SOUTHERN CALIFORNIA ASSOCIATION OF GOVERNMENTS

YEAR 2001 POST-CENSUS REGIONAL TRAVEL SURVEY

Final Report of Survey Methodology

June 30, 2003

NuStats

3006 Bee Caves Rd., Suite A-300 . Austin, Texas 78746
(512) 306-9065 . fax (512) 306-9077 . *www.nustats.com*

Contact: Johanna Zmud, President

The preparation of this report was financed in part through a grant from the United States Department of Transportation and the Federal Highway Administration. It is a result of a study being conducted by NuStats, LP on behalf of the Southern California Association of Governments, with support from the Los Angeles County Metropolitan Transportation Authority, California Air Resources Board, Riverside County Transportation Commission, and the San Bernardino Association of Governments.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

The Year 2001 Post-Census Regional Travel Survey was a large-scale regional household travel survey conducted in six counties in Southern California, under contract to the Southern California Association of Governments (SCAG). In addition to providing basic information about each household and its members, the survey captured specific characteristics of activities and trips made, including number, purpose, the time of day, travel mode and questions specific to the travel modes used.

The Household Travel Survey was designed by SCAG and its partner agencies to include several different sample types. These types included both probability and non-probability samples. This report focuses on the survey components that comprised the largest portion of the regional survey effort (see bullets below). Because these sample surveys were conducted via telephone using Random Digit Dial (RDD) sampling methods, cases from each can be combined to produce a larger base from which to conduct analyses. For purposes of this report, these samples are collectively referred to as the Household Travel Survey. The Household Travel Survey included the following probability samples:

- **Base Sample** of households in the SCAG region across its constituent counties (i.e., Los Angeles; Orange; Riverside; Imperial; San Bernardino; Ventura). This sample was used to collect 24-hour (weekday) diaries.
- **Caltrans Sample**, which comprised the six-county Southern California portion of the greater statewide Caltrans household travel survey. This sample was drawn from the same sampling frame as the base sample (above) at the same time and was conducted with a field period that overlapped with the SCAG study.
- **Regional Statistical Area (RSA) Augment**, which was an oversample of subregional geographic areas (RSAs) in the SCAG region with the intent to increase sample at the subregional level.[1] The RSA sample was drawn independently from the Base, Caltrans, and Weekend samples during the latter stages of data collection.
- **Weekend Sample** was designed to measure travel behavior on weekends using a 48-hour diary, whereas the Base, Caltrans, and RSA augment samples of households were designed to measure weekday household travel behavior. The weekend sample was drawn during the same "pull" as the Base and Caltrans samples. All weekend households recorded travel during a 48-hour period that spanned either Friday/Saturday or Sunday/Monday.

A fourth household survey component covered a Mode User Augment. The Mode User Augment entailed the collection of seven independent samples of specific types of mode users for the purpose of implementing mode specific statistical models of travel volume and behavior. The results of this survey are presented as a separate report. A fifth household survey component entailed the collection of information using global positioning system (GPS) technology from a sub-sample of households who completed the 24-hour travel logs. The results of this survey component are presented as a separate report.

The details of the Household Travel Survey have been provided in two volumes based on expected interest. This document, the *Final Report of Survey Methods*, is focused on the conduct of the survey and primarily would be of interest to those investigating survey methods or planning their own survey. A separate volume, the *Final Report of Survey Results*, is focused on the survey results and primarily would be interest to data users.

[1] This effort was curtailed during the fieldwork period due to contractual issues, and thus, resulted in far less oversampling at the subregional level than specified in the original sample design.

The Household Travel Survey began with sample design, instrument development, and pilot testing in Fall 2000. Data collection was conducted during Spring 2001, Fall 2001, and Spring 2002. Cleaning, geocoding, and data delivery was done on a continuous basis. The Final Work Plan for the survey effort is provided as Appendix A. This plan was developed early in the project. While it accurately depicts the tasks that were accomplished to achieve the survey objectives, the schedule was quite different due to contractual issues.

In total, 27,476 households were recruited to participate in the Household Travel Survey. Interviewing was conducted in English and Spanish through a computer-assisted telephone interviewing (CATI) system. Of these 17,775 households completed travel diaries, and the information was retrieved from all household members. These data were captured using place-based, activity-focused diaries. In total, 16,506 households passed all of the quality assurance criteria. This is the total number of households delivered to SCAG, and the survey results reported in this document are based only on these 16,506 households. The map on the next page plots the location of these households throughout the region.

NuStats conducted the Household Travel Survey, in association with DataSource, Batelle, and Resource Systems Group. NuStats designed the survey, managed data collection, analyzed the survey data, and documented the results in a series of reports. DataSource, a NuStats' affiliate, conducted the telephone interviews and mailed the diary packets. Batelle designed a GPS data collection strategy for a subsample of the population, managed the data capture, and analyzed the resulting data. Resource Systems Group supported survey data collection effort by building and managing an Internet data collection option. In addition, Cheryl Stecher, Franklin Hill Group, was contracted by SCAG to provide ongoing quality assurance assistance.

Map of home locations

SAMPLING DESIGN AND PROCEDURES

This chapter provides specifications for the sampling approaches used for the SCAG suite of population surveys (also see Appendix B: Technical Memorandum on Sampling). The integrated population surveys comprised the largest portion of the regional survey effort. They were integrated in the sense that all samples was drawn from a single sampling frame so that cases from each sample could be combined to produce a larger base from which to conduct analyses. All of these population surveys were conducted via telephone using Random Digit Dial (RDD) sampling methods. Three of the four population surveys were used to examine 24-hour weekday travel behavior among household members. For this report, the three surveys are known collectively as the Main Survey. The Main Survey included:

- A base sample of households in the SCAG region;
- The Caltrans sample, representing the SCAG regional portion of a statewide Caltrans survey of households regarding travel behavior; and
- The Regional Statistical Area (RSA) augment, an oversample of subregional geographic areas in the SCAG region to increase sample in specific geographic areas that would allow the production of nominal estimates from specific RSA subregions of the 6-county region.[2]

The Weekend Survey sample was drawn as part of the integrated population surveys; however, the weekend sample denotes a separate metropolitan area sample of households that was designed to measure 48-hour travel behavior on weekends.

SURVEY UNIVERSE

The population of inference for the population surveys comprises households and household residents in the six-county SCAG region (i.e., Los Angeles, Orange, Riverside, Imperial, San Bernardino, Ventura). The universe of households in the SCAG region totals 5,386,491 occupied housing units. The inferential capabilities of the SCAG survey rests on probability sampling from a frame covering all members of the Universe (in this case, occupied housing units). Because a telephone survey was used to recruit households into the sample, the survey population was households with telephones in the SCAG region. The sampling universe was total household telephone numbers, and the sampling frame was the collection of ten-digit telephone numbers from the listed working banks that fell into the SCAG region.

SAMPLE DESIGN AND SELECTION

The Household Travel Survey was intended to represent the diverse population and travel patterns of residents of the SCAG six-county region. The methodological cornerstone of such a program is the sample design and the manner in which it is implemented. The primary criteria that should be met by the sample design are:

- To produce a statistically adequate number of observations at a geographic level that meets SCAG's (and its partners') modeling objectives;
- To produce data depicting the diverse travel behavior in the region;
- To minimize selection bias across subgroups in the population, particularly those that are more difficult to reach;

[2] The RSA augment data collection was curtailed during the fieldwork period due to contractual issues, and thus, resulted in far less oversampling at the subregional level than was specified in the original Technical Memorandum on Sampling.

- To maximize overall participation rates.

The first two criteria are fundamental to the design of the sample. They are primarily addressed via the statistical method applied, the sample frame used, the actual process for drawing the sample, and the documentation maintained. These are addressed below. Meeting the other two criteria depends on a combination of sample design and the procedures applied for managing and controlling the sample. These procedures are discussed further in this section.

The surveys were conducted using telephone sampling for household selection and screening, followed by diaries mailed to eligible households and completed by all residents. A highly stratified RDD sample of households (with telephones) was used. The sampling frame for the stratified probability sample included both listed and unlisted telephone numbers from working blocks of numbers that make up the six-county SCAG region. "Blocks" of numbers were defined as groups of 100 consecutive telephone numbers whose area code, exchange and stem (i.e., the last four digits) were identical except for the last (rightmost) two digits (which ranged in value from 00 to 99). "Working blocks" denote all blocks for which at least one telephone number in the block is a listed residential telephone number. Working blocks included both listed and unlisted telephone numbers, and collectively covered a high percentage of all residential number in the area. The sampling frame was defined as the collected of telephone numbers in all working blocks which covered the six-county region.

Data collection targets by county that were identified by SCAG guided the selection of elements from the frame. The sample for the Base, Weekend, and Caltrans studies was drawn at the same time, and then divided into partitions based on the data collection targets. The sample was drawn at the county level, as shown in Table 1, and represents disproportionate sampling by county when considering the universe of occupied housing units at the county level.

TABLE 1: DATA COLLECTION TARGETS BY COUNTY

County	Occupied Households	% of Occupied Households	Base Target	Weekend Target	Caltrans Target	Total Target	% of Total
Imperial	39,384	0.7	500	21	500	1,021	5.6
Los Angeles	3,133,774	58.2	5,500	1,887	500	7,887	43.6
Orange	935,287	17.4	1,500	508	500	2,508	13.8
Riverside	506,218	9.4	1,500	254	500	2,254	12.5
San Bernardino	528,594	9.8	1,500	294	500	2,294	12.7
Ventura	243,234	4.5	1,500	137	500	2,137	11.8
Total	5,386,491	100.0	12,000	3,100	3,000	18,100	100.0

Source: Table 1 of Technical Memorandum on Sampling, April 2001.

The original sample design specified that a data collection target of 2,217 households for the RSA Augment portion of the survey. This target was expected to provide the marginal increases in sample by RSA to achieve the nominal desired sample sizes (i.e., a minimum of 200 households per RSA), excepting those RSAs for which a nominal representation was not desired. The fielding of the RSA augment sample was to be the last fieldwork component of the Household Travel Survey. However, this data collection was curtailed early in its field period due to contractual issues, such that the RSA augment data collection target was eliminated.

DATA WEIGHTING

This section presents the process by which the resultant data were weighted to account for the probability of selection and adjusted for non-response bias. The expansion factor calculation is also presented. (See also Appendix C: Technical Memorandum on Weighting and Expansion.)

Probability of Selection (Sampling Weight)

A probability sample assigns known nonzero chances of selection to every member of the survey population. Inequalities in the selection frame and procedures may create unequal selection probabilities. These are corrected with weights inverse to those probabilities. The probability of selection is defined as the probability of being selected from the frame for the county of residence. The population parameter for this calculation (i.e., total household telephone numbers) was derived from Acxiom's InfoBase telephone directories. The probability was calculated by dividing the number of sample pieces selected for a given county by the universe of household telephone numbers in that same geography. The six probabilities of selection or sampling weights by county are:

1. Imperial County: 6,798 telephone numbers drawn from a universe of 79,600 household telephone numbers = 0.0854.
2. Los Angeles County: 222,191 telephone numbers drawn from a universe of 6,971,600 household telephone numbers = 0.0318.
3. Orange County: 82,875 telephone numbers drawn from a universe of 1,572,700 household telephone numbers = 0.0526.
4. Riverside County: 58,684 telephone numbers drawn from a universe of 1,038,200 household telephone numbers = 0.0565.
5. San Bernardino County: 54,070 telephone numbers drawn from a universe of 995,600 household telephone numbers = 0.0543
6. Ventura County: 41,990 telephone numbers drawn from a universe of 515,600 household telephone numbers = 0.0814.

Sampling weights were also computed for the RSA augment as is documented in the technical memorandum in Appendix B. Due to small sample sizes in some RSAs, the RSAs were collapsed for this calculation. The sampling weights for the combined sample (Main Survey, plus RSA) considered the overall probability of selection ($PT_{i=} P_{1+} P_{2i}$ for all sample households in the i*th* RSA). So for example, a household in Los Angeles County RSA 7 would need to receive an overall probability of selection that was the sum of the sampling weights from the Main Survey and the RSA Augment. For RSAs in which there is no RSA Augment sample households, the probability of selection was simply the Main Survey sampling weight. Once PT_i is determined, the corresponding weight was $W_i = 1/PT_i$.

Post Stratification Weights

Post-stratification is a way of improving survey sample estimates by the proper utilization of *ancillary* sources of information. Post-stratification requires (a) information on the proportions of the universe on the variables of interest, and (b) information for classifying the sample cases into the same strata. For our purposes, the information on the universe was derived from Census 2000, Summary File 2.

Typically, post-stratification cells are formed by cross-classifying demographic cells (thought to be related to nonresponse). Unfortunately, at the time of this weighting procedure, the necessary population parameters from Census 2000 were not available in cross-classification cells. A "raking" methodology was used that, in mathematical terms, is called iterative proportional fitting. As the name implies, this is an iterative technique that required sequential matching of weighted sample marginals to first one

variable, then another, and so on until the sample proportions of variables of interest matched the proportions for those variables in Census 2000. The variables involved were household size and number of vehicles. These were chosen because:

- They are closely related to the travel variables of interest
- Census data are available for these variables
- Every household has complete data for these variables (there is no non-response).

In addition, an adjustment was made to bring the sample distributions in line with the population parameters by county since it was important to produce volume estimates at the county level. The technical memorandum in Appendix B reveals the process that was involved in the "raking" methodology.

Expansion Factor

The final weight computed for these data was the expansion factor. It was calculated by dividing the number of universe households by the number of households in the sample (5,385,491 / 16,506 or 326.3353). This final weight for each household unit will be unique based on its geographic location (both county and RSA), as well as its household size and vehicle availability. However, it is a "stand alone" weight in that it does not change if the unit of analysis falls below the county level.

.

SURVEY DESIGN AND PILOT TESTS

The objectives of the Main Survey required comprehensive survey instruments to collect demographic and socioeconomic details about households and persons, vehicle information, details of school and work travel, and detailed data of all trips made on an assigned travel day. The survey employed a generally accepted research method for household travel behavior research, in which all household members were asked to record all trips for a specified 24-hour (or 48-hour) period (from 3 a.m. to 2:59 a.m.) using a specially designed travel log. Parents or guardians were asked to complete the travel logs of household members under the age of 12. Households, completing 24-hour logs, were randomly assigned to non-Holiday weekdays for recording their travel (Monday-Friday), while those completing 48-hour logs were randomly assigned to non-Holiday pairs of Friday-Saturday or Sunday-Monday. The survey began with an in-depth review of data needs that would satisfy the modeling requirements and analysis plans that would be relying on the survey data. Those data needs are summarized in Table x.

TABLE X: DATA ITEMS COLLECTED

Household Data File	Person Data File	Vehicle Data File	Travel / Activity Data File
For each household	**For each person in HH**	**For each HH vehicle**	**For each person trip**
Household Size	Gender	Year	Destination Address
# Vehicles Owned	Age	Make	Departure time
# Bicycles Owned	Disability Status	Model	Arrival Time
Household Income	Ethnicity	Fuel Type	Trip Purpose / Activity (1, 2)
Dwelling Unit Type	Licensed Driver	Body Type	Trip Duration
Tenure (Own/Rent)	Employment Status	Owned / Leased	Mode
Home Address	Job Location (1st, 2nd)	Lessee	# HH members traveling with
Language Spoken in Home	# Paying Jobs	Usual Driver	ID of HH members
# Telephone Lines	# Hours / Wk (1st, 2nd)	Year Acquired	# Non-HH members traveling
# Dedicated Fax/Modem Lines	Fixed/ Varied Schedule		HH vehicle used
# Shared Phone Lines	Industry (1st, 2nd)		Parking Cost
Discontinued Phone Service	Occupation (1st, 2nd)		Parking Payment Method
Length of Discontinued Service	Required to Drive Veh		Bus Stop Boarding Location
Technology in HH	Parking Costs at Job		Bus Stop Arrival Time
	Parking Subsidy at Job		Bus Stop Access
	Transit Subsidy at Job		Bus Stop Boarding Time
	Usual Mode to Work		Bus System
	Student Status		Bus Route
	School type		Transit Fare
	# Days / Wk		Bus Stop Alighting Location
	School Location		Bus Stop Alighting Time
	Educational Attainment		Bus Stop Egress
	Relationship to Resp.		Used Toll Road
			Toll Road Name
			Toll Cost
			Toll Payment Method
			Carpool or Diamond Lane Use

Data collection was via a multi-mode method consisting of telephone recruitment of households, the placement of respondent materials through the U.S. Mail, and telephone retrieval of travel data after the assigned travel day. This survey commenced shortly after the start of the statewide household travel survey being conducted by Caltrans. It was the intention of both SCAG and Caltrans that the regional and statewide surveys use methods and instruments, which would permit the later combination of data. As the contractor for the Caltrans survey, NuStats fulfilled this intention.

