

DATA PRODUCTION REPORT:

A report on data collection, checking, cleaning, and archiving for

Phase One of the Southern California Beach Project

APPENDICES

Prepared by

The Research Team

Submitted to

Project Funders

March 29, 2002

Appendix 1. Chico SRC Production Report

Southern California Beach Survey
Data Collection
Production Report

Prepared for

Michael Hanemann Ph.D.

By

The Survey Research Center
California State University, Chico

January 2002

TABLE OF CONTENTS

Introduction.	3
Questionnaire Development.	3
Sample & Production Management.	8
Production Statistics.	10
Interviewer Training.	18
Data Set Construction.	19
Appendix A: Outcomes For All Panel Members For All Diaries.	22
Appendix B: Panel Correspondence.	50
Appendix C: Examples Of The Web Pages.	64
Appendix D: Explanation Of Response Rates.	68

TABLE OF TABLES

Table 1: Results Of The Recruitment Survey.	11
Table 2: Results Of The Recruitment Survey (Breakdown by Eligibility).	12
Table 3: Results Of The Replenishment Survey.	13
Table 4: Results Of The Replenishment Survey (Breakdown by Eligibility).	13
Table 5: Results For Diaries.	17
Table 6: Table 6: Key To Codes Used In Reporting.	22
Table 7: Diary by diary outcomes for individual cases.	23

INTRODUCTION

In June of 1999, The Survey Research Center (SRC) at California State University, Chico was contracted to collect data and prepare data files for the Southern California Beach Project. The two contractors were University of California, Berkeley and the Santa Monica Bay Restoration Foundation. The project involved an initial recruitment survey, six panel diary surveys, and one replenishment survey.

Data collection was completed in January of 2001. As early as November 2000, the SRC began delivering data. Because of the complexity of the data, it was necessary to go through several reconstructions of the data matrix prior to final delivery. All data were delivered to Michael Hanemann by April 2001. As of June 2001, the SRC remained in consultation with the research team to assist in analysis of the submitted data.

QUESTIONNAIRE DEVELOPMENT

General Procedures:

All questionnaires were developed by members of the research team outside of the SRC, in consultation with the funding agencies.

The research team sent electronic versions of each questionnaire to the SRC, which then would be changed to a CASES-formatted questionnaire to be used during the telephone interview. Several subsequent versions of the questionnaire were submitted to the SRC, and the CASES version was updated to reflect each set of revisions.

Once a final version of the CASES survey was in place, the program was tested to insure skip patterns were sound and that the data were being stored correctly. Callers would then be called in to conduct a pretest of the questionnaire in real survey conditions with randomly selected respondents. Members of the research team and representatives of the funding agencies monitored these calls. Any problems were corrected and reprogrammed, and an additional pretest was conducted when it was determined to be necessary by the research team.

Once the pretest was completed, the questionnaire was put into production. The initial recruitment survey and the first two diaries were programmed in CASES 3.8, The last four diary surveys, as well as the replenishment survey, were

programmed in CASES 4.3.

The Initial Recruitment Survey:

On August 9, 1999 the SRC received the “final” version of the Screener questionnaire. This version was 22 pages in length. The final CASES version of the survey was well over 300 pages. From August to mid-November, the SRC received and tested five additional versions of the screener. One reason for the changes in the interview was the closing of Huntington Beach, due to sewage spills, during the Summer of 1999. The research team wanted to gather additional data from respondents whose typical beach-going behavior was altered due to the Huntington Beach closures. These questions were included in the final version of the survey.

The Diary Questionnaire:

After initial recruitment, respondents were re-interviewed at an average interval of every two months. The main purpose of these repeated interviews was to get a picture of the respondent’s beach-going activity over a complete year. To do this, the respondent was mailed a calendar upon which he or she could record his/her trips to the beach. Each calendar covered a two-month period. After a particular two-month period was ended, the respondent was called and asked about his or her beach going-activities for the two months in question.

Some respondents felt uncomfortable providing their address to the interviewer. Those respondents were offered the opportunity to retrieve the survey materials through the Internet. A web site was established from which the calendars could be downloaded. Appendix C contains examples of these web pages¹. A total of 50 respondents from the original recruitment initially chose to participate through the internet. Of those, 26 completed at least one diary survey. A total of 14 completed the entire slate of interviews. A total of 27 respondents from the replenishment survey chose to participate through the internet, however those cases were not included in the remaining diaries.

