K-12th)	81,282	0.8%
School (Post-secondary - college, vocational)	35,237	0.3%
Childcare, day care, after school care	434,007	4.1%
Eat out (restaurant, drive-through, etc.)	142,442	1.4%
Medical	157,341	1.5%
Fitness activities (Gym/health club/participating in	75,499	0.7%
Recreational (vacation, camping, etc.)	115,906	1.1%
Entertainment (movies, dance club, bar, spectator	225,522	2.1%
Visit friends / relatives	40,947	0.4%
Community meeting, political or civic event, public	20,641	0.2%
Occasional volunteer work	51,929	0.5%
Church, temple, religious meeting	59,751	0.6%
Buy gas	661,931	6.3%
Incidental shopping (groceries, housewares, medicine,	71,602	0.7%
Major shopping (furniture, clothes, autos, etc.)	172,388	1.6%
ATM, banking, post office, utilities	445,486	4.2%
Other personal or household business away from home	97,637	0.9%
Other specify	375	0%
Be with another person at their activity	1,505	0%
Don't know / Refused	9,122	0.1%
Total	**10,534,870**	**100%**

Base: 10,534,870 Trip records. May not add to 100% due to rounding

At home activities make up nearly half (47.4%) of all daily activities. Serve passenger is the second most frequent activity with nearly fifteen percent of responses. Work and work-related activities account for over five percent of the daily activity distribution.

The overall average trip length is nineteen minutes, and the average travel time to work is slightly longer than the overall average at twenty-one minutes. The longest travel time is over fifty-six minutes to "fitness activity." The average trip time to school (K-12th) is twenty minutes, while trips to college or vocational schools are shorter at sixteen minutes.

TABLE 21
TRAVEL TIME (IN MINUTES) BY TRIP PURPOSE (TRAVEL TO)

TRIP PURPOSE	MEAN
Working at home (related to main or second job)	14.0
Eating / preparing meals at home	20.8
Watching TV / Videos at home	15.8
Shopping by phone/TV	20.3
Exercising at home	15.1
Other activities at home	19.8
Travel by car, bus walk, etc.	18.0
Wait for / get on a vehicle	22.7
Leave / park a vehicle	15.7
Boarding activities for airplane, rail, intercity bus	26.3
Getting off airplane, rail, intercity bus	15.7
Pick up someone or get picked up	14.8
Drop off someone or get dropped off	22.0
Work (includes regularly scheduled volunteer work)	21.2
Work-related (sales calls, meetings, errands, etc.)	14.9
School (Preschool, K-12th)	19.9
School (Post-secondary - college, vocational)	15.9
Childcare, day care, after school care	15.0
Eat out (restaurant, drive-through, etc.)	20.5
Medical	18.6
Fitness activities (Gym/health club/participating in	56.7
Recreational (vacation, camping, etc.)	21.6
Entertainment (movies, dance club, bar, spectator	23.1
Visit friends / relatives	12.1
Community meeting, political or civic event, public	30.9
Occasional volunteer work	16.5
Church, temple, religious meeting	12.1
Buy gas	16.0
Incidental shopping (groceries, housewares, medicine,	21.5
Major shopping (furniture, clothes, autos, etc.)	12.4
ATM, banking, post office, utilities	19.6
Other personal or household business away from home	21.9
Be with another person at their activity	42.4
Don't know / Refused	21.2
Total	**19.0**

Base: 10,534,870 Trip records. May not add to 100% due to rounding

TABLE 22
TRIP DEPARTURE TIME

DEPARTURE TIME	FREQUENCY	PERCENT
6 a.m. to 9:59 a.m.	2,048,852	19%
10 a.m. to 3:59 p.m.	3,382,003	32%
4 p.m. to 7:59 p.m.	2,104,407	20%
8 p.m. to 4:59 a.m.	2,999,607	29%
Total	**10,534,870**	**100%**

Base: 10,534,870 Trip records. May not add to 100% due to rounding

The most frequent departure time for all trips is between 10am and 3:59pm (32%). Trips departing during the morning peak hours (6 am-9:59 am) account for nearly twenty percent (19%) of all trips. The average vehicle occupancy for all weekday trips, regardless of mode, is 1.49 persons, while driving trips have an average of 1.45 persons per vehicle.

TABLE 23
MEMBERS IN TRAVEL PARTY (ALL MODES)

Mean Occupancy	1.49

Base: 8,239,410 trip records.

TABLE 24
DRIVER VEHICLE OCCUPANCY

Mean Occupancy	1.45

Base: 5,734,258 Driver trip records.