An Internet-based data retrieval program was developed for this survey. However the timing of the implementation of Internet retrieval was interrupted due to contractual issues. Therefore, only 50 households used the Internet option. Those data have not been broken out for analysis in this report or in the *Final Report of Survey Results.*

PILOT TESTS

A pilot test was conducted in November and December 2000. The pilot test was designed as a full "dress rehearsal" and allowed for the full evaluation of the survey procedures from sample generation to data file preparation. The goal was to obtain 250 completed households, among several specific sample types. These sample types included the main sample, weekend sample, and mode augment sample. The objectives of the pilot test were to test and refine the survey procedures, survey materials and the Computer-Assisted Telephone Interviewing (CATI) programs. Specific pilot objectives included:

- Explore differential levels of respondent cooperation and response rates.
- Examine all stages of data flow procedures and quality assurance processes.
- Compare the geocoding results using GIS databases from Thomas Brothers Maps (TBM) and Geographic Data Technologies (GDT).
- Test two different travel diary formats.
- Evaluate mode augment sampling efficiency.

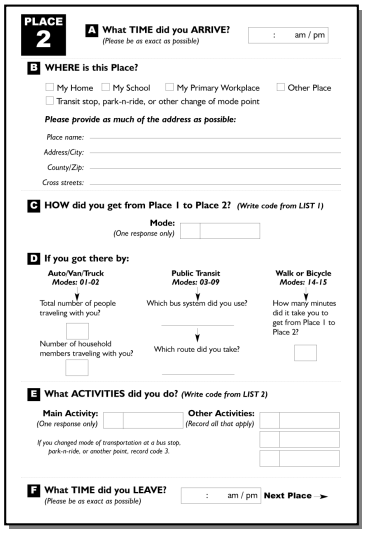
The pilot test produced data from 286 recruited households and 216 completed households. Our recruitment of households was not as productive as we expected. This situation was a result of the timing of our pilot (proximate to the holidays), the abbreviated schedule for pilot, and also a reflection of the challenge of conducting telephone surveys, in general. More than 70% of all call attempts were non-contacts. The maximum number of call attempts to a single sampled household was 17 – well above the RFP requirement of eight (8). However, our analyses indicated that such persistent calling was effective in reaching a household member. Once we reached a household member, our success in recruiting a household relative to obtaining a final refusal to participate was basically one of three. Completed interviews accounted for 38% of "known" eligible household contacts. Based on these findings, we revised our introduction script to be more compelling to potential survey participants. Quite a few households experienced a problem in receiving their advance mailing or diary packets in a timely manner. Thus, the mail-time was increased from 7 days (pilot) to 10 days (full survey).

Our quality control checks found that the CATI programs had few bugs that resulted in problem data. Most of the data that required "cleaning" during processing were the result of interviewer errors that could be dealt in training. We did not achieve 100% geocoding of home addresses or primary work addresses due primarily to missing addresses. Modifications to the diary instrument (see below) will also be effective in reducing missing or insufficient addresses. We did come close to the 90% requirement for trip ends. The geocoding outcomes were virtually identical for GDT and TBM coverages.

Early in the instrument design process, NuStats developed a place-based diary in which respondents were asked to record information about travel to each location as a separate page in the diary. The one-day diary contained 14 pages and the two-day diary contained 26 pages. This diary design was similar to those NuStats had designed for use in other travel surveys, including those conducted in the greater New

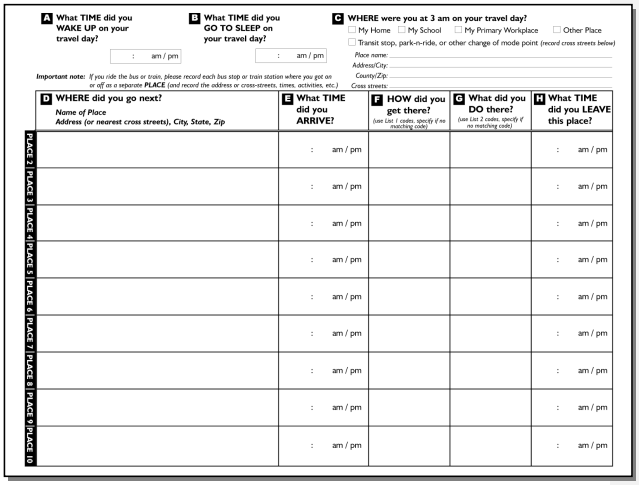
York and Northern New Jersey metropolitan areas, Seattle (WA) metropolitan area, Columbus (OH) metropolitan area, and Sacramento (CA) metropolitan area. Upon review, members of the Travel Survey Committee thought the diary format would be confusing to respondents. The concept of "place-rostering" surfaced. This diary format resembled a travel log in which each place was a "row" in the log. NuStats had used similar place-rostering formats for travel surveys in Nashville and Knoxville, Tennessee. (See Figure 1 and Figure 2 for examples of the two diary formats.)

FIGURE 1



ONE PLACE PER PAGE DIARY FORMAT

FIGURE 2



PLACE-ROSTERING LOG FORMAT

Quantitative and qualitative research methods were used to determine whether these objectives were met. Sampled households were flagged as to whether the "place-rostering" or "one place per page" diary format was used. Statistical analyses of participation rates and trip information were done comparing these two groups of households. In addition, mini-focus groups (3-4 participants) were conducted in mid-January 2001.

Our evaluation indicated that the Place-Rostering format outperformed the One Place per Page format. The roster format appeared to capture more trips, capture more information on transit trips, be used more frequently during the travel day, and also to retain more single, mobile households. For this reason, we proposed to use a modified roster format in the actual survey. Our recommendation was to develop a modified Place-Rostering format that provides space for more detailed address information. The diary "checklist" and the "example" were placed in more prominent, up-front locations in the diary. Most importantly, we clearly distinguished between trips and stops in the diary materials so that respondents comprehend the importance of recording exact address information for their incidental stops and other short trip segments.

The pilot also assessed the efficiency of intensive telephone screening for allocating households to the mode augment samples. It was found that a two-stage sampling procedure would be necessary. Phase one would be the RDD telephone sample to generate the local and community bus augment. This method would generate some households that will be eligible for allocation to the other target modes. Intercept methods would be used in Phase two for all other target modes.

DATA COLLECTION & PROCESSING

The SCAG Household Travel Survey was conducted during Spring 2001, Fall 2001, and Spring 2002. Cleaning, geocoding, and data delivery was done on a continuous basis. Data collection followed a six-stage survey process that included: pre-notification, recruitment, placement of materials, retrieval, processing and geocoding. In this section, outcomes are provided for recruitment and retrieval interviews.

THE RECRUITMENT QUESTIONNAIRE AND INTERVIEW

The recruitment interview was administered using a computer-assisted telephone-interviewing (CATI) program (see Appendix D for the recruitment script). These interviews were conducted in English and Spanish. Each household was telephoned by an interviewer to determine if they qualified for the study. The respondent was then asked (on behalf of the entire household) to participate in the study. If the respondent agreed, demographic information was collected from the household including income, household size, vehicle ownership, and other household characteristics. In addition, demographic characteristics were obtained for each member of the household such as age, gender, employment and school status. The average interview length was XX.xx minutes, obviously increasing as household size increased.

TABLE X: AVERAGE RECRUITMENT INTERVIEW LENGTH

Household Size	Count	Mean
1		
2		
3		
4		
Total		

In total, 27,476 households were recruited to participate in the Main Survey. Each recruited household was notified that it would receive a package in the mail that included a travel diary for each member in the household. This effort represented a recruitment response rate of 38 percent. Due to contractual issues, dialing was interrupted several times during the course of this project. We believe this had a deleterious impact on the recruitment rate. About 16 percent of all contacted households refused to participate in the study. The response rates were calculated under standards of the Council of American Survey Research Organizations (CASRO). These were derived by dividing the number of households that agreed to participate by the sum of the total number of "eligible" households and a portion of the households for whom "eligibility" was unknown. These response rate formulas are shown below.

$$RR = \left(\frac{a}{A+(C * ER)} \right)$$

Where,

RR is the response rate,
a is the number of completed surveys,
A is the number of eligible telephone numbers,
C is the number of eligibility unknown, and
ER is the eligibility rate.

$$RR = \frac{27,476}{48,096 + (120,364*.20)} = \frac{27,476}{48,096 + 24,072} = \frac{27,476}{72,168} = 38\%$$

TRAVEL LOG

A total of 27,476 travel-log packages were mailed to recruited households in the Main Survey. These packages consisted of a cover letter, brochure, and one travel log for each member of the household. (See Appendix E for diary materials.) The travel log collected information about each trip made on the assigned travel day, including place name and address, time of travel, travel mode, and purpose.

RETRIEVAL QUESTIONNAIRE AND INTERVIEW

The night prior to the assigned travel day, interviewers attempted to contact each household with a reminder call. The purpose of the reminder call was threefold: to remind the household of the upcoming travel day, to ensure that the household had received the survey packet, and to answer any questions that the household members may have. In non-contact experiences, reminder call messages were left on answering machines.

The day following each household's assigned travel day, the household was contacted by telephone (or attempted to be contacted) to retrieve the travel information. (See Appendix F for the Retrieval Questionnaire.) The interviews were guided using CATI programs and conducted in both English and Spanish. The average interview lasted xx minutes.

In total, 17,775 households completed recruitment and retrieval activities. For most of these households, the information was collected within seven days of the assigned travel day. Most persons (57%) used their travel log to record their activities on their travel day. This is a retrieval rate of 65% (17,775 retrieved / 27,476 recruited). The overall response rate for the study is determined by multiplying the recruitment rate (38%) by the retrieval rate (65%). For this study, the response rate was 25%. This means that 25% of all eligible households contacted about participation in the Household Travel Survey completed all activities associated with the project. This rate was low compared to other comparable surveys undertaken by NuStats. The interruptions in data collection (starting and stopping) had a deleterious impact on the response rate.

DATA PROCESSING

In cooperation with SCAG, criteria were established that defined a completed record. The data contained in the *Final Report of Survey Results* represent only those records that meet the criteria for a completed record. These criteria included:

- Household size variable in Household file must equal the number of persons in the Person file for that household.
- Number of vehicles variable in Household file must equal the number of vehicles in the Vehicle file for that household.
- In trip/ place file, arrival time at a place is prior to departure time for that place.
- In trip / place file, there is a travel mode recorded for all activities involving travel.
- In trip/ place file, there is an activity /purpose code for all places.
- 100% of all home addresses are geocoded.
- 100% of all primary work (main job) addresses are geocoded.
- 90% of all non-home, non-work addresses are geocoded.
- For any household, no more than 60% of non-home, non-work addresses can be missing geocode.
- All persons in households with 1, 2, or 3 persons must provide information on activities.
- No more than 20% of the households with 4+ persons per county may be missing activities.

When these criteria were applied to the 17,771 households that completed the recruitment and retrieval portion of the survey, the total number of deliverable households decreased to 16,506 confirmed and completed households. These households represent 39,264 persons.

TABLE X: DELIVERED HOUSEHOLDS BY COUNTY OF RESIDENCE

County	Frequency	Percent
Imperial	911	5.5
Los Angeles County	6,982	42.3
Orange County	2,271	13.8
Riverside	2,286	13.8
San Bernardino	2,139	13.0
Ventura	1,917	11.6
TOTAL	16,506	100.0

Note: Data are neither weighted nor expanded.

TABLE X: PERSONS IN DELIVERED HOUSEHOLDS BY COUNTY OF RESIDENCE

County	Frequency	Percent
Imperial	2,606	6.6
Los Angeles County	15,543	39.6
Orange County	5,296	13.5
Riverside	5,978	15.2
San Bernardino	5,371	13.7
Ventura	4,470	11.4
TOTAL	39,264	100.0

Note: Data are neither weighted nor expanded.

TABLE X: DELIVERED HOUSEHOLDS BY SAMPLE TYPE

Sample Type	Frequency	Percent
Base	10,875	65.9
Caltrans	3,060	18.5
Weekend	2,416	14.6
RSA	155	0.9
TOTAL	16,506	100.0

Note: Data are neither weighted nor expanded.

GEOCODING

Geocoding was conducted using coverage files obtained from SCAG. Home addresses and trip-end addresses were geocoded subsequent to the retrieval interview. The retrieval interview collected multiple location information such as address, nearest landmark, nearest cross street or street intersection to facilitate geocoding. City name and zip code were used to distinguish duplicated street names in different geographies. U.S. Postal Office Standard Address Format, which matched the address style of the street network reference database, was used to record address information.

Out of the xxxx addresses that were recorded by households as "traveled to", and were within the study area, xx percent were successfully matched to some level of geography. xx percent were matched to an X/Y coordinate, less than x percent were matched to a zip code centroid, and x percent were matched to a city centroid. Table XX presents geocode match rates by location type. As shown, the school addresses had the lowest overall match rate at x percent.

TABLE X: GEOCODING MATCH RATES

Address Type	Frequency	Percent
Home		
Work		
School		
Trip Ends		
Total		100.0

DATA FILE CREATION

After completion of data collection and data editing tasks, the Main Survey data were contained in seven files.

1. ***Household data file*** – the household is the unit of analysis, with 16,506 records. Contains data elements relating to household demographics such as household size, vehicles available to household and household income.

2. ***Person data file*** – persons within households are the units of analysis, with xxxxx records. Contains data elements relating persons, such as age, gender, work and school status.

3. ***Vehicle data file*** – vehicles owned by households are the units of analysis, with xxxxx records. Contains information relating to vehicles, such as make, model, and year.

4. ***Place data file –*** represents a particular geographic location as the unit of analysis with xxxx records. Contains information relating to travel, such as locations, purpose, mode, and time of travel.

5. ***Unlinked Trip data file*** – represents trips made by persons within households with an origin and destination for each row of data, with xxxx records. Contains information relating to travel, such as locations, purpose, mode, and time of travel.

6. ***Linked Trip data file*** – same as above, but trips have been linked to remove intermediate activities (such as change mode and serve passenger) thought to be secondary to the subsequent trip. File has xxxx records.

7. ***Location data file*** – all locations pertinent to households and trips made by persons within households, with xxxx records. Contains a location number that links to trip, person and household files.

All data files contain certain variables, such as sample number (unique number assigned to each household), and the weight variables "finwgt" and "expwgt".

DATA ITEM COMPLETION RATES

Table xx presents completion rates for the most important non-geographic variables. As shown, these rates are excellent.

TABLE XX: ITEM COMPLETION RATES

Variable	Frequency	Percent
Household Data		
Household Size	16,506	100.0
Vehicles Available	16,506	100.0
Bikes Available	16,502	100.0
Residence Type	16,486	99.9
Own/Rent Status	16,462	99.7
Languages Spoken in Home	16,500	100.0
Household Income	14,493	87.8
Person Data		
Gender	39,054	99.5
Age	38,850	98.9
Driver License (age 16+)	30,609	99.6
Employment Status (age 16+)	30,743	99.3
Number of Jobs (employed)	18,959	99.6
Business Type, Main Job (employed)	17267	90.7
Occupation, Main Job (employed)	17,063	89.6
Total Hours Worked Main Job (employed)	16,895	88.8
Usual Mode to Main Job (employed)	15,490	87.5
Type of Non-Employment (not employed)	11,664	97.8
Ethnicity	36,576	93.1
Disability Status	36,925	94.0
Educational Attainment	36,519	96.5
Student Status	39,085	99.7
Toll Used (base)	24,873	96.2
Carpool Lane Used (base)	24,934	96.2
Trip Data		
Arrival Time		
Departure Time		
Trip Purpose		
Activity (*Both origins and destinations)		
Mode		

Vehicle Data		
Year	29,015	97.4
Make	29,278	98.2
Model	29,803	100.0
Body type	29,362	98.5
Fuel Type	29,537	99.1
Year Acquired	28,147	94.4

APPENDICES

The Appendices section contains the following:

A. Work Plan,

B. Technical Memorandum on Sampling,

C. Technical Memorandum on Weighting and Expansion

D. Recruitment Questionnaire

E. Diary Package Materials

F. Data Retrieval Script

APPENDIX A – WORK PLAN

The NuStats team has developed a work plan for accomplishing the tasks associated with the Year 2000 Post-Census Regional Travel Survey for the Southern California Association of Governments (SCAG). The purpose of the project is to collect regional travel data that will support the estimation of statistical models of travel demand and mode choice and support analyses of travel characteristics of the region's residents. To achieve these objectives, a nine-task work scope was developed.

1. Design the survey instruments and data collection procedures
2. Finalize the sampling design
3. Train interviewers and implement quality control
4. Conduct and analyze the pre-test of all survey procedures
5. Revise and finalize the survey instruments and procedures
6. Conduct the main survey
7. Clean and geocode the data on a continuous basis
8. Weight the data set
9. Deliver the final data set, analyses, and final report.

Each of these tasks is discussed below, along with the task milestones, invoices, and meetings planned.

Task 1. Project Work Program Plan and Management Plan

The purpose of this task is to finalize the work program plan and the management plan. Technical Memorandum #1 (Work Program Plan) tracks assignments, costs, schedules, and survey milestones. It allows management of the activities to be completed so as to keep the entire project on schedule and is derived primarily from the RFP, NuStats proposal, and the Statement of Work.

Technical Memorandum #2 (Management Plan) will address issues of coordination among NuStats, its subcontractors, travel survey committee, SCAG project management staff and contractor. It will present the project organization chart and methods for monitoring and reporting of the project's progress and schedule. The role of each subcontractor will be presented, in addition to that of each key NuStats staff member.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Kick-off Meeting	NuStats/SCAG	8/23/00	8/23/00
Tech Memo #1: Project Work Plan	NuStats	9/25/00	10/10/00
Review Tech Memo #1	SCAG	10/11/00	10/25/00
Finalize Tech Memo #1	NuStats	10/26/00	10/31/00
Tech Memo #2: Project Management Plan	NuStats	9/25/00	10/10/00
Review Tech Memo #2	SCAG	10/11/00	10/25/00

Task 2. Survey Instrument Design and Procedures

This task involves the design and development of the survey instruments and the data collection methodology, including the CATI programming and all survey materials that will be mailed to the respondents. The SCAG and Caltrans survey instruments must be carefully coordinated for data compatibility.

Main Survey Materials

The main survey will consist of a 24-hour diary that is patterned after the one currently being used in the Caltrans survey. Diaries will be distributed to each household member. Travel days will be weekdays only and will be uniformly *distributed within each county* over each of the five weekdays. The list of data elements to be included in the survey has already been shared with SCAG and the Travel Survey Committee. The household survey materials will include:

- Screening interview script (household and respondent demographics)
- Introductory letter and study brochure
- Scheduling interview script (vehicle data and household member demographics)
- Diary, along with cover letter, travel day reminder sheet, and return envelope
- Retrieval interview script.

All materials will be translated into Spanish. In addition, NuStats will develop special enhancements to facilitate the participation of speakers of Asian languages. These include: translation of introductory letter, identification for interviewers of the Asian language phrases used to answer household telephone so that the incidence of various population groups in the sample can be tracked, etc.

WEEKEND SURVEY MATERIALS

The weekend survey will use the materials developed for the main survey with the exception that the diary will cover a 48-hour period. The diary will be designed for two-day pairs Friday/Saturday and Sunday/ Monday.

MODE AUGMENT SURVEY MATERIALS

The materials for respondents in the mode augments may need slight modification in the screening interview scripts, depending on the sampling methodology (e.g., intensive telephone screening or intercept). The decision on the methodology will be made in conjunction with SCAG and the Travel Survey Committee.

GPS SURVEY MATERIALS

Survey materials for respondents in the GPS survey will be identical to those used in the main survey. Additional materials and / or programs associated with fulfilling the requirements of the GPS survey may be developed as the GPS objectives and work plan dictate. Decisions on which additional materials will be required will be made in conjunction with SCAG and the Travel Survey Committee.

DATA COLLECTION PROCEDURES

NuStats will conduct the SCAG data collection in 15 continuous stages as noted in Table 1. A system of customized software programs (our continuous data flow [CDF] system) provides the quality checks at each stage that identify data records that do not meet project standards so that these can be investigated. The priority is to keep data records flowing smoothly through the stages.

Table 1

SCAG SURVEY STAGES

Stage	Day	Stage Description
1	1	Sample Generation
2	2	Geocode Home Addresses
3	3	Screening interview to collect household demographics and demographics on respondent (telephone)
4	4	Introductory letter and project brochure mailed
5	8	Geocode respondent's work and school addresses
6	9	Scheduling interview to collect vehicle information and demographics of other household members (telephone, mail, or Internet)
7	11	Diary package mailed to household that contains a personal diary for each household member
8	16	Contact with all households to remind the household of travel day
9	18	Travel day
10	19-24	Retrieval interview to collect travel data from each household member (telephone)
11	19-24	Field edits to check retrieved travel data for completeness
12	19-24	Data processing to prepare data for geocoding and edit checks
13	19-28	Geocoding of trip ends
14	29	Data quality checks
15	30	Process complete – data ready for delivery

The CDF system will incorporate the specific checks required by the SCAG project. These checks include ensuring that the following data items have no missing data:

- Number of persons in the household
- Number of vehicle owned or used regularly by household members
- Type of housing unit
- Start and end times for all activities in each diary
- Mode of travel for all activities involving travel
- Purpose for all activities
- Purpose for all trips if a place-based trip diary is used.

The following criteria will apply to geocoding:

- 100% geocoding of home address to x- and y- coordinates
- 100% geocoding of primary work (main job) address to x- and y- coordinates
- 90% geocoding of all non-home, non-work activity locations to x- and y-coordinates[3]

[3] Except if there is a missing street or street segment in the SCAG –provided street files.

- All non-home, non-work activity locations that cannot be coded to x- and y- coordinates will be coded to census tract.

In addition, the following criteria will apply on a per household level:

- No household should have more than 60 percent of non-home, non-work activity location that are not geocoded
- For households with one, two, or three persons, all persons (regardless of age) must provide completed information on all activities for the diary period
- For households with four or more persons, up to 20 percent of the households per county may be missing activities from one household member.

SCAG will review data in a timely fashion (with 2 weeks of biweekly data deliveries) to make a final determination on whether a household is considered a complete.

NuStats intends to develop a "customer friendly" survey approach that not only will give respondents multiple ways of providing data (telephone, mail, Internet) but also will provide information to respondents about the project in multiple ways (telephone, mail, internet).

NuStats will use *INFO ZERO UN* CATI software to conduct the telephone interviews. This software has advanced functions for field management, including importation of sample data, editing respondent records, monitoring the phone-room (remote online monitoring), frequency lists, respondent record lists, call-back assignment and lists, and productivity reports. The functions for sample management and tracking will be particularly useful for the SCAG project. These include real-time management of: multiple sample strata cells, replicates, and scheduling of call-backs on different times of day and days of week.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Travel Survey Committee Meeting	NuStats	9/20/00	9/20/00
Draft Survey Materials	NuStats	9/25/00	10/12/00
Review Survey Materials	SCAG	10/13/00	10/25/00
Review Survey materials	Travel Survey Committee	10/13/00	10/25/00
Finalize Survey Materials	NuStats	10/26/00	11/6/00
Travel Survey Committee Meeting	NuStats	10/25/00	10/25/00

Task 3. Sample Design

The purpose of this task is to finalize the sample design presented by SCAG in the RFP to ensure that it will meet the following objectives:

- Produce statistically adequate observations at a geographic level that meets the modeling and administrative objectives of the agency;
- Produce data depicting the diverse travel patterns and mode usage in the region;
- Minimize selection bias across subgroups in the population, particularly those that are more difficult to reach – high income households, the very poor, mobile persons such as renters, and others; and
- Maximize participation rates overall.

The SCAG request for proposal identified five distinct sample types. The budget assumed the following distribution of households among these sample types:

Type 1: Main sample (stratified random sample from 6-county area) 12,000

Type 2:	RSA sample (enrichment sample at the RSA level)	2,200
Type 3:	Weekend sample (proportional distribution across 6 counties)	3,100
Type 4:	GPS sample (proportional distribution across 6 counties)	500
Type 5:	Mode Augment sample (enrichment sample of transit users)	6,000

SCAG will also receive data from 3,000 households in the 6-county region that will be completed as part of the Caltrans Statewide Survey. Of the 6,000 households comprising the Mode Augment sample, 4,000 will be from Los Angeles County.

NuStats will prepare a technical memorandum that documents the process of selecting and preparing the sample for each sample type. This includes the determination of universe sizes for each sample type (control and target) and the establishment of stratum quotas. This memorandum will also contain documentation of any disproportionate sampling ratios, which will be used for subsequent expansion of the data.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #3: Sampling Design	NuStats	9/25/00	11/14/00
Review Tech Memo #3	SCAG	11/13/00	11/20/00
Review Tech Memo #3	Travel Survey Committee	11/13/00	11/20/00
Final Tech Memo #3	NuStats	11/21/00	11/30/00

Task 4. Interviewer Training and Quality Assurance

The purpose of this task is to ensure that the data delivered to SCAG is of the highest quality. In this task, NuStats will concentrate on the following five areas when explaining NuStats' techniques for interviewer training and quality assurance; survey staff, office procedures, training of survey staff, quality assurance, and data management.

Survey Staff

NuStats is proposing a task-based approach for the study team. Under this structure, the project will have a principal who is ultimately responsible for ensuring the rigor, quality, and integrity of the data and a project manager who is responsible for ensuring that work done on all tasks is done on time, on budget, and without any surprises. NuStats will perform all the data collection by its own staff in the Austin Survey Operations Center and in the newly developed Survey Lab at California State University at Northridge. NuStats will apply the same oversight and training in both locations and will generate an integrated status and diagnostic reporting system. Prior to the pretest ("dry run"), NuStats will prepare an Interviewer Training Manual for SCAG review.

Office Procedures

NuStats has designated staff for monitoring surveyors and is also prepared to have off-site or remote monitoring take place via the telephone. Remote monitoring can take place anytime with prior notice to a data collection supervisor.

NuStats will randomly select ten percent of completed interviews and verify them for accuracy of the number or persons in the household; the number of household vehicles; the home address; the work place and industry of employed persons; and activities performed by household members during the diary period. Documentation of this verification will be presented with the delivery of the final data set.

Training of Survey Staff

The Interviewer Training Manual will detail the training procedures that should be used to prepare interviewer personnel for the project. 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 this plan. The project specific training, as will be documented in the training plan, begins with an overview of the project, its objectives, glossary of terms, and the end product. That is followed by a detailed "local knowledge" course led by the Geocoding Manager. This course will include a Power Point presentation with specific tables and maps showing political geography, geographic terrain, transportation networks, and major trip attractions. The initial training sessions will be given in-person, video recorded, and added to the Power Point presentation. Each interviewer added to the project after the initial training will be required to view this presentation as part of training prior to each stage of collection. The third section of training deals with the intricacies of the survey instruments themselves, with separate training for each survey instrument involved. Mock interviews are used 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 will address common questions and objections. After passing a project-specific test, interviewers begin to work on the project and are monitored frequently and receive specialized one-on-one training with supervisors and the project manager. Project Trainers regularly debrief with interviewers and other project staff to keep everyone informed.

An additional training component will be included for NuStats' bilingual staff. This component will provide Spanish-speaking interviewers with the background on the Spanish-speaking population, additional review and practice time with the Spanish instruments and documentation, and key terms and phrases to ensure proper translation of the technical terms required to properly convey the study objectives.

This plan will also describe how the interviewers will be supervised during the data collection process. Monitoring criteria and standards will be detailed, as well as procedures for providing feedback and additional training for interviewers. Data checks will be specified to provide insight into the data quality and sufficiency of address information.

The initial plan will be developed and reviewed by SCAG prior to the pretest. Its use during the pretest will be evaluated and updated prior to the start of the full study. After the pretest, all interviewers will receive training for the full study, including those who were involved with the pretest. Potential problems and important issues will be clarified and emphasized during each training session.

Quality Assurance

A Quality Control Plan will be produced for SCAG review that details the steps taken to assure that accurate travel data will be delivered to SCAG. This plan will specify the duties and responsibilities of all organizations involved in the project, define project milestones, and establish appropriate review cycles. The plan will also outline the specific steps that will be taken in order to reduce or eliminate survey bias from sample selection, item non-response, and unit non-response. The following elements will be addressed: sample management, data collection procedures, data processing procedures, geocoding procedures, and survey results reporting procedures.

Geocoding Quality Control

NuStats will prepare a geocoding quality control plan for SCAG review. This plan will describe the resources to be used in geocoding, the expected turn-around for geocoding once data is collected, the specific edit checks, final cleaning, and delivery preparation.

A data transmittal document will be attached to each data delivery file. The memo will include the number of cases in each file/table, a statement that quality checks have been performed, geocoding match rates, and other necessary documentation on the quality of the data.

Data Management Plan

In order to ensure the highest quality of data in the resulting data set, a comprehensive reporting program is required to track the status of each household on a daily basis, identify problems quickly and easily, and ensure that key variables and indicators are reviewed on a daily basis. The Data Management Plan that is prepared for SCAG will document the various stages of data management, the requirements that would allow a completed survey to proceed to the next stage, and the process for correcting data that fail to meet the requirements and therefore are prevented from moving forward. All reports (including those for tracking daily movement by sample quota, travel day, and geocoding progress) will be contained in this plan as well. In addition to the flow of the sample, the data management plan will also detail the required elements of the computer-aided telephone interviewing (CATI) survey instruments, including all data "skips" and instructions for how they are programmed. Sample management instructions will be detailed, both for release and tracking of sample as well as the scheduling of callbacks in the computer-aided system. In addition, specification of data to be maintained for households that do not participate in the study or drop out at some point during the study will be outlined, as well as information on how this data can be used to enhance response rate and sample bias analyses.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #4: Interviewer Training Manual	NuStats	10/12/00	11/10/00
Review Tech Memo #4	SCAG	11/13/00	11/17/00
Final Tech Memo #4	NuStats	11/18/00	11/23/00
Tech Memo #5: Quality Assurance Plan	NuStats	10/16/00	11/13/00
Review Tech Memo #5	SCAG	11/14/00	11/17/00
Final Tech Memo #5	NuStats	11/17/00	11/21/00
Tech Memo #6: Geocoding Quality Plan	NuStats	10/12/00	11/10/00
Review Tech Memo #6	SCAG	11/13/00	11/17/00
Final Tech Memo #6	NuStats	11/18/00	11/23/00
Tech Memo #7: Data Management Plan	NuStats	10/16/00	11/16/00
Review Tech Memo #7	SCAG	11/17/00	11/27/00
Final Tech Memo #7	NuStats	11/28/00	11/30/00

Task 5. Conduct and Analyze Pretest

The purpose of this task is to pretest the survey instruments and procedures. It will evaluate the process for computerized sample management under actual field conditions, including household recruitment in a computer-assisted telephone interviewing environment, activity and trip data collection, geocoding, and data processing. Survey specialists will be trained in all procedures relating to the study, which will also be documented in a training manual. The objective of the pretest is to uncover any problems in the data collection forms, instructions for respondents and interviewers, and other materials. It will also allow an opportunity to assess interviewer performance and explore local levels of respondent cooperation and response rates.

In order to ensure that the pretest fully meets SCAG objectives, NuStats will produce a pretest plan that will present the methods and finalize the schedule for pretesting among the main survey sample, mode augment samples, and the GPS sample.

We will focus pretest objectives on refining our methods and materials to be "respondent friendly" to two key sample subgroups – Spanish language speakers and speakers of Asian languages. While materials will only be translated into Spanish, there will be special enhancements to our methods to facilitate the participation in the survey of Asian-language speakers.

The pretest's specific production goal will be to collect complete data from approximately 250 study area households distributed among the counties in the survey area. The proposed distribution of the 250 households in the pilot sample is as follows:

- 185 households, Main Sample
- 25 households, Weekend Sample
- 40 households, Mode Augments.

Technical Memo #8: Pretest Plan will provide more detail on the specific aspects of the pilot test sample. This sample will include a few "planted" households of individuals associated with SCAG or the Travel Survey Committee, who can provide additional feedback about their experience as respondents.

There will also be a pretest of methods to identify households in the mode-augment sample. There are two possible methods for selecting our samples: telephone screening and intercept interviews. The pretest criteria will include objective methods for assessing the results of these two methods, to identify the one will be employed in the full survey. The current budget assumes telephone screening; however, further discussions with the Travel Survey Committee have indicated that this method may not be feasible for some of the mode augments (e.g., park-n-ride to bus or rail).

While NuStats original budget included funding for focus groups that would have taken place during the pretest, our current schedule and pretest objectives do not lend themselves easily to this mode of evaluation. NuStats plans to conduct a separate GPS pilot test in early 2001. In this pilot test, we will test the effectiveness of three different types of GPS units: the Air Resources Board units, units proposed by Batelle in our proposal, and electronic travel logs. The GPS portion of the actual survey will be used to estimate underreporting of trips. These data will be used to calculate objective parameters for the magnitude of trip underreporting, uncover types (location types, trip purpose, activity patterns) of unreported trips, describe demographics of populations with the highest rates of underreported trips, and calculate adjustment factors that "could" be applied to trip rates. The GPS pilot test will determine:

- which units are most reliable in the field,
- to which units respondents are least resistant,
- which units provide data that is most efficiently analyzed, and
- which units best meet the GPS application objectives.

NuStats will use project funds that were allocated in our budget to the focus groups to conduct this separate GPS test. Any "focus group" funds that are remaining will be used to conduct follow-up mini-groups (focus group of 3-4 respondents who participated in the pretest) to do in-depth debriefings with them. We expect these mini-groups will be both in English and Spanish. NuStats will elaborate upon this evaluation in Tech Memo #8: Pretest Plan.

The pretest evaluation data will include the following:

- All questions and responses on all survey instruments to confirm skip patterns.
- Notes and/or transcripts of monitoring by supervisory staff, the project manager and task leaders. These are in addition to the scorecards described above.
- Tabulations of each question, showing response frequencies for all pretest respondents.
- Debriefing notes by project manager of all "planted" respondents (SCAG or related staff that will volunteer to be respondents in the pilot) as well as interviewing staff.

- Respondent debriefings conducted by the interviewing staff at the conclusion of the study. This will include both those that completed the study as well as those that did not.
- Debriefings with households who were "non-responding" during the pretest.

The above information will be the material on which the evaluation of the pretest will be based. The evaluation itself will probe a series of questions that we will resolve fully prior to recommending the initiation of the actual study. Prior to beginning the pretest, NuStats will submit a detailed list of questions we will ask and answer as part of the pretest evaluation. Among these questions, we anticipate including the following:

- Did the item elicit the range of responses expected? Where any responses out of range?
- Did the interviewer need to repeat the 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, monitoring staff, and respondent plants 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 hit 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 time lines met? What adjustments to the data flow process are required?