Tables 25 and 26 display the top five makes and models of vehicles owned by household members. Honda is among the top five makes with eleven percent of responses, and accounts for two of the top five models owned by respondents. Toyota and Ford are the two most frequently owned makes with nearly fifteen percent of responses each. Most (65%) vehicles are automobiles, while nearly ten percent (9%) of vehicles in the area are sport utility vehicles. Gasoline accounts for ninety-six percent of all fuel types used for vehicles owned. More than ninety percent (94%) of vehicles are owned by a household member, while three percent of vehicles are leased. Over six in ten (63%) vehicles are more than five years old. The average age for vehicles is 9.5, and vehicles less than three years old account for less than fifteen percent.

TABLE 25
TOP FIVE VEHICLE MAKES

VEHICLE MAKE	FREQUENCY	PERCENT
Toyota	216,487	14%
Ford	211,324	14%
Honda	163,026	11%
Chevrolet	115,093	7%
Dodge	90,907	6%

TABLE 26
TOP FIVE VEHICLE MODELS

VEHICLE MAKE	FREQUENCY	PERCENT
Civic	62,134	4%
Camry	55,950	4%
Accord	53,913	3%
1/2 Ton Truck	47,640	3%
Corolla	31,818	2%

TABLE 27
VEHICLE BODY TYPE

BODY TYPE	FREQUENCY	PERCENT
Auto	1,001,878	65%
Van	117,265	8%
RV	12,438	1%
Sport Utility Vehicle	141,587	9%
Pick-up Truck	196,161	13%
Other Truck	16,223	1%
Motorcycle/Moped	18,655	1%
Other (Specify)	2,440	0%
DK/RF	34,850	2%
Total	1,541,498	100%

TABLE 28
VEHICLE FUEL TYPE

FUEL TYPE	FREQUENCY	PERCENT
Gasoline	1,485,422	96%
Diesel	15,301	1%
Electricity	6,935	<1%
Other	4,818	<1%
DK	3,294	<1%
RF	25,729	2%
Total	1,541,498	100%

TABLE 29
VEHICLE OWNERSHIP

OWNERSHIP	FREQUENCY	PERCENT
Owned by a household member	1,451,851	94%
Owned by a person not in household	16,386	1%
Leased	45,013	3%
DK	22,250	1%
RF	5,999	0%
Total	1,541,498	100%

TABLE 30
AGE OF VEHICLE

VEHICLE AGE	FREQUENCY	PERCENT
<3 Years	207,315	14%
3-5 Years	345,365	23%
6-10 Years	397,460	27%
>10 Years	543,158	36%
Total	1,493,298	100%
Mean Vehicle Age	9.5	

7. 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 7.1
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	$\sqrt{}^{4}$	$\sqrt{}$
Language	$\sqrt{}^{9}$	Not Available

Person File Comparison

Travel Survey Variable	2000CSS	1995 NPTS
Gender	$\sqrt{}$	$\sqrt{}$
Age	$\sqrt{}^{9}$	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
[4] Recategorizing of data required to allow comparison

The following are the comparisons for each variable listed above.

Table 7.2
Data Comparison

Travel Survey Variable	AMBAG	CalTrans	2000CSS	1995 NPTS
Household Size	2.9	2.83	2.87	2.63
Vehicles	1.9	1.86	Not Available	1.78
Gender:				
Male	51%	48.90%	49.80%	48.70%
Female	49%	50.70%	50.20%	51.30%
Dwelling type:				
Single	71%	67.90%	67.50%	
Multiple	24%	31.20%	32.40%	
Owner Status:				
Own	70.3%	54.80%	56.90%	
Rent	28.5%	44.20%	43.10%	
Income:				
Less than $10,000	2%	9%	8%	
$10,000 to $24,999	8%	19%	18%	
$25,000 to $34,999	10%	11%	11%	
$35,000 to $49,999	9%	13%	15%	
$50,000 to $74,999	20%	17%	18%	
$75,000 to $99,999	18%	11%	11%	
$100,000 to $149,999	20%	11%	10%	
$150,000 or more	13%	8%	7%	
Age (median)	40	33	33.3	
Industry:				
Agriculture, forestry, fishing and hunting, and mining	1.80%	3.80%	2.10%	
Construction	5.10%	6.10%	6.30%	
Manufacturing	7.40%	5.10%	13.00%	
Wholesale trade	0.40%	0.90%	4.10%	
Retail trade	5.10%	8.00%	11.20%	
Transportation, warehousing & utilities	2.80%	4.90%	4.60%	
Information	5.50%	2.50%	3.90%	
Finance, insurance, real estate, and rental and leasing	3.90%	5.70%	6.90%	
Professional, scientific, management, management services	30.60%	16.00%	11.60%	
Educational, health and social services	19.50%	21.50%	17.50%	
Arts, entertainment, recreation, accommodation and food services	5.90%	7.80%	8.70%	
Other services (except public administration)	6.50%	1.00%	5.20%	
Public Administration	3.80%	6.30%	4.50%	