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #8a: Pretest Plan	NuStats	10/18/00	11/7/00
Review Tech Memo #8a	SCAG	11/8/00	11/13/00
Final Tech Memo #8a	NuStats	11/14/00	11/15/00
Identify "plants"	SCAG	11/1/00	11/15/00
Conduct Pretest:			
Generate sample	NuStats	11/15/00	11/16/00
Recruitment Process	NuStats	11/17/00	12/12/00
Assigned Travel Days	respondents	12/12/00	12/19/00
Data Processing	NuStats	11/17/00	1/8/00
Mini-Groups	NuStats	1/4/00	1/5/00
Tech Memo #9a: Pretest Results	NuStats	11/20/00	1/12/01
Review Tech Memo #9a	SCAG	1/13/01	1/23/01
Final Tech Memo #9a	NuStats	1/25/01	1/31/01
Pretest Data Set	NuStats	11/15/00	1/12/01
Travel Survey Committee & Peer Review Meeting	NuStats	1/24/01	1/24/01
Tech Memo #8b: GPS Pretest Plan	NuStats	11/6/00	11/20/00
Review Tech Memo #8b	SCAG	11/21/00	11/28/00
Final Tech Memo #8b	NuStats	11/29/00	11/30/00
Conduct GPS Pretest:			

Recruitment Process	NuStats	1/29/01	2/16/01
Assigned Travel Days	respondents	2/19/01	12/28/01
Data Processing	NuStats	2/26/01	3/15/01
Tech Memo #9b: Pretest Results	NuStats	2/19/01	3/19/01
Review Tech Memo #9b	SCAG	3/20/01	3/27/01
Final Tech Memo #9b	NuStats	3/28/01	4/4/01
Travel Survey Committee & Peer Review Meeting	NuStats	3/28/01	3/28/01

Task 6. Finalize Survey Design and Procedures

The purpose of this task is to finalize all changes recommended from the pilot survey that have been approved by SCAG. Final versions of all instruments (diaries, forms, scripts, letters, and other relevant materials) will be prepared in this task.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Produce final survey instruments	NuStats	1/15/00	1/31/01

Task 7. Conduct Spring 2001 Survey

The purpose of this task is to collect complete and acceptable accounts of activities performed by households with travel days assigned during the time period of February 5, 2001 through June 15, 2001. A review of school calendars for Southern California indicated that schools generally end the second or third week in June. Travel days will not be assigned on major holidays. NuStats expects to complete approximately 10,000 households during the spring survey effort. Unless further refined in Technical Memo #3, the number of households of each sample type collected in the Spring are:

Type 1:	Main sample (stratified random sample from 6-county area)	6,000
Type 2:	RSA sample (enrichment sample at the RSA level)	0
Type 3:	Weekend sample (proportional distribution across 6 counties)	1,300
Type 4:	GPS sample (proportional distribution across 6 counties)	150
Type 5:	Mode Augment sample (enrichment sample of transit users)	2,500

There is approximately 1 month between the delivery of the pretest results and the pretest data file, and the start of travel days. To provide as much time as possible for making any necessary changes, NuStats intends to start the "screening and scheduling interviews" in early January to get a jumpstart on recruitment. We believe that the screening and scheduling interview scripts will few changes. We can take extra time, if needed, to make all necessary adjustments to the travel diary materials.

Under this task, NuStats will be responsible for conducting the survey, coding all survey responses, and correcting the resulting computer files as necessary. A high standard of data collection and coding will be upheld as documented in the Quality Assurance plans produced under Task 4. Prior to implementation of the survey, NuStats will prepare a technical memorandum that specifies the production plan for the spring.

Activities within the survey implementation sequence include:

1. Generate sample and geocode home address (listed)
2. Conduct screening interview
3. Geocode unlisted home address and work/school address (respondent)

4. Mail advance letter and brochure
5. Conduct scheduling interview
6. Geocode work/school addresses other household members
7. Place diary packet via mail
5. Make reminder call.
8. Conduct retrieval interview
9. Finalize geocoding of address data
8. Conduct follow-up calls to verify or collect missing data, if necessary.
9. Ask respondents to return survey documents in postage paid envelopes provided.

Progress reports on survey progress and data quality will be reported on nearly a real-time basis during and after data collection. During the data collection task, statistical reports summarizing participation rates, average trip rates and standard deviations of trip rates will be produced at least once a week. However, since NuStats' CDF process is based on continuous data processing methods, reports can be generated at any time. NuStats does not intend to execute a "blanket distribution" of incentives. Incentives, if distributed, will be distributed during Fall survey execution. Incentives will be targeted to categories of households for which we attained a lower than expected response rates during the Spring survey execution. By targeting incentives in this way, NuStats can distribute slightly higher monetary incentives (e.g., $5 per person in household); thereby, eliciting stronger effects on participation. NuStats also intends to offer a "menu" of incentives. It has been our experience that different types of incentives appeal to different types of people. At a minimum, our menu will include a cash incentive and the opportunity to donate the cash to a charitable organization. NuStats will include a field in the database that indicates whether a given household did or did not receive a monetary incentive.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #10: Spring Production Plan	NuStats	1/15/01	1/19/01
Review Tech Memo #10	SCAG	1/22/01	1/26/01
Final Tech Memo #10	NuStats	1/29/01	1/31/01
Weekly progress reports	NuStats	2/15/01	6/29/01

Task 8. Interim Delivery of Data

The purpose of this task is to prepare and deliver interim data sets that have been geocoded and meet the project quality standards in terms of logic, household consistency and other measures. NuStats intends to do this via monthly data deliveries beginning in March 2001. NuStats assumes that SCAG will review the interim delivery of households for completeness and provide critical feedback on data quality at regular intervals. NuStats would then deliver a final interim data set at the end of July 2001.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Monthly data deliveries	NuStats	3/16/01	6/29/01
Data file reviews	SCAG	3/19/01	7/17/01
Final Interim Data File	NuStats	6/20/01	7/30/01
Data file review	SCAG	7/30/01	8/31/01
Tech Memo #11: Mid-point analysis	NuStats	6/30/01	7/30/01
Review Tech Memo #11	SCAG	8/1/01	8/31/01

Final Tech Memo #11	NuStats	9/1/01	9/07/01
Travel Survey Committee Meeting	NuStats	9/19/01	9/19/01

Task 9. Conduct Fall Survey

During summer 2001 and in close consultation with SCAG, NuStats will perform a comprehensive review of the distribution and key characteristics of the spring data households in each of the sample types. The actual review will be an on-going process during spring data collection; however, the end of season comprehensive review will be geared to establish any key priorities or adjustments for the Fall survey. Particular attention will be given to the following (and others as needed):

- Precise distribution in all stratification cells (household size by vehicles in household; RSA; mode-specific) and measurement of any gaps from the completion goal for the Spring.
- Demographic profiles to identify any significant discrepancies between actual data and expected distribution for major demographic variables (ethnicity, age).
- Participation rates overall, including recruitment and final retrieval rate patterns, with a special emphasis on comparative retrieval rates for target groups (renters, large households, non-English speakers).
- Data quality indicators, as defined in Tasks 4.

Based on actual performance during the Spring and on the assessment of the data, NuStats plans on completing the remaining 14,000 (approximate) households in the Fall, with all data collected by the end of 2001. NuStats will implement a Fall schedule that has screening and scheduling interviews beginning in early August and travel days beginning the week of September 10. A review of the school calendars for the region indicated that schools generally begin September 6. Unless further refined in Technical Memo #3, the number of households of each sample type collected in the Fall are:

Type 1:	Main sample (stratified random sample from 6-county area)	6,000
Type 2:	RSA sample (enrichment sample at the RSA level)	2,200
Type 3:	Weekend sample (proportional distribution across 6 counties)	1,800
Type 4:	GPS sample (proportional distribution across 6 counties)	350
Type 5:	Mode Augment sample (enrichment sample of transit users)	3,500

These households will be distributed among the main sample, the RSA augmenting over-samples, mode-specific augments, and GPS samples in such a way that the final Fall/Spring totals are on the targets set in the study design. NuStats intends to deliver a slight overage, which will be used to compensate for any problems regarding final quality assurance passage or completion criteria standards.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #12: Fall Production Plan	NuStats	8/15/01	8/29/01
Review Tech Memo #12	SCAG	8/30/01	9/7/01
Final Tech Memo #12	NuStats	9/10/01	9/14/01
Weekly progress reports	NuStats	9/4/01	12/14/01

Task 10. Delivery of Final Data Set

The purpose of this task is to produce a clean, final deliverable data set at the completion of the project. NuStats will be delivering ready-to-use data on a continuous schedule as indicated in Task 8. These periodic, interim deliveries will be constantly reviewed by SCAG so that the deliveries will fully meet all SCAG

expectations, as communicated to NuStats. As with all the partial deliveries during the project, the final data set will be a group of relational files, totally linked by identification numbers uniquely associated with each household. In addition to the four principal files (household, vehicle, person, activity or trip), NuStats will maintain and deliver a place file that will have an unduplicated list of all locations reflected in the entire survey. Each unique place will have a unique numeric identifier and a code for whether it is a home, work, school, or other location. The value of the place file is that it allows for several useful quality control tests of geocoding and is a useful reference resource for subsequent analysis. The activity/trip file will use the unique place identifier to document where the trip or activity took place. The unique place identifier would be in addition to the required set of descriptors and geocoding information that is provided for in the data structure documentation elsewhere.

Each of the data files will have calculated summary statistics reflecting frequencies of observations in subordinate files. For example, the household file will have a vehicle inventory count of the number of vehicles in the household; this count will be consistent with the number of vehicle records in the vehicle file that are associated with that household. Similar summary counts will be generated for number of total persons, number of licensed drivers, number of employed persons, number of activities/trips, and other useful indicators of data consistency.

Each data file will have a detailed data structure in SPSS format (or a similar statistical software) and in a tab delimited ASCII format. We anticipate that both interim deliveries of the data and the final data set will be delivered in CD-ROM medium. In addition to the data file documentation, NuStats will produce a Users Manual that can guide prospective users with varying levels of expertise to access and use the data set.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Monthly data deliveries	NuStats	9/28/01	11/30/01
Data file reviews	SCAG	10/15/01	12/14/01
Final Data File	NuStats	11/30/01	1/11/02
Data file review	SCAG	12/14/01	2/28/02

Task 11. Analysis of Results

Under this task, NuStats will analyze and document the quality of the survey data. The task of reporting is based upon a mutual end view or visualization of the deliverables. NuStats will produce an Analysis Plan for SCAG review early in the Fall 2001. This plan will provide detailed specifications of the project and the outputs that will be provided. Currently, NuStats assumes that its analysis will be divided into three main parts.

1. The first part is a final compilation of on-going reporting of participation rates, trip rates and standard deviations, consistency checks, geocoding rates, plausibility checks, and other indicators of data quality.
2. The second component of analysis is related to field work and participation patterns. A file that provides all attempted calls to each household will be provided and analyzed in a brief technical memorandum. This file will contain all completed households, as well as any that dropped out or became unusable for any reason. The summary of this analysis will include exact item response rates, responses by stratification cell, and additional variables that may be agreed to with SCAG.

Finally, the reporting will provide documentation on the following:

- All edit checking and the subsequent correction steps taken to reconcile data.
- Calculation of recruitment, retrieval, and overall response rates according to standards prescribed by the Council of American Survey Research Organizations.

- Analyses of non-response for both the sample estimates as compared to accepted population parameters and specific data items. Census and NPTS data will be used as resources.
- Flags that document data irregularities that result from respondent provided information (such as rounding of travel times) to guide the use of the data set in the modeling effort.

One of the more valuable components of the analysis of results is the compendium of tables to be generated for dissemination. The proposed approach from NuStats is to generate a comprehensive compendium of tables. The tables should be able to combine into a single comprehensive document and to be easily extracted as single tables for distribution to interested parties or for presentations.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #13: Analysis Plan	NuStats	10/16/01	10/26/01
Review Tech Memo #13	*SCAG*	10/29/01	11/9/01
Final Deliverable	NuStats	12/3/01	1/11/02

Task 12. Data Expansion

The data expansion plan will be incorporated into the sample design plan in Task 3. The expansion factors will include, at a minimum, all Type 1 and 2 households collected as part of this effort plus the 3,000 households to be collected by Caltrans. Because some key variables used for sample stratification will rely on data that are ten years old (from the 1990 US Census), it will be valuable to provide an expansion model that can be refreshed with 2000 US Census data when those become available. The basic objectives of the data expansion model are:

- Compensate for the differential sampling fractions applied to various strata (so that the final data observations are balanced to their true representation in the population); and,
- Provide target population numbers that reflect the full population statistics that the survey sample(s) represent.

The SCAG region data collection to be undertaken by Caltrans will also need to be considered in the expansion model. Assuming that the data are completed and available and that the characteristics of the Caltrans data set are known, the expansion weights generated by NuStats can take into consideration these 3,000 additional households.

NuStats anticipates separate expansion weights for the weekend data set. The expanded data set will provide population estimates for Saturday and Sunday combined, as well as for the collected day pairs and for a comparison of weekend peak (Saturday/Sunday) and weekend shoulder (Friday/Monday).

Expansion weights for the GPS data set will be somewhat more limited due to the relatively smaller sample size of the GPS sample. However, a set of weights, with appropriate limitations on their utility will be generated and provided to SCAG.

The most complicated weighting issue will involve the mode-specific augments. If the sampling approach used is the intensive screening method recommended by NuStats or the intercept approach indicated in the RFP (a decision to be made after the pilot survey), the weighting would be geared to reflect the alternative approaches.

A detailed Technical Memorandum with a set of relational spreadsheets for the development and calculation of the expansion weights and parameters will be delivered to SCAG. An early version of the Technical Memorandum and the worksheets will be made available for review by SCAG and for mutual discussion with

NuStats. The final version will thus reflect the full reality of the data set (with or without the Caltrans data) at the time that this Technical Memorandum is delivered.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #14: Expansion & Weight Documentation	NuStats	1/07/02	1/11/02
Review Tech Memo #14	*SCAG*	1/14/02	1/28/02
Final Tech Memo #14	NuStats	1/29/02	2/1/02

Task 13: Final Report

The purpose of this task is to prepare a final household survey report documenting the methodology and survey tabulations. At minimum, the NuStats final report will include the following sections.

- **General Final Report** - comprehensive document that presents the survey findings using the expanded data. The report will focus on measures of travel activity patterns estimated for the region and for county as well.
- **Methods and Implementation** – documentation of survey methods, interviewing outcomes, response rates, weighting, and survey instruments and materials.
- **Compendium of Tables** – tabular presentation of results by region and county of the study area. This would likely include as many as 150-200 tables.
- **GPS Analysis** – special analysis of the GPS efforts with both the in-vehicle approach and the generic approach of using new technology.

The draft versions of the final report(s) will be delivered to SCAG within two months after the final data collection is completed. Twenty paper copies will be provided to SCAG for the review, along with an electronic version in CD-ROM. After receipt of comments and requested changes from SCAG, NuStats will produce a final set of deliverables. This Final Report(s) will be delivered in one high-quality paper original for reproduction, fifty bound copies for dissemination and archiving, and an electronic version in Microsoft Office in a CD-ROM medium. Alternatively, or additionally, an electronic version in Acrobat format can be provided to SCAG.

NuStats anticipates a final meeting with the Regional Travel Survey Committee to present the final results and to provide a final review and retrospective on the project. In addition, at the discretion of SCAG, NuStats can make or assist in making any presentations of final outcomes and results to SCAG executive officers and boards, to the media, or to other groups.

Schedule and Milestones:

Milestone	Action By	Start	Complete
Tech Memo #15: Final Report Outline	NuStats	10/16/01	10/26/01
Review Tech Memo #15	SCAG	10/29/01	11/12/01
Draft Final Report	NuStats	1/15/02	2/15/02
Review Final Report	SCAG	2/18/02	3/14/02
Final Report	NuStats	3/18/02	3/29/02

Appendix B: Technical Memorandum on Sampling[4]

Introduction

The purpose of this memorandum is to provide specifications for our recommended final sampling plans for the SCAG suite[5] of population and mode augment studies. The recommended sampling plans are presented in two parts:

- a sample design for the *integrated population surveys*; and
- our recommended sampling designs for the multiple *mode augments*.

The integrated population surveys denote a collection of four surveys of the SCAG region. They are to be integrated in the sense that cases from each sample can be combined to produce a larger base from which to conduct analyses. Three of the four population surveys are used to examine *weekday* travel behavior among household members. They include:

- a base sample of households in the SCAG region;
- the Caltrans sample, representing the SCAG regional portion of a statewide Caltrans survey of households regarding travel behavior; and
- the Regional Statistical Area (RSA) augment, an oversample of subregional geographic areas in the SCAG region that will ensure the ability to produce nominal estimates from specific RSA subregions.

A fourth household survey component covers *weekend* travel behavior and is called the *weekend augment*. The weekend augment will not be combined with the other three samples (which focus on weekday travel).

All population surveys are to be conducted via telephone using Random Digit Dialed (RDD) sampling methods. In this memorandum we detail how that will be accomplished.

The *mode augment* samples represent a collection of seven mode specific samples of mode user households (and their residents) for the purpose of implementing mode specific statistical models of travel volume and behavior. The samples will focus on the following seven public transit modes:

- Local and community bus
- transit-express bus
- Commuter Rail (Metrolink)
- Metro (Red, Green and Blue lines)
- Auto Access to Express bus (including Park-n-Ride and Kiss-n-Ride)
- Auto Access to Rail (including Park-n-Ride and Kiss-n-Ride)
- Toll Road User.

The mode augments will employ a combination of RDD telephone sampling and intercept sampling to achieve the desired survey goals.

[4] Submitted on January 19, 2001 and Revised on April 18, 2001. This appendix only contains text pertaining to the population surveys. Documentation of the mode augment sampling approach appears as a separate document.
5 The SCAG suite of studies also includes a GPS pilot; the technical approach for that study appears as a separate document.

Below a separate section is devoted to each component of the SCAG research suite. These are organized as:

- Integrated Population Surveys
- Mode Augment Samples
- Analytic Weights.

Integrated Population Surveys

SCAG population surveys involve surveys of households and their residents in the six- county SCAG region. Separate samples are desired for:

- ***the base sample*;** the base sample represents the six-county SCAG region in proportion to population distributions across its constituent counties (i.e., Los Angeles; Orange; Riverside; Imperial; San Bernardino; Ventura)
- ***the Caltrans sample;*** the Caltrans sample denotes the six-county portion of the greater statewide Caltrans household survey that is being conducted with a field period that overlaps with the SCAG study; accordingly, an effort is being made to integrate the Caltrans into the SCAG in a way that achieves the overall Caltrans analytic goals and can concurrently enhance and complement the SCAG survey;
- ***the RSA augment;*** separate estimates are desired for specific subregions of the SCAG region; these subregions are called RSAs, and nominal samples are desired for all but a few specific RSAs; the objective of attaining the nominal representation gives rise to the need for an RSA Augment sample of households;
- ***the weekend sample;*** the base, Caltrans, and RSA augment samples of households are designed to measure weekday household travel behavior; the *weekend sample* denotes a separate metropolitan area sample of households that is designed to measure travel behavior on weekends.

Population Surveys – the Population of Inference

Generally, the population of inference for the population surveys comprises households and household residents in the six-county SCAG region (i.e., Los Angeles, Orange, Riverside, Imperial; San Bernardino; Ventura). Because a telephone survey will be used to recruit households into the sample, the survey population will be households with telephones in the SCAG region.

Population Surveys – Survey Objectives

The survey objectives for the population surveys call for specific targeted sample sizes of households for each sample type. The sample size targets are illustrated in Table 1. Columns A-D present the county specific targets for the base, Caltrans and RSA Augment samples, respectively. And Column D exhibits the cumulation of these samples by county. Overall, Column D shows that 17,217 households are to be sampled for the combined Base, Caltrans and RSA augment samples:

- The base sample accounts for 12,000
- the Caltrans sample involves 3,000 equally allocated across counties
- the RSA augment has differential sample sizes across counties in order to achieve the nominal representation of at least 200 per RSA (excepting four specific RSAs that are not a part of the RSA augment).

Column E shows the county specific sample size targets for the weekend sample. This sample is allocated roughly in proportion to the distribution of population across counties in the SCAG region.

Population Surveys – Recommended Design

The population surveys are to be conducted using telephone sampling for household selection and screening, followed by diaries mailed to eligible households and completed by all residents. We recommend the use of a highly stratified RDD telephone probability sample of households (with telephones).

The sampling frame for the stratified probability sample will include both listed and unlisted telephone numbers from working blocks of numbers that make up the six-county SCAG region. "Blocks" of numbers are defined as groups of

100 consecutive telephone numbers whose area code, exchange and stem (i.e., the last four digits) are identical except for the last (rightmost) two digits (which range in value from 00 to 99). "Working blocks" denote all blocks for which at least one telephone number in the block is a listed residential telephone number. Working blocks include both listed and unlisted telephone numbers, and working blocks collectively cover a high percentage of all residential numbers in the area. The sampling frame is defined as the collection of telephone numbers in all working blocks that cover the six-county SCAG region.

The probability sample of telephone numbers for the SCAG samples will be drawn using a highly stratified geographic design. First, the address information from the listed numbers in the frame will be used to create assignment rules for all numbers (listed and unlisted) based on their area code, exchange and stem. This will be used to construct 55 RSA strata that comprise the six-county geographic region. The assignment of telephone numbers to strata is not 100% accurate but should be satisfactory for stratification purposes. Additionally, telephone numbers will be stratified by *listed status* (i.e., listed; unlisted). Combining these two stratification variables yields 55 x 2 = 110 strata. These 110 strata will be the basis for a stratified probability sample of telephone numbers. Separate, independent samples will be drawn for: (1) the base sample; (2) the Caltrans sample; and (3) the RSA Augment sample. The respective samples will be drawn to effect the relative distributions shown in Table 1.

Using the recommended sample design, Table 2 presents the expected distribution of sample households across RSAs for the SCAG base sample, the Caltrans sample and the RSA Augment. These distributions aggregate precisely to the desired distributions by sample type and County shown in Table 1. Columns A and B in Table 2 show the expected distribution of households across RSAs for the Base sample and the Caltrans sample. Column D presents the expected distribution across RSAs of the RSA Augment. The last column, Column E, shows the resulting expected sample sizes by RSA and demonstrates that the recommended design is expected to attain the nominal desired sample sizes (i.e., a minimum of 200 households per RSA), excepting those RSAs for which a nominal representation is not desired.

The *Weekend Augment* sample will be drawn from the 110 strata so as to yield the expected HH distribution shown in Column E of Table 1. The sample of households is expected to total 3,100. The weekend augment sample will be uniformly distributed across two-day pairs (Friday-Saturday; Saturday-Sunday).

Population Survey Weights

From a finite population sampling theory perspective, analytic weights are needed to develop estimates of population parameters and more generally to draw inferences about the population that was sampled. Without the use of analytic weights, population estimates are subject to biases of unknown (possibly large) magnitude. It would be inappropriate, for instance, to treat the survey data as a simple random sample of households in the SCAG region, since unequal probability sampling (via the RSA augment, for instance) must be reflected in the construction of estimators, in the evaluation of statistical precision and in other statistical inferences (e.g., hypothesis testing). Weighting compensates for these "departures" from simple random sampling. Consequently, analytic weights will be developed for the population surveys.[6]

Table 1
Sample Targets for the SCAG Population Surveys

County	**(A) SCAG Base Sample**	**(B) Caltrans Sample**	**(C) RSA Augment**	**(D) (A+B+C) *Subtotal***	**(E) Weekend Augment**
Los Angeles	5,500	500	826	***6,826***	1,887
Orange	1,500	500	346	***2,346***	508
Riverside	1,500	500	492	***2,492***	254

[6] Documentation of Analytic Weights appears in this report as Appendix X.

San Bernardino	1,500	500	394	***2,394***	294
Ventura	1,500	500	159	***2,159***	137
Imperial	500	500	0	***1,000***	21
Total	***12,000***	***3,000***	***2,217***	***17,217***	***3,100***

TABLE 2
Expected Distribution of SCAG Sample

By Sample Type and RSA

RSA	COUNTY	(A) SCAG Base	(B) Caltrans Statewide	(C) (A+B) Combo	(D) RSA Augm	(E) (A+B+D) 1.1.1
1	Ventura	2	1	2	0	2
2	Ventura	383	128	511	0	511
3	Ventura	510	170	680	0	680
4	Ventura	305	102	406	0	406
5	Ventura	270	90	360	0	360
6	Ventura	31	10	41	159	200
7	Los Angeles	37	3	41	159	200
8	Los Angeles	98	9	107	93	200
9	Los Angeles	92	8	100	100	200
10	Los Angeles	81	7	88	112	200
11	Los Angeles	2	0	2	0	2
12	Los Angeles	466	42	509	0	509
13	Los Angeles	225	20	245	0	245
14	Los Angeles	183	17	199	0	199
15	Los Angeles	13	1	14	186	200
16	Los Angeles	269	24	293	0	293
17	Los Angeles	809	74	882	0	882
18	Los Angeles	382	35	417	0	417
19	Los Angeles	285	26	311	0	311
20	Los Angeles	337	31	368	0	368
21	Los Angeles	479	44	523	0	523
22	Los Angeles	392	36	428	0	428
23	Los Angeles	82	7	89	111	200
24	Los Angeles	329	30	359	0	359
25	Los Angeles	481	44	524	0	524
26	Los Angeles	336	31	367	0	367
27	Los Angeles	123	11	134	66	200
28	*San Bernardino*	553	184	737	0	737
29	San Bernardino	522	174	696	0	696
30	San Bernardino	59	20	78	122	200
31	San Bernardino	13	4	17	183	200
32	San Bernardino	260	87	346	0	346
33	San Bernardino	83	28	110	90	200
34	San Bernardino	11	4	14	0	14
35	Orange	97	32	129	71	200
36	Orange	120	40	160	40	200
37	Orange	231	77	307	0	307

38	Orange	221	74	295	0	295
39	Orange	156	52	208	0	208
40	Orange	158	53	211	0	211
41	Orange	95	32	126	74	200
42	Orange	244	81	325	0	325
43	Orange	108	36	145	55	200
44	Orange	70	23	93	107	200

TABLE 2
Expected Distribution of SCAG Sample

By Sample Type and RSA (cont.)

RSA	COUNTY	(A) SCAG Base	(B) Caltrans Statewide	(C) (A+B) Combo	(D) RSA Augm	(E) (A+B+D) 1.1.1
45	Riverside	75	25	100	100	200
46	Riverside	542	181	723	0	723
47	Riverside	137	46	182	18	200
48	Riverside	153	51	204	0	204
49	Riverside	142	47	189	11	200
50	Riverside	55	18	73	127	200
51	Riverside	13	4	18	0	18
52	Riverside	260	87	347	0	347
53	Riverside	89	30	118	82	200
54	Riverside	34	11	45	155	200
55	Imperial	500	500	1,000	0	1,000
	TOTAL	12,000	3,000	15,000	2,217	17,217

APPENDIX C – TECHNICAL MEMORANDUM ON WEIGHTING AND EXPANSION

This technical memo documents the methodology NuStats will use to weight and then expand the Post-Census Regional Travel Survey sample to represent the total population of households in the SCAG region. The issues addressed include: universe estimation, adjustments for probability of selection, post stratification adjustments, determination of final weights, and expansion factors. The memo focuses on the probability samples ~~ated~~ that comprised the largest portion of the regional survey effort. Because these sample surveys were conducted via telephone using Random Digit Dial (RDD) sampling methods, cases from each can be combined to produce a larger base from which to conduct analyses. The samples include:

- **Base Sample** of households in the SCAG region across its constituent counties (i.e., Los Angeles; Orange; Riverside; Imperial; San Bernardino; Ventura). This sample was used to collect 24-hour (weekday) diaries.
- **Caltrans Sample**, which comprised the six-county Southern California portion of the greater statewide Caltrans household travel survey. This sample was drawn from the same sampling frame as the base sample (above) at the same time and was conducted with a field period that overlapped with the SCAG study.
- **Regional Statistical Area (RSA) Augment**, which was an oversample of subregional geographic areas (RSAs) in the SCAG region with the intent to increase sample at the subregional level.[7] The RSA sample was drawn independently from the Base, Caltrans, and Weekend samples during the latter stages of data collection.
- **Weekend Sample** was designed to measure travel behavior on weekends using a 48-hour diary, whereas the Base, Caltrans, and RSA augment samples of households were designed to measure weekday household travel behavior. The weekend sample was drawn during the same "pull" as the Base and Caltrans samples. All weekend households recorded travel during a 48-hour period that spanned either Friday/Saturday or Sunday/Monday.

A fourth household survey component covered a **Mode User Augment**. It entailed the collection of seven independent samples of specific types of mode users for the purpose of implementing mode specific statistical models of travel volume and behavior. The original sample design recommended a combination of RDD telephone sampling and intercept sampling to achieve the desired survey goals. The RDD portion would have been a probability sample and subject to weighting and expansion. However, the RDD methodology was dropped during fieldwork due to cost considerations. Instead, the Mode Augment effort relied heavily on intercept sampling. This was supplemented by a small portion of telephone interviews that relied upon highly targeted listed sample. Therefore, the Mode Augment data file (3,050 households) was not subject to any weighting or expansion process since it was not based on a probability sample.

1. DOCUMENTATION AND VALIDATION OF ANALYTICAL WEIGHTS

From a finite population sampling theory perspective, analytic weights are needed to develop estimates of population parameters and more generally to draw inferences about the population that was sampled. Without the use of analytic weights, population estimates are subject to biases of unknown magnitude. It would be inappropriate, for instance, to treat the survey data as a simple random sample of households in the SCAG region, since unequal probability sampling must be reflected in the construction of estimators, in the

[7] This effort was curtailed during the fieldwork period due to contractual issues, and thus, resulted in far less oversampling at the subregional level than specified in the original sample design.

evaluation of statistical precision and in other statistical inferences. Weighting compensates for these "departures" from simple random sampling. Consequently, analytic weights were developed for the population surveys as documented in subsequent paragraphs. After these weights were applied to the final samples, an expansion factor was created to expand the survey data from the sample of households to the regional population of households they represent. The documentation for the expansion factor calculation and its application is provided in Section 2 of this memo.

A. Estimation of Survey Universe

The population of inference for the population surveys comprises households and household residents in the six-county SCAG region (i.e., Los Angeles, Orange, Riverside, Imperial; San Bernardino; Ventura). Through discussions with SCAG, it was decided to use 2000 Census data to estimate population parameters. Estimates in this memo are from Census 2000, Summary File 2. The universe of households in the SCAG region totals 5,386,491 occupied housing units.[8]

The inferential capabilities of the SCAG survey rests on probability sampling from a frame covering all members of the universe (in this case, occupied housing units). Because a telephone survey was used to recruit households into the sample, the survey population was households with telephones in the SCAG region. The sampling universe was total household telephone numbers, and the sampling frame was the collection of ten-digit telephone numbers from the listed working banks that fell into the SCAG region. 2000 Census data indicates that 1.6% of occupied housing units in the SCAG region are without telephones, and this percentage is a fair indication of the coverage bias introduced by the choice of survey population and sampling frame.

~~B.~~2. Probability of Selection or Sampling Weight

Formatted: Bullets and Numbering

A probability sample assigns known, nonzero chances of selection to every member of the survey population. Inequalities in the selection frame and procedures may create unequal selection probabilities. These are corrected with weights inverse to those probabilities.

Data collection targets by county that were identified by SCAG guided the selection of elements from the frame. The sample for the Base, Weekend, and Caltrans studies (hereafter, referred to cumulatively as the "Base" sample) was drawn at the same time, then divided into partitions based on the data collection targets. The sample was drawn at the county level, as shown in Table 1, and represents disproportionate sampling by county when considering the universe of occupied housing units at the county level. The sample for the RSA augment was drawn later in the study schedule, and thus, requires a separate probability of selection calculation.

Table 1: Data Collection Targets by County

County	Occupied Households	% of Occupied Households	Base Target	Weekend Target	Caltrans Target	Total Target	% of Total
Imperial	39,384	0.7	500	21	500	1,021	5.6
Los Angeles	3,133,774	58.2	5,500	1,887	500	7,887	43.6
Orange	935,287	17.4	1,500	508	500	2,508	13.8
Riverside	506,218	9.4	1,500	254	500	2,254	12.5
San Bernardino	528,594	9.8	1,500	294	500	2,294	12.7
Ventura	243,234	4.5	1,500	137	500	2,137	11.8
Total	5,386,491	100.0	12,000	3,100	3,000	18,100	100.0

Source: Table 1 of Technical Memorandum on Sampling, April 2001.

The probability of selection in the Base Sample is defined as the probability of being selected from the frame for the county of residence. The population parameter for this calculation (i.e., total household telephone numbers) was derived from Acxiom's InfoBase telephone directories. The formula is shown below, followed by Table 2, which shows the formula inputs and probability calculation for each county.

$P_1 = n_1 / N$

where

P_1= probability of selection
n_1 = the total HH telephone numbers selected
N = the total HH telephone numbers in the Sampling Universe.

Table 2: Probability of Selection (Sampling Weights) by County

County	Sampling Universe (*N*)	Sample Selected (n_1)	Probability of Selection (P_1)
Imperial	79,600	6,798	0.0854
Los Angeles	6,971,600	222,191	0.0318
Orange	1,572,700	82,875	0.0526
Riverside	1,038,200	58,684	0.0565
San Bernardino	995,600	54,070	0.0543
Ventura	515,600	41,990	0.0814
Total HH Telephone Numbers	11,173,300	465,629	

The probability of selection of a household telephone for the RSA augment is the probability of being selected in any particular RSA within the counties of study. This probability is calculated as $P_2=n_{2i}/N_{2i}$ where the numerator is the verified number of households in the sample selected in the *i*th RSA. Sample needed to be dialed before it was determined to be in a given RSA. Therefore, sample selected is equal to verified households and not household telephone numbers. N_{2i} is the total number of households in the *i*th RSA. The values of n_{2i}, N_{2i}, and P_2 are presented in Table 3. (Due to small sample sizes in some RSAs, the RSAs are collapsed for this calculation as indicated in the table. The full set of RSAs and collapsed RSAs is presented in Table 5 below.)