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 survey data collection was conducted during spring 2002. Telephone calls were randomly made to households in the San Benito and Monterey Counties, California. Recruitment calls were made six days per week during the recruitment phases and seven days per week during the retrieval phase.

During the recruitment phase, household and person demographic data were collected and households who agreed to participate in the survey were assigned a one-day (24-hour) travel date. Travel days were assigned Monday, Tuesday, Wednesday, Thursday or Friday April 2 – Thursday April 26, 2002.

A total of 820 households in Monterey County were recruited and 180 were recruited in San Benito. In Monterey County, 561 households provided diary data. In San Benito, 136 provided diary data.

The survey had an overall response rate of 15%. Response rates for similar household travel surveys (and using the same formula) ranges between 20% and 35% depending on the study area and length of the survey.

The response rate is the total number of households completing a dairy as a percentage of all households *contacted* (not just recruited) during recruitment. 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.

Since the Monterey and San Benito Household Travel Survey used a multistage survey process (i.e., household recruitment and household retrieval), a rate is calculated for each survey stage — called the recruitment and retrieval component response rates, then the overall response rate is the product of the two rates.

Recruitment Component Response Rate

From the sampling frame, 9,894 telephone numbers were randomly generated in the two counties. These were the basis from which households were identified and recruited for inclusion in the study.

As shown in Exhibit 2.1, a total of 2,074 telephone numbers was found to be ineligible for the study (disconnects, business/government, and computer/fax), respectively. Dividing the total eligible units (2,038) by the sum of the total eligible and ineligible units (4,112) is the eligibility rate for the recruitment phase (49.6%). This rate is used to determine the number of eligibility unknown numbers for each region to allocate to the response rate calculation. Therefore, 60.1% of the total eligibility unknown of 5,782 – no answer, busy, answering machine, callbacks, etc. (2,868) is added to the denominator of the formula.

The recruitment response rate is therefore calculated using the following formula:

Recruitment Response Rate = Recruits / Eligible+49.6% of Eligibility Unknown Units)
Recruitment Response Rate = 1,000 / 2,038+2,868
Recruitment Response Rate = 20.4%

Exhibit 2.1
Recruitment Call Outcomes

Call Outcome	Frequency
ELIGIBLE UNITS	
Recruited	1,000
Refused to participate	964
Partial completes	74
SUB-TOTAL ELIGIBLE	2,038
INELIGIBLE UNITS	
Disconnected/non-working	496
Business/Government	505
Facsimile	575
Over Quota	6
Out of area	492
SUB-TOTAL INELIGIBLE UNITS	2,074
ELIGIBILITY UNKNOWN UNITS	
No answer	4,353
Call Back	1,141
Answering machine	62
Busy	38
Communications Barrier	188
SUB-TOTAL ELIGIBILITY UNKNOWN UNITS	5,782
GRAND TOTAL:	*9,894*

Retrieval Component Response Rate

The retrieval rate is then calculated using the same formula as the recruitment survey component rate. Of the total 1,000 recruited households virtually all are eligible since the vast majority had been contacted previously (the only ineligible households are those in which the home phone was not in service at the time the retrieval contact was attempted). The retrieval component response rate is therefore the number of completed surveys divided by the total sample (i.e., all recruits minus disconnects). The retrieval component response rate is 70.9%.

Exhibit 2.2
Retrieval Call Outcomes

Call Outcome	Frequency
ELIGIBLE UNITS	
Completed	697
Refused to participate	117
Pending (no answer, call backs, answering machines)	169
SUB-TOTAL ELIGIBLE	983
INELIGIBLE UNITS	
Disconnected/non-working	17
SUB-TOTAL INELIGIBLE UNITS	17
GRAND TOTAL:	*1,000*

Overall Response Rate

The overall response rate is calculated after the recruitment and retrieval component rates are calculated. The overall response rate is 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 (note that in the retrieval phase all households, with the exception of disconnects, 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 15% (0.204 * 0.709). The response rate calculation uses the same formula prescribed by the Council of American Survey Research Organizations (CASRO) and the Federal Highway Administration (FHWA).