Table 3: Probability of Selection (Sampling Weights) by RSA

RSA (i)	Total Number of Households (N_{2i})	Verified Households in Sample Selected (n_{2i})	Probability of Selection (P_2)
7, 12	291,994	4,742	0.0162
8, 9, 10, 11	161,292	9,200	0.0570
15, 16	158,966	1,641	0.0103
23, 24	234,648	3,057	0.0130
25, 26, 27	533,325	1,886	0.0035
30, 32	116,541	663	0.0057
31, 33, 34	28,210	2,884	0.1022
35, 36	128,985	2,413	0.0187
41, 43, 44	191,173	3,711	0.0194

45, 46	202,729	2,772	0.0137
48, 50, 51	70,440	1,894	0.0269
52, 53, 54	123,259	2,992	0.0243
Total HHs	2,241,562	37,855	

To calculate a probability of selection for the combined samples (Base plus RSA), the probability of selection in each of the separate samples needs to be considered. The overall probability of selection will be $PT_{i=} P_{1+} P_{2i}$ for all sample households in the i*th* RSA. So for example, a household in Los Angeles County RSA 7 would need to receive an overall probability of selection of 0.048 (the sum of 0.0318 from Table 2 and 0.0162 from Table 3). For RSAs in which there are no RSA Augment sample households, the probability of selection will be simply PT_i =P1. (For example, a household in Los Angeles County RSA 17 would only receive the Los Angeles County probability of selection of 0.0318). Once PT_i is determined, the corresponding weight will be $W_i = 1/PT_i$.

~~B.~~2 Tables of Sample Statistics – With and Without Probability of Selection Weight

Formatted: Bullets and Numbering

The tables below illustrate the utility of applying the probability of section weight to the combined samples -- Base plus RSA samples. In the tables, unweighted and weighted sample statistics are compared with population parameters from the 2000 Census. The unweighted completed sample size is less than that in the data collection target (16,506 versus 18,100) due to the curtailment of the RSA augment during the fieldwork period due to contractual issues, and thus, far less oversampling at the subregional level was conducted than was specified in the original sample design.

Table 4: Completed Households by County:
Unweighted and Weighted by Probability of Selection

County	Universe: Occupied Households	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
Imperial	39,384	0.7	2.9	5.6	911
Los Angeles	3,133,774	58.2	58.3	43.6	6,982
Orange	935,287	17.4	11.5	13.8	2,271
Riverside	506,218	9.4	10.8	12.5	2,286
San Bernardino	528,594	9.8	10.3	12.7	2,139
Ventura	243,234	4.5	6.3	11.8	1,917
Total	5,386,491	100.0	100.0	100.0	16,506

Table 5: Completed Households by RSA:
Unweighted and Weighted by Probability of Selection

RSA	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
2	1.11	2.28	4.18	689
3	1.52	2.24	4.12	679
4, 6	1.02	1.04	1.92	317
5	0.88	0.76	1.39	230

7, 12	5.41	3.30	2.40	396
8, 9, 10, 11	2.99	2.72	2.15	355
13, 14	4.41	2.68	1.93	318
15, 16	2.95	5.04	3.68	607
17	8.49	10.93	7.87	1298
18	4.05	5.74	4.14	682
19	3.03	4.59	3.31	546
20	3.48	4.62	3.33	549
21	5.00	5.86	4.23	697
RSA	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
22	4.09	4.11	2.97	489
23, 24	4.35	4.82	3.49	576
25, 26, 27	9.89	3.91	2.84	468
28	3.83	3.55	4.32	712
29	3.26	2.83	3.48	574
30, 32	2.16	3.01	3.88	640
31, 33, 34	0.52	0.87	1.25	206
35, 36	2.39	1.58	1.90	314
37	2.46	2.00	2.38	393
38	2.44	2.65	3.14	518
39	1.72	1.33	1.59	262
40	2.14	1.12	1.33	220
41, 43,44	3.54	0.83	1.07	176
42	2.69	1.96	2.33	385
45, 46	3.76	4.21	5.40	890
47	0.89	1.20	1.53	253
48, 50, 51	1.31	1.49	1.95	321
49	1.20	1.47	1.87	309
52, 53, 54	2.28	2.39	3.10	512
55	0.73	2.86	5.52	911
Total	100.00	100.00	100.00	16,506

Table 6: Completed Households by Household Size:
Unweighted and Weighted by Probability of Selection

Household Size	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
One Person	22.7	31.9	30.2	4,979
Two Persons	27.4	34.8	35.2	5,812
Three Persons	16.0	13.6	14.0	2,316
Four Persons	15.4	11.6	12.1	1,992
Five Persons	9.0	5.3	5.5	908

Six Persons	4.8	2.0	2.1	346
7+ Persons	4.7	0.9	0.9	153
Total Percent	100.0	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506	16,506

Table 7: Completed Households by Number of Vehicles:
Unweighted and Weighted by Probability of Selection

Household Vehicles	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
Zero Vehicle	10.1	6.2	5.5	907
One Vehicle	34.8	36.4	35.3	5,826
Two Vehicle	37.2	39.6	40.3	6,658
Three Vehicles	12.6	12.7	13.3	2,200
Four Vehicles	3.8	3.7	4.0	655
Five+ Vehicles	1.5	1.5	1.6	260
Total Percent	100.0	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506	16,506

Table 8: Completed Households by 2000 Annual Household Income:
Unweighted and Weighted by Probability of Selection

Income	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
Up to $9,999	9.1	6.4	6.2	893
$10,000 to $24,999	18.0	16.1	16.0	2,322
$25,000 to $34,999	11.7	13.0	13.0	1,883
$35,000 to $49,999	15.2	14.5	14.7	2,127
$50,000 to $74,999	18.8	22.4	22.5	3,261
$75,000 to $99,999	11.1	12.7	12.9	1,866
$100,000 to $149,999	9.7	9.5	9.5	1,374
More than $150,000	6.3	5.5	5.3	767
Missing	N/A	---	---	2,013

Total Percent	100.0	100.0	100.0	100.0
Total Households	5,386,491	14,468	14,493	16,506

Table 9: Completed Households by Type of Dwelling Unit:
Unweighted and Weighted by Probability of Selection

Type of Dwelling Unit	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
Single Family	53.8	60.7	63.1	10,407
All Other Types	46.2	39.3	36.9	6,079
Missing	N/A	---	---	20
Total Percent	100.0	100.0	100.0	100.0
Total Households	5,386,491	16,486	16,486	16,506

Table 10: Completed Households by Tenure:
Unweighted and Weighted by Probability of Selection

Tenure	% of Universe	% of Sample Weighted	% of Sample Unweighted	Completed Households in Sample Unweighted
Owner	54.8	60.2	63.1	10,383
Renter	45.2	39.4	36.5	6,007
Other	N/A	.4	.4	72
Missing	N/A	---	---	44
Total Percent	100.0	100.0	100.0	100.0
Total Households	5,386,491	16,463	16,462	16,506

The tables indicate that there are differences in nonresponse and noncoverage among important geographic or demographic parts of the final sample due to the effect of the choice of sampling frame, the sample selection process, and the field operations. Post-stratification weights would be useful in compensating for these inadequacies and consequently, reducing some of the biases associated with them.

~~B.~~3 Post-Stratification Weights

Formatted: Bullets and Numbering

~~□~~Post-stratification is a way of improving survey sample estimates by the proper utilization of *ancillary* sources of information. In our case, we are most concerned about the deviation of the sample from the universe in terms of the distribution of household size. This is a variable that can have a significant impact on household trip rate estimates. The bias in our sample against larger size households due to the extra burden these households face when asked that all household members complete diaries would result in artificially depressed household trip rate estimates for the region. Household size is also an attractive choice for post-stratification adjustment due to the lack of item nonresponse.

Formatted: Bullets and Numbering

Post-stratification requires (a) information on the proportions of the universe on the variables of interest, and (b) information for classifying the sample cases into the same strata. Typically, post stratification cells are formed by cross classifying demographic cells (thought to be related to both nonresponse and our dependent measures). Unfortunately, the necessary population parameters are not available in cross-classification cells. In these adjustments, we will use ~~weighted~~ survey data ~~with weights~~weighted by the probability of selection weight W_i (see section 1D above) to calculate weighted totals X(k) within post stratification adjustment cell *k*. And since we will have population totals Z(*k*) for each cell *k* from 2000 census data, the post stratification adjustment PS(*k*) is defined as the ratio of the two; so, for each cell *k*:

PS(k) = Z(k) / X(k).

The household size dimension will consist of the following categories: 1, 2, 3, 4, 5, 6, and 7+ persons per household. We want to use the greater specificity of categories when constructing this weight (i.e., up to 7+ persons) because this will improve the effect of the post-stratification adjustment when expanding the data to all occupied housing units in the region.

Formatted: Bullets and Numbering

~~B.~~4 Tables of Sample Statistics – With Probability of Selection Weight and Household Size Weight

Tables 11 – 17 compare the distribution of delivered ~~main~~Base Sample households compared to the census distribution for the 6-county region.

Table 11: Completed Households by Household Size:
Weighted by Probability of Selection and Household Size

Household Size	% of Universe	% of Sample Combined Wgt	Completed Households HH Size Wgt
One Person	22.7	22.7	3,747
Two Persons	27.4	27.4	4,523
Three Persons	16.0	16.0	2,641
Four Persons	15.4	15.4	2,542
Five Persons	9.0	9.0	1,486
Six Persons	4.8	4.8	792
7+ Persons	4.7	4.7	776
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 12: Completed Households by County:
Weighted by Probability of Selection and Household Size

County	Universe: Occupied Households	% of Universe	% of Sample Combined Weight	%of Sample Just Prob. of Selection Wgt
Imperial	39,384	0.7	3.2	2.9

Los Angeles	3,133,774	58.2	56.8	58.3
Orange	935,287	17.4	11.2	11.5
Riverside	506,218	9.4	11.6	10.8
San Bernardino	528,594	9.8	10.8	10.3
Ventura	243,234	4.5	6.2	6.3
Total	5,386,491	100.0	100.0	100.0

Table 13: Completed Households by RSA:
Weighted by Probability of Selection and Household Size

RSA	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection Wgt.
2	1.11	2.1	2.28
3	1.52	2.3	2.24
4, 6	1.02	1.1	1.04
5	0.88	0.7	0.76
7, 12	5.41	3.0	3.30
8, 9, 10, 11	2.99	2.8	2.72
13, 14	4.41	2.5	2.68
15, 16	2.95	4.2	5.04
17	8.49	9.6	10.93
18	4.05	5.5	5.74
19	3.03	4.5	4.59
20	3.48	4.4	4.62
21	5.00	7.3	5.86
22	4.09	4.5	4.11
23, 24	4.35	4.5	4.82
25, 26, 27	9.89	3.9	3.91
28	3.83	3.9	3.55
29	3.26	3.0	2.83
30, 32	2.16	3.2	3.01
31, 33, 34	0.52	0.8	0.87
35, 36	2.39	1.5	1.58
37	2.46	2,1	2.00
38	2.44	2.4	2.65
39	1.72	1.2	1.33
40	2.14	1.0	1.12
41, 43,44	3.54	0.8	0.83
42	2.69	2.2	1.96
45, 46	3.76	4.8	4.21
47	0.89	1.4	1.20
48, 50, 51	1.31	1.5	1.49
49	1.20	1.6	1.47
52, 53, 54	2.28	2.3	2.39

55	0.73	3.2	2.86
Total	100.00	100.0	100.00

Table 14: Completed Households by Number of Vehicles:
Weighted by Probability of Selection and Household Size

Household Vehicles	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection Wgt.
Zero Vehicle	10.1	5.8	6.2
One Vehicle	34.8	31.8	36.4
Two Vehicle	37.2	40.6	39.6
Three Vehicles	12.6	14.8	12.7
Four Vehicles	3.8	4.8	3.7
Five+ Vehicles	1.5	2.1	1.5
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 15: Completed Households by 2000 Annual Household Income:
Weighted by Probability of Selection and Household Size

Income	% of Universe	% of Sample Combined Wgt	% of Just Prob. of Selection Wgt.
Up to $9,999	9.1	6.1	6.4
$10,000 to $24,999	18.0	16.1	16.1
$25,000 to $34,999	11.7	13.4	13.0
$35,000 to $49,999	15.2	14.4	14.5
$50,000 to $74,999	18.8	22.4	22.4
$75,000 to $99,999	11.1	12.9	12.7
$100,000 to $149,999	9.7	9.2	9.5
More than $150,000	6.3	5.6	5.5

Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	14,571	14,468

Table 16: Completed Households by Type of Dwelling Unit:
Weighted by Probability of Selection and Household Size

Type of Dwelling Unit	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection Wgt
Single Family	53.8	64.1	60.7
All Other Types	46.2	35.9	39.3
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,487	16,486

Table 17: Completed Households by Tenure:
Weighted by Probability of Selection and Household Size

Tenure	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection Wgt.
Owner	54.8	60.9	60.2
Renter	45.2	38.7	39.4
Other	N/A	0.4	0.4
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,467	16,463

Because household size is correlated with other geographic and demographic variables (as indicated with the validation variables above), the impact of applying this weight is to introduce other biases. In this case, we have increased the proportion of single-family households and increased the proportion of households with two ore more vehicles. Also geography is important since one key objective of the survey data is to be able to provide volume estimates at the county level. RSAs comprise the "building blocks" for the county samples. At this point, the distributions on such variables as dwelling type and number of vehicles are "off" to a greater extent than geography. It is therefore decided to adjust by number of vehicles and then, if needed, by geography. Number of vehicles is selected as the adjustment variable because it has no missing data and because it is known to be associated with trip-making behavior.

~~B.~~5 Tables of Sample Statistics – With Probability of Selection Weight, Household Size Weight, and Number of Vehicles Weight

Formatted: Bullets and Numbering

Tables 17 – 22 compare the distribution of delivered ~~main~~Base Sample households compared to the census distribution for the 6-county region with the additional weight by Number of Vehicles.

Table 17: Completed Households by Number of Vehicles:
Weighted by Probability of Selection, Household Size, Number of Vehicles

Household Vehicles	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection & HH Size Wgt
Zero Vehicle	10.1	10.1	5.8
One Vehicle	34.8	34.8	31.8
Two Vehicle	37.2	37.2	40.6
Three Vehicles	12.6	12.6	14.8
Four Vehicles	3.8	3.8	4.8
Five+ Vehicles	1.5	1.5	2.1
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 18: Completed Households by Household Size:
Weighted by Probability of Selection, Household Size, and Number of Vehicles

Household Size	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection & HH Size Wgt
One Person	22.7	22.8	22.7
Two Persons	27.4	27.4	27.4
Three Persons	16.0	16.0	16.0
Four Persons	15.4	15.4	15.4
Five Persons	9.0	9.0	9.0
Six Persons	4.8	4.8	4.8
7+ Persons	4.7	4.7	4.7
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 19: Completed Households by County:
Weighted by Probability of Selection, Household Size, and Number of Vehicles

County	Universe: Occupied Households	% of Universe	% of Sample Combined Weight	% of Sample Just Prob. of Selection & HH Size Wgt
Imperial	39,384	0.7	3.3	3.2
Los Angeles	3,133,774	58.2	58.0	56.8
Orange	935,287	17.4	11.0	11.2
Riverside	506,218	9.4	11.3	11.6
San Bernardino	528,594	9.8	10.6	10.8
Ventura	243,234	4.5	5.9	6.2
Total	5,386,491	100.0	100.0	100.0

For expediency sake, we did not include a geography table by RSA.

Table 20: Completed Households by 2000 Annual Household Income:
Weighted by Probability of Selection, Household Size, and Number of Vehicles

Income	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection & HH Size Wgt.
Up to $9,999	9.1	7.9	6.1
$10,000 to $24,999	18.0	18.3	16.1
$25,000 to $34,999	11.7	13.9	13.4
$35,000 to $49,999	15.2	14.0	14.4
$50,000 to $74,999	18.8	20.9	22.4
$75,000 to $99,999	11.1	11.7	12.9
$100,000 to $149,999	9.7	8.3	9.2
More than $150,000	6.3	5.0	5.6
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	14,539	14,571

Table 21: Completed Households by Type of Dwelling Unit:
Weighted by Probability of Selection, Household Size, and Number of Vehicles

Type of Dwelling Unit	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection & HH Size Wgt.
Single Family	53.8	61.1	64.1
All Other Types	46.2	38.9	35.9
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,488	16,487

Table 22: Completed Households by Tenure:
Weighted by Probability of Selection, Household Size, and Number of Vehicles

Tenure	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection & HH Size Wgt..
Owner	54.8	57.0	60.9
Renter	45.2	42.6	38.7
Other	N/A	0.4	0.4
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,468	16,467

The use of the additional weight did serve to mediate a large portion of the demographic biases that were evident in the sample. We hesitate to adjust by other demographic variables as we believe we have selected those most associated with trip-making behavior and also those that do not have any missing data. Our distributions on the validation variables are within tolerant levels with the exception of Type of Dwelling. As mentioned previously, a significant adjustment is the one to bring the sample distributions in line with the population parameters by county since it will be important to be able to produce volume estimates at the county level. Therefore, the final weight variable proposed is geography.

Formatted: Bullets and Numbering

~~B.~~6 Tables of Sample Statistics – With Probability of Selection Weight, Household Size Weight, Number of Vehicles Weight, and Region Weight

Tables 23-28 compare the distribution of delivered ~~main~~Base Sample households compared to the census distribution for the 6-county region with the additional weight by region.

Table 23: Completed Households by County:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

County	Universe: Occupied Households	% of Universe	% of Sample Combined Weight	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
Imperial	39,384	0.7	0.7	3.3
Los Angeles	3,133,774	58.2	58.2	58.0
Orange	935,287	17.4	17.4	11.0
Riverside	506,218	9.4	9.4	11.3
San Bernardino	528,594	9.8	9.8	10.6
Ventura	243,234	4.5	4.5	5.9
Total	5,386,491	100.0	100.0	100.0

For expediency sake, we did not include a geography table by RSA.

Table 24: Completed Households by Number of Vehicles:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

Household Vehicles	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
Zero Vehicle	10.1	10.1	10.1
One Vehicle	34.8	34.8	34.8
Two Vehicle	37.2	37.2	37.2
Three Vehicles	12.6	12.6	12.6
Four Vehicles	3.8	3.8	3.8
Five+ Vehicles	1.5	1.5	1.5
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 25: Completed Households by Household Size:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

Household Size	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
One Person	22.7	22.9	22.8
Two Persons	27.4	27.3	27.4
Three Persons	16.0	16.0	16.0
Four Persons	15.4	15.4	15.4
Five Persons	9.0	9.0	9.0
Six Persons	4.8	4.8	4.8
7+ Persons	4.7	4.7	4.7
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	16,506

Table 26: Completed Households by 2000 Annual Household Income:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

Income	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
Up to $9,999	9.1	7.6	7.9
$10,000 to $24,999	18.0	18.0	18.3
$25,000 to $34,999	11.7	13.8	13.9
$35,000 to $49,999	15.2	13.8	14.0
$50,000 to $74,999	18.8	21.1	20.9
$75,000 to $99,999	11.1	11.9	11.7
$100,000 to $149,999	9.7	8.6	8.3
More than $150,000	6.3	5.3	5.0
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	14,518	14,539

Table 27: Completed Households by Type of Dwelling Unit:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

Type of Dwelling Unit	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
Single Family	53.8	60.3	61.1
All Other Types	46.2	39.7	38.9
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,488	16,488

Table 28: Completed Households by Tenure:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region

Tenure	% of Universe	% of Sample Combined Wgt	% of Sample Just Prob. of Selection, HH Size, # Veh Wgt
Owner	54.8	56.7	57.0
Renter	45.2	42.9	42.6
Other	N/A	0.4	0.4
Missing	N/A	---	---
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,469	16,468

The final weight is the product of the selection probability weight and the post stratification weights. This weight will be applied to each sampled case. The final weight will be normalized in two ways:

1. to ensure that the sum of the weights is equal to the final sample size (so that when weighted, the final data set contains the same number of households as the unweighted data set).
2. to ensure that the sum of the weights in each RSA is equal to the total population of households in each RSA.

2. DOCUMENTATION OF EXPANSION FACTOR

Once the final weight is normalized, each weight is multiplied by 326.3353 (the total study area population of 5,386,491 divided by the number of surveyed households 16,506) to create the expansion factor for each record.

The final weight for each household unit will be unique based on its geographic location (both county and RSA), as well as its household size and vehicle availability. However, it is a "stand alone" weight in that it does not change if the unit of analysis falls below the regional level. If SCAG desires to perform an analysis of data at the regional level or for Orange County households only, the same weight would be used. The same is true of the expansion factor.

Tables 29 – 34 provide validation of the application of the expansion factor.

Table 29: Completed Households by County:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

County	Universe: Occupied Households	% of Universe	% of Expanded Sample	Counts for Expanded Sample
Imperial	39,384	0.7	0.7	39,384
Los Angeles	3,133,774	58.2	58.2	3,133,774
Orange	935,287	17.4	17.4	935,287
Riverside	506,218	9.4	9.4	506,218
San Bernardino	528,594	9.8	9.8	528,594
Ventura	243,234	4.5	4.5	243,234
Total	5,386,491	100.0	100.0	5,386,491

For expediency sake, we did not include a geography table by RSA.

Table 30: Completed Households by Number of Vehicles:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

Household Vehicles	% of Universe	% of Expanded Sample	Counts for Expanded Sample
Zero Vehicle	10.1	10.1	544,175
One Vehicle	34.8	34.8	1,874,588
Two Vehicle	37.2	37.2	2,003,677
Three Vehicles	12.6	12.6	678,617
Four Vehicles	3.8	3.8	204,650
Five+ Vehicles	1.5	1.5	80,784
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	5,386,491

Table 31: Completed Households by Household Size:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

Household Size	% of Universe	% of Expanded Sample	Counts for Expanded Sample
One Person	22.7	22.9	1,231,005
Two Persons	27.4	27.3	1,473,049
Three Persons	16.0	16.0	860,042
Four Persons	15.4	15.4	827,441
Five Persons	9.0	9.0	483,833
Six Persons	4.8	4.8	258,252
7+ Persons	4.7	4.7	252,868
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,506	5,386,491

Table 32: Completed Households by 2000 Annual Household Income:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

Income	% of Universe	% of Expanded Sample	Counts for Expanded Sample
Up to $9,999	9.1	6.7	361,418
$10,000 to $24,999	18.0	15.8	852,659
$25,000 to $34,999	11.7	12.1	651,723
$35,000 to $49,999	15.2	12.1	652,023
$50,000 to $74,999	18.8	18.5	997,781
$75,000 to $99,999	11.1	10.5	565,154
$100,000 to $149,999	9.7	7.6	407,108
More than $150,000	6.3	4.6	250,086
Missing	N/A	12.0	648,538
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	14,518	5,386,491

Table 33: Completed Households by Type of Dwelling Unit:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

Type of Dwelling Unit	% of Universe	% of Expanded Sample	Counts for Expanded Sample
Single Family	53.8	60.2	3,244,659
All Other Types	46.2	39.7	2,135,937
Missing	N/A	0.1	5,896
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,488	5,386,491

Table 34: Completed Households by Tenure:
Weighted by Probability of Selection, Household Size, Number of Vehicles, and Region and Expanded to Population Parameters

Tenure	% of Universe	% of Expanded Sample	Counts for Expanded Sample
Owner	54.8	56.6	3,048,365
Renter	45.2	42.8	2,305,024
Other	N/A	0.4	21,063
Missing	N/A	0.2	12,039
Total Percent	100.0	100.0	100.0
Total Households	5,386,491	16,469	5,386,491

Appendix D: Recruitment Questionnaire

Hi my name is_____ and I'm calling on behalf of the Southern California Travel and Congestion Study. The information gathered in this study will be used for transportation planning in<COUNTY >.
May I speak with someone 18 years of age or older?

Your household has been selected to participate in a very important survey of travel patterns. The information will be used by transportation planners to identify where and when traffic is most congested and to come up with solutions to reduce traffic congestion. You should have received a letter and brochure about the study. For this survey, first we will conduct the phone interview with you that I would like to start on right now. Second, we will mail you specially prepared logs for your household to track their travel for one day. And finally we will call you back to collect the information that you recorded in your logs. Unless you have some questions I will start right now.

MODE
I'd like to know if anyone in your household has used any of the following forms of transportation within Southern California in the last 7 days.

NONE	0
Park-n-Ride to Metrorail	1
Park-n-Ride to Express Bus	2
Metrolink	3
Toll Road	4
Metrorail (Red, Green or Blue)	5
Express Bus	6
Local or Community Bus	7
DK/RF	9

HHVEH
Now, how many vehicles are presently available to members of your household? This includes all cars, vans, trucks, RVs, SUVs, motorcycles and mopeds, whether owned or leased or provided by an employer.

ZERO	00
ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT+ [specify]	08
DK	98
RF	99

VEH YEAR
Now think about your vehicles in order from being driven the most to being driven the least. What is the year of<VEH > RANGE:
1900 TO 2002
DK 8888
RF 9999

MAKE
What is the make of that vehicle?

ACURA	01
AUDI	02
BMW	03
BUICK	04
CADILLAC	05
CHEVROLET	06
CHRYSLER	07
DAEWOO	41
DODGE	08
FORD	09
GEO	10
GMC	11
HARLEY DAVIDSON	12

HONDA	13
HYUNDAI	14
INFINITI	15
ISUZU	16
JAGUAR	17
JEEP	18
KAWASAKI	19
KIA	20
LEXUS	21
LINCOLN	22
MAZDA	23
MERCURY	24
MERCEDES	25
MITSUBISHI	26
NISSAN	27
OLDSMOBILE	28
PLYMOUTH	29
PONTIAC	30
PORSCHE	31
RANGE ROVER	32
SAAB	33
SATURN	34
SUBARU	35
SUZUKI	36
TOYOTA	37
VOLKSWAGEN	38
VOLVO	39
YAMAHA	40
OTHER, SPECIFY	97
DON'T KNOW	98
REFUSED	99

MODEL
What is the model?

BTYPE
What is the body type of your vehicle?

Auto	01
Van	02
RV	03
Sport Utility Vehicle	04
Pick-up Truck	05
Other Truck	06
Motorcycle/Moped	07
OTHER, SPECIFY	97
DK/RF	99

FUEL
What type of fuel does the vehicle use?

Gasoline	1
Diesel	2
Electric	3
OTHER, SPECIFY	7
DK	8
RF	9

CIGAR
Does this vehicle have a working 12-volt outlet or cigarette lighter adapter? IF RESPONDENT ASKS WHY WE WANT TO KNOW, SAY: To determine if this vehicle is GPS-compatible.

YES	1
NO	2
DK	8
RF	9

VOWN
Is this vehicle owned or leased?

Owned by a household member	1
Owned by a person not in your household	2
Leased	3

OLEAS
Who leases this vehicle?

A household member	1
An employer	2
Someone else, SPECIFY	7

ACQUR
What year did you acquire this vehicle? ENTER ALL 4 DIGITS RANGE: 1900 TO 2001

DK	8889
RF	9999

DRIVE
Who usually drives this vehicle? Can you give me their first name or initials?

CONTINUE WITH NEXT VEHICLE 1

BIKES
Now, how many bicycles in working condition are available to members of your household for use in their daily travel?

ZERO	00
ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT+ [specify]	08
DK	98

RESTY
Do you live in a...

Unattached Single Family Home	1
Duplex	2
Apartment	3
Condominium or townhouse	4
Mobile home or trailer	5
Group quarters (dorms, barracks, etc. (eligible respondents based on the "room" to which phone assigned.)	6
OTHER, SPECIFY	7
DK/RF	9

OWN
Do you own or rent this home?

Own/buying	1
Rent	2
OTHER, SPECIFY	7
DK	8
RF	9

HLIVE
How many months out of the year do you live at this address? FORMAT: 1 TO 12

HLANG
What language or languages are spoken in your home? READ LIST, MULTIPLE RESPONSE

English	01
Spanish	02
Tagalog	03
Chinese (Mandarin/Cantonese)	04
Japanese	05
Vietnamese	06
Cambodian	07
Korean	08
OTHER, SPECIFY	97

HHSIZ
Now I need to get information about the persons in your household. How many people, including yourself, live in your household?
CLARIFY: Includes all persons who sleep there at least 3 nights per week.

ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT+ [specify]	08
DK	98

What is your first name? VERIFY THAT THIS IS THE RESPONDENT
What is the first name of the other person living in your home?
What is the first name of the next person in your home, from oldest to youngest

GEND
What is<FNAME >'s gender? DON'T ASK FOR RESPONDENT

MALE	1
FEMALE	2
DK/RF	9

AGE
What is<YOUR>age in years? IF LESS THAN 1, ENTER 0

98 years and older	98
DK/RF	99

RELAT
What is<YOUR>relationship to you? DON'T ASK FOR RESPONDENT

SELF	1
Spouse/partner	2
Son/Daughter	3
Mother/Father/Mother In-law/Father In-law	4
Other relative	5
Live-in Help	6
Not Related	7
DK/RF	9

LIC (Age 15+)
Do/Does<YOU>have a valid driver's license?

Yes	1
No	2
DK	8
RF	9

DISAB
Do/Does<YOU>have a physical, mental, or other health disability that has lasted 6 or more months and which makes it difficult for<YOU2>to go outside the home alone, for example to shop or visit a doctor's office?

Yes	1
No	2
DK/RF	9

DISTY
What type of disability is that?

None	1
Difficulty standing, walking or climbing stairs	2
Visually impaired/blind	3
Hearing impaired/deaf	4
Wheelchair	5
Require cane/walker	6
OTHER, SPECIFY	7
DK/RF	9

ETHN
Which of the following best describes<YOUR>race/ethnicity?

White/Not Hispanic	1
Hispanic	2
African American	3
Asian/Pacific Islander	4
OTHER, SPECIFY	7
DK/RF	9

EMPLY (Age 16+)
Are<YOU>employed full-time or part-time?

Yes	1
No	2
DK/RF	9

PRIMA
Which of the following best describes<YOUR>current situation?

Retired	1
Disability status	2
Homemaker	3
Unemployed	4
Other, SPECIFY	7
DK/RF	9

JOBS
How many paying jobs do/does<YOU>hold?

ONE	01
TWO	02
THREE+	03
DK/RF	99

JOBF1
On average, how many hours a week do/does<YOU>work at<YOUR><MAINW>job?

97+ hours	97
DK	98
RF	99

FTPT (computed variable)

FULL-TIME	1
PART-TIME	2

HRVR1
Are<YOUR>work hours fixed or do they vary?

FIXED	1
VARIED	2

WHYV1
Which of the following best describes<YOUR>situation?

Hours vary at my choice	1
Hours allowed to vary within fixed limits	2
Fixed starting time, but variable ending time	3
Fixed hours, but different hours on different days	4
Variable depending on work	5
OTHER, SPECIFY	7
DK/RF	9

BUST1
In what type of business or industry is<YOUR><MAINW>employer? DO NOT READ LIST. SELECT BEST MATCH AND READ BACK TO VERIFY.

Agriculture/Forestry/Fishing	11
Mining	21
Utilities	22
Construction	23
Manufacturing - Nondurable goods	31
Manufacturing - Durable goods	32
Wholesale trade	41
Retail trade	44
Transportation and Warehousing	48
Information	51

Finance or Insurance	52
Real Estate	53
Professional, Scientific, and Technical services	54
Management of companies and enterprises	55
Administrative and support services	56
Educational Services	61
Healthcare and social assistance	62
Arts, entertainment, and recreation	71
Accommodations and food services	72
Other Services (excluding public administration)	81
Public Administration	92
OTHER, SPECIFY	97
DK	98
RF	99

OCCU1
What do/does<YOU>do there? DO NOT READ LIST. SELECT BEST MATCH AND READ BACK TO VERIFY.

Executive, Admin, or Managerial	01
Professional specialty	02
Technicians or related support	03
Sales	04
Administrative support, clerical	05
Private Household	06
Protective Services	07
Service, except protective and household	08
Farming, Forestry, or Fishing	09
Precision, Production, Craft, or Repair	10
Machine operator, assembler, or inspector	11
Transportation, or material moving	12
Handler, equipment cleaner, helper, or laborer	13
OTHER, SPECIFY	97
DK	98
RF	99

JLOC1
Which of the following best describes<YOUR>job location?

Home	1
Fixed Address	2
No fixed address (e.g. traveling salesman, repairman, etc.)	3
DK/RF	9

WCNT1
Is<YOUR><MAINW>job in<HCNTY >? SELECT APPROPRIATE COUNTY

Imperial County	025
Los Angeles County	037
Orange County	059
Riverside County	065
San Bernardino	071
Ventura	111
OTHER, SPECIFY	997
DK/RF	999

PRIMARY WORKPLACE
What is the name of the place where you/he/she work(s)?
@WNAM1
What is the street address?
@WADD1
What is the name of a street that crosses near
<WADD1 >
@WXST1
What is a landmark that is near there?
@WLND1
What city is that in? @ PRESS F2 FOR TABLE; 7777=OTHER, SPECIFY 9999=DK/RF
What is the Zip Code? @WZIP1 99999=DK/RF

VEHR1
Does<YOUR>employer require that one of the household vehicles be driven to work?

YES	1
NO	2
DK/RF	9

FREE1
Is parking free at the workplace?

YES	1
NO	2
DK/RF	9

EMPP1
Does the employer offer to pay any of the parking costs?

YES	1
NO	2
DK/RF	9

EMPP2
Do/Does<YOU>use the employer's offer to pay?

YES	1
NO	2
DK/RF	9

EMPT1
Does<YOUR>employer offer to pay for any of the cost for public transit?

YES	1
NO	2
DK/RF	9

EMPT2
Do/Does<YOU>use the employer's offer to pay?

YES	1
NO	2
DK/RF	9

MODE1
How do/does<YOU>usually get to<YOUR><MAINW>job?

Drive alone	1
Auto, Van shared ride - 2+ persons	2
Public transit	3
Walk	4
Bike	5
DK/RF	9

PRKP1
How much, if anything, do/does<YOU>personally pay to park? YOU CAN ENTER CHANGE AMOUNTS, LIKE "50 CENTS"=0.50

FREE	000000
DK/RF	999999

PARU1
What time period does that payment cover?

Daily	1
Weekly	2
Monthly	3
Annualy	4
OTHER, SPECIFY	7
DK/RF	9

BUSP1
How much do/does<YOU>personally pay for a transit pass or ride? YOU CAN ENTER CHANGE AMOUNTS, LIKE "50 CENTS"=0.50

FREE	000000
DK/RF	999999

BUSU1
What time period does that payment cover?

Per Ride	0
Daily	1
Weekly	2
Monthly	3
Annualy	4
OTHER, SPECIFY	7
DK/RF	9

BUST2
In what type of business or industry is<YOUR>second workplace? DO NOT READ LIST. SELECT BEST MATCH AND READ BACK TO VERIFY.

Agriculture/Forestry/Fishing	11
Mining	21
Utilities	22
Construction	23
Manufacturing - Nondurable goods	31
Manufacturing - Durable goods	32
Wholesale trade	41
Retail trade	44
Transportation and Warehousing	48
Information	51
Finance or Insurance	52
Real Estate	53
Professional, Scientific, and Technical services	54
Management of companies and enterprises	55
Administrative and support services	56
Educational Services	61
Healthcare and social assistance	62
Arts, entertainment, and recreation	71
Accommodations and food services	72
Other Services (excluding public administration)	81
Public Administration	92
OTHER, SPECIFY	97
DK	98
RF	99

OCCU2
What is<YOUR>occupation or what do/does<YOU>do at work? DO NOT READ LIST. SELECT BEST MATCH AND READ BACK TO VERIFY.

Executive, Admin, or Managerial	01
Professional specialty	02
Technicians or related support	03
Sales	04
Administrative support, clerical	05
Private Household	06
Protective Services	07
Service, except protective and household	08
Farming, Forestry, or Fishing	09
Precision, Production, Craft, or Repair	10
Machine operator, assembler, or inspector	11
Transportation, or material moving	12
Handler, equipment cleaner, helper, or laborer	13
OTHER, SPECIFY	97
DK	98
RF	99

HWOR2
How many hours a week on average, do/does<YOU>work at this second job?

JLOC2
Is the location of this second job...

Home	1
Fixed Address	2
No fixed address (e.g. traveling salesman, repairman, etc.)	3
DK/RF	9

WCNT2
Is<YOUR><2nd>job in<HCNTY >? SELECT APPROPRIATE COUNTY

Imperial County	025
Los Angeles County	037
Orange County	059
Riverside County	065
San Bernardino	071

Ventura	111
OTHER, SPECIFY	997
DK/RF	999

SECONDARY WORKPLACE
What is the name of the place where you/he/she work(s)?
@WNAM1
What is the street address?
@WADD1
What is the name of a street that crosses near
<WADD1 >
@WXST1
What is a landmark that is near there?
@WLND1
What city is that in? @ PRESS F2 FOR TABLE; 7777=OTHER, SPECIFY 9999=DK/RF
What is the Zip Code? @WZIP1 99999=DK/RF

STUDE
Do/Does<YOU>go to any type of school, including day-care, technical institutes, adult classes, or senior care?

YES	1
NO	2
DK	8
RF	9

SCHOL
What type of school is that?

Daycare/Pre-School	1
K - 6	2
7 - 12	3
Trade/Technical	4
College Undergraduate Studies	5
College Graduate Studies	6
OTHER, SPECIFY	7
DK/RF	9

SDAYS
How many days per week do/does<YOU>go to this place? 1 TO 7 DAYS

SLOC1
Do/Does<YOU>attend school at home or at another location?

HOME	1
ANOTHER SCHOOL LOCATION	2
DK/RF	9

SCNT1
Is the school in<HCNTY >? SELECT APPROPRIATE COUNTY

Imperial County	025
Los Angeles County	037
Orange County	059
Riverside County	065
San Bernardino	071
Ventura	111
OTHER, SPECIFY	997
DK/RF	999

SCHOOL LOCATION

What is the name of the school?
@SNAM1

What is the name of the street it is on?
@SADD1

What is the name of a street that crosses near
<SADD> @SXST1

What city is that in? @SCIX1 PRESS F2 FOR TABLE; 7777=OTHER, SPECIFY; 9999=DK/RF
Do you know the zip code? @SZIP1 99999=DK/RF
106: SNAM1

EDUCA
What is the highest level of education<YOU>have/has attained?

11th grade or less	1
High school graduate	2
2 years of college/Associates Degree	3
4 years of college/Bachelors degree	4
Post-Graduate	5
OTHER, SPECIFY	7
DK/RF	9

GO TO NEXT PERSON 1
DONE WITH HH MEMBERS 2

EQUIP
I just have a few more questions about your household. First, which of the following pieces of equipment or services are used in your home, including a home office?

NONE OF THESE	0
A portable cellular telephone	1
A fax machine	2
A desktop or laptop computer	3
Web TV	4
Answering Machine/Voice Mail	5
Caller ID	6
Call Blocking	7
Internet Service	8
RF	9

INCA
Was your total household income in 2000 from all sources before taxes, for all members of your household ...?

Above 50K	1
Below 50K	2
DK/RF	9

INCB
I will read you a series of income ranges. Please stop me when I read the range that is closest to your household's. IF DON'T KNOW OR REFUSED, REMIND RESPONDENT OF THE IMPORTANCE OF THE SURVEY AND THE IMPORTANCE OF THE INFORMATION TO MAKE SURE WE INCLUDE ALL TYPES OF HOUSEHOLDS

$50,000 to $74,999	5
$75,000 to $99,999	6
$100,000 to $149,999	7
$150,000 or more	8
DK/RF	9

INCC
I will read you a series of income ranges. Please stop me when I read the range that is closest to your household's. IF DON'T KNOW OR REFUSED, REMIND RESPONDENT OF THE IMPORTANCE OF THE SURVEY AND THE IMPORTANCE OF THE INFORMATION TO MAKE SURE WE INCLUDE ALL TYPES OF HOUSEHOLDS

Less than $10,000	1
$10,000 to $24,999	2
$25,000 to $34,999	3
$35,000 to $49,999	4
DK/RF	9

PHLNS
How many telephone lines do you have?

ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT	08
DK	98
RF	99

DEDIC
How many of those telephone lines are only used for a fax or modem?

ZERO	00
ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT	08
DK	98
RF	99

SHRE2
How many other households share a phone line with your household?

ZERO	00
ONE	01
TWO	02
THREE	03
FOUR	04
FIVE	05
SIX	06
SEVEN	07
EIGHT	08
DK	98
RF	99

NOPHN
Have there been times within the past 12 months when the home you were living in did not have telephone service?

YES	1
NO	2
DK	8
RF	9

LENGT
How long were you without phone service? Was it . . .

Less than 1 month	1
1 month to less than 1 year	2
1 year or longer	3
DK	8
RF	9

ASSN
Thank you for completing Part One of the survey. In Part Two, people in your household are asked to record their travel for one day in diaries that we'll mail to you. The diary package will contain instructions and more information about the survey. Are there any questions about the diary that I can answer now? Okay. Your household should keep track of their travel on READ DATE. Is this okay? LIST OF ELIGIBLE ASSN DATES

As part of this study, we are asking a select number of households to help evaluate new technology that's just beginning to be used in travel studies. The new technology uses GPS to give us greater insight into how people travel. As part of the GPS study, you will receive a GPS unit for each household vehicle along with information on how to use the units. As a token of our appreciation, your household will receive $10 for each unit that is installed and used. In addition to recording your trips in the diary, will you help us evaluate this new GPS technology?

Thank you for helping us. We'll call you on the day before your travel day to make sure you received your diaries and to answer any questions you might have. Have a nice evening/day. Goodbye.

APPENDIX E – DIARY PACKET MATERIALS

APPENDIX F – RETRIEVAL SCRIPT

Hi - my name is _____ and I'm calling on behalf of the Southern California Association of Governments about the travel survey your household recently completed. May I please speak with someone over the age of 18? I'm calling to collect your travel information. First I need to verify that the information we show for your household is correct. I'd like to start by verifying the address where you live.

HHSIZ
Our records show that there are READ NUMBER BELOW people living in your household. Is this correct? MAKE CORRECTIONS AS NEEDED

VERHM
Okay- now I need to confirm the name, age and gender we have for each household member. VERIFY NAME, AGE, AND GENDER FOR EACH MEMBER OF THE FAMILY.

HHVEH
In terms of vehicles available to your household, we show that you have READ NUMBER BELOW available. Is that right?

VISIT
Did you have any overnight visitors at your home on your travel day?

YES	1
NO	2
DK/RF	9

VISNO
How many? RANGE: 1 TO 20

INT03
Great, now I'd like to collect the trip information your household recorded for<ASSN > We'll begin with your information. Do you have your diaries handy?

DIARY
Did<YOU >use<YOUR >diary to record<YOUR >travel?

Yes	1
No	2
DK/RF	9

TOTPL
How many total places did<YOU >visit over the course of the travel day?

GETUP
What time did<YOU >wake up on<YOUR >travel day? ENTER IN MILITARY TIME HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

IF PLACE 1: Okay - Where were/was<YOU >at 3 am on<YOUR >travel day?
OTHERWISE: Where did<YOU >go next?

AFTER THEY TELL YOU, SAY: Did<YOU >make any stops along the way for any reason, such as to get coffee or gas, drop someone off, or change travel modes?

IF PLACE 1: Okay - Where were/was<YOU >at 3 am on...

HOME	01
SCHOOL LOCATION	02
NEW SCHOOL ADDRESS	12
SECOND SCHOOL LOCATION -deleted	03
PRIMARY WORK PLACE	04
NEW PRIMARY WORK ADDRESS	14
SECOND WORK PLACE	05
NEW SECOND WORK ADDRESS	15
PREVIOUSLY ENTERED PLACE/SOMEONE ELSE'S HABITUAL ADDRESS	07
NEW PLACE	08
OUT OF THE TRAVEL STUDY AREA	99

LOCAT
Does the place have a name? IF THEY GAVE YOU AN EXACT NAME IN PREVIOUS QUESTION, ENTER IT HERE

PLCTY
Was that in<HCNTY >? THIS IS NOT PRE-LOADED, SELECT COUNTY

Imperial	025
Los Angeles	037
Orange	059
Riverside	065
San Bernardino	071
Ventura	111
OUT OF STUDY AREA (IN CALIFORNIA)	995
OUT OF STATE	996
DK/RF	999

CITYX
Can you tell me which city that is in?

STATE
Is this in California?

ADDR
What is the street address there?

XSTRT
Can you tell me the names of a street that crosses near<ADDR >? FORMAT: STREET 1

LAND
Can you tell me a nearby landmark that can be found easily on a map?

PLZIP
And the zip code at that location?

MODE
What was<YOUR >main means of travel to this place?

Walk	01
Bicycle	02
Drove	03
Passenger in car/truck/van	04
Local bus or community bus	05
Express bus	06
Metro Blue Line	07
Metro Green Line	08
Metro Red Line	09
Commuter Rail (Metrolink, Amtrak)	10
Dial-A-Ride/Paratransit	11
School Bus	12
Greyhound Bus	13
Taxi/Shuttle Bus/Limousine	14
Motorcycle/Moped	15
OTHER, SPECIFY	97
DK/RF	99

ARRTM
IF PLACE 1, ENTER 0300 OTHERWISE: What time did<YOU >get there? ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

TPUR1
And what was<YOUR >main activity there? DO NOT READ LIST. SELECT BEST MATCH AND CONFIRM.

NO OTHER ACTIVITIES	00
CHANGE MODE OF TRANSPORTATION	01
PICK UP SOMEONE OR GET PICKED UP	02
DROP OFF SOMEONE OR GET DROPPED OFF	03
ATM, BUY GAS, QUICK STOP FOR COFFEE, NEWSPAPER, ETC	04
SHOPPING	05
BANKING, POST OFFICE, PAY BILLS	06
WORK (INCLUDE REGULAR SCHEDULED VOLUNTEER WORK)	07
WORK-RELATED (SALES CALL, MEETING, ERRAND, ETC.)	08

SCHOOL (ATTENDING CLASSES)	09
OTHER SCHOOL ACTIVITIES (SPORTS, EXTRA-CURRICULAR)	10
CHILDCARE, DAY CARE, AFTER SCHOOL CARE	11
EAT MEAL (RESTAURANT, DRIVE THROUGH, TAKE-OUT)	12
MEDICAL	13
FITNESS ACTIVITY (PLAYING SPORTS, GYM, BIKE RIDE)	14
RECREATIONAL (VACATION, CAMPING, ETC.)	15
ENTERTAINMENT (WATCHING SPORTS, MOVIES, DANCE, BAR, ETC.)	16
VISIT FRIENDS/ RELATIVES	17
COMMUNITY MEETINGS, POLITICAL/CIVIC EVENT, PUBLIC HEARING	18
OCCASIONAL VOLUNTEER WORK	19
CHURCH, TEMPLE, RELIGIOUS MEETING	20
WITH ANOTHER PERSON AT THEIR ACTIVITY OUT OF HOME	21
OTHER PERSONAL (SPECIFY)	22
WORKING AT HOME (RELATED TO MAIN OR SECOND JOB)	23
OTHER AT HOME ACTIVITIES, SPECIFY	24
OTHER ACTIVITY (SPECIFY)	98
DK/RF	99

TPUR2
And what other activities did<YOU >do there?

OTHTR
What was the total number of people traveling with<YOU >? NOT INCLUDING THE PERSON YOU'RE ON

HHMEM
Of those, how many were household members?

PERTP
PERSON<PERNO> PLACE<PLANO>
Who was/were the person(s)?

VEHAV
Was a household vehicle used to make this trip?

YES	1
NO	2
DK	8
RF	9

VEHNO
Which vehicle did<YOU >use? REFER TO SAMPLE SHEET FOR VEHICLE NUMBER

PRKCO
Did<YOU >pay to park? How much?

PRKUN
And was that... READ LIST

Hourly	1
Daily	2
Weekly	3
Monthly	4
Quarterly	5
Annually	6
None	7
Other	8
DK/RF	9

PRKMO
Which method did<YOU >use to pay for<YOUR >parking?

DID NOT USE ANY METHOD TO PAY	0
Cash	1
Credit Card	2
Payroll Deduction/Employee Pass	3
Validated Parking Ticket	4
OTHER, SPECIFY	7
DK/RF	9

BOARD
At which bus/rail stop did<YOU >board? ENTER CROSS STREETS OF INTERSECTION

ARRBS
What time did<YOU >arrive at the bus/rail stop? ÿ ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

RGETT
How did<YOU >get to the bus/rail stop?

Walk	01
Bicycle	02
Drove	03
Passenger in car/truck/van	04
Local bus or community bus	05
Express bus	06
Metro Blue Line	07
Metro Green Line	08
Metro Red Line	09
Commuter Rail (Metrolink, Amtrak)	10
Dial-A-Ride/Paratransit	11
School Bus	12
Greyhound Bus	13
Taxi/Shuttle Bus/Limousine	14
Motorcycle/Moped	15
OTHER, SPECIFY	97
DK/RF	99

RBLOC
What was the distance of this trip in blocks?

GETON
What time did you board (get on) the bus/train? ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

SYSTM
Which bus system/transit system did<YOU >use?

ROUTE
What route did<YOU >take?

FARE
How much did it cost to ride the bus or train? FORMAT: 999.99

FTYPE
How did<YOU >pay for the trip?

Cash	1
Credit Card	2
Bus/Train Pass	3
Transfer	4
OTHER, SPECIFY	7
DK/RF	9

ALGHT
At which bus/rail stop did<YOU >get off? ENTER CROSS STREETS OF INTERSECTION

LEVBS
What time did<YOU >get off at the next stop? ÿ ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM NOTE!!!!

RGETF
How did<YOU >get from the bus/rail stop to<YOUR >next place?

Walk	1
Picked up car and drove	2
Was picked up	3
Rode bike	4

Transferred to another bus or train	5
OTHER, SPECIFY	7
DK/RF	9

WLKBL
What was the distance of this trip in blocks? IF THEY DON'T KNOW, TRY TO GET THEM TO GUESS.

CHECK
Did<YOU >go anywhere else that day?

GOBED
What time did<YOU >go to sleep on<YOUR >travel day? ENTER IN MILITARY TIME HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

UTOLL
Did<YOU >use a toll road at all for<YOUR >trips today?

YES	1
NO	2
DK/RF	9

TOLRD
Which toll road did<YOU >use?

SAN JOAQUIN TOLL ROAD (SR 73)	1
FOOTHILL/EASTERN TOLL ROAD (SR 241,261,133)	2
SR 91 EXPRESS LANES	3
DK/RF	9

TOLCO
How much did<YOU >pay for the toll? FORMAT: 999.99

TOLMO
Which method did<YOU >use to pay for<YOUR >tolls?

Cash	1
Fastrak	2
OTHER, SPECIFY	7
DK/RF	9

CARPL
Did<YOU >use a carpool lane or diamond lane during<YOUR >travel today?

YES	1
NO	2
DK/RF	9

WAHOM (If no work travel)
Did<YOU >work at home at all during your travel day>?

YES	1
NO	2
DK/RF	9

WHSTR
What time did<YOU >start<YOUR >work at home? ÿ ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

WHSTO
PERSON<PERNO> PLACE<PLANO>
What time did<YOU >stop<YOUR >work at home? ÿ ENTER IN MILITARY TIME, HHMM 0000-MIDNIGHT, 1200-NOON, 2359-11:59PM, 0030-12:30AM

NOGO
So,<YOU >made no trips, including for work or school?

NEXT
Next Place, Day, Person 1
Done With This Household 2

STUDY
Would you be willing to participate in future research studies on this subject?
Okay, great. Those are all the questions I have for you today. We appreciate you for taking the time to help us with this important study. Thank you and good day/evening.