In order to collect this data in as complete a manner as possible, a diary questionnaire was developed that would

¹ Note that due to formatting problems, this appendix uses end notes that can be found at the very end of this document.

cover all the possible ways in which an individual beach trip could occur. So, for instance, a person who went to a single beach on a single day was asked a different (but equivalent) set of questions than a person who went to multiple beaches. All told, there were five different modes in which a beach trip could occur:

1. A single beach on a single day.
2. Multiple beaches on a single day.
3. A single beach over multiple days.
4. Multiple beaches over multiple days.
5. Multiple beaches gone past in a single day due to a recreational mode of transportation (i.e. bicycling past several beaches).

To simplify the interview for the respondents and interviewers, each of the different types of beach trips was covered in a different set of questions. For the most part, the questions were equivalent, but fewer questions were asked about more complicated trips.

Another way in which the diary survey varied by respondent had to do with the number of beach trips that a given respondent reported taking during the period covered by a given diary. Again, the reasoning behind this difference was a desire to reduce the burden on the respondent. Those respondents who reported having gone to the beach less than five times over a two month period were asked a series of questions about activities that they participated in for each trip they took.

If a respondent said he/she went to the beach five times or more, the program collected only the beach name and the date of each individual tripⁱⁱ. A list of individual beaches that the respondent had gone to was created, and the respondent was then asked about what activities he/she engaged in on a typical trip to a given beach. The respondent was asked about activities only once for each individual beach he/she had visited over the period covered by the diary. For example, if a respondent went to Huntington beach five times and Newport beach twice during the diary period, he/she was asked about his/her typical activities for each beach only once.

This arrangement meant that there were a total of ten different modes in which a given beach trip could be

recordedⁱⁱⁱ and an almost infinite number of combinations that could occur over a two-month period. This resulted in a first draft of the CASES-formatted survey that was literally thousands of pages in length. Though the diary remained the same throughout the panel, the method of execution within the CASES program evolved. The original intent was to have an individual set of questions for each day of the study period, but this led to a program so long that the computer could not process it. It was determined that a very small percentage of respondents (approximately 8%) would go to the beach more than 5 times during the months of December and January, so for the first diary, any trips beyond the 5 were recorded on a paper version of the survey (See Appendix A.).

Beginning with the second diary, the diary portion of the program was “rostered.” A rostered program allowed the same set of questions to be asked multiple times, with the data being stored separately for each incident. The program was set to collect up to 30 beach trips. This was the first time that the SRC had used the rostering feature of the CASES program. At the time the second diary survey was conducted, the SRC was using version 3.8 of the CASES program. It was discovered that this version of the rostering program made data difficult to output. The SRC therefore updated their software to CASES 4.3 for the remainder of the study.

The Additional Survey Modules:

In addition to the diary portion of the survey, each of the diary interviews also involved an additional module designed to collect information to be used in the final economic model based on the study data. Each module covered distinct topics, and each followed different diary surveys. The particular module types and the diary survey each followed (as indicated by the number that precedes it) were:

1. Substitution effects. A substitution module designed to determine what the respondent would have done if the last beach they visited had been closed at the time of the last trip.
2. Health effects. A medical module was added in an effort to determine if beach activities could be associated with illnesses reported by respondents.
3. Choice set and familiarity. A module was added to

determine which of all the beaches in L.A. and Orange counties the respondent had been to or was familiar with. Questions were also asked about respondent perceptions of water and sand quality and parking availability.

4. Expenditures. This module involved a detailed accounting of money expended by the respondent on his/her last two trips to the beach.
5. Contingent behavior and contingent valuation. This module involved asking the respondents to decide if they were likely to go to the beach within 4 weeks of the date they were being interviewed, and what they would do if the beach they intended to visit was closed. Contingent valuation questions were aimed to elucidate respondent willingness to pay for improved water quality in Southern California.
6. San Onofre/power plant recreation effects questions. If the respondent did not go to the beach during October or November, they were asked what for the reason. In addition, a short module was added asking the respondent if they had gone to San Onofre beach in the last two years, and, if not, why.

Modules were developed in the same manner as the rest of the survey. A proposed questionnaire was developed by the research team; the SRC then programmed the module and the module was pretested with the research team and representatives of the funding organizations listening in. Once the team approved the module, it was placed into production.

The Replenishment Survey:

During the month of August 2000, the SRC began recruiting 641 new respondents into the sample. This action was undertaken in order to counteract respondent attrition. The research team wanted to ensure the existence of a large sample size through the completion of the study. The decision was made to conduct a diary survey for the months of June and July with those respondents who qualified for the study. Because these respondents did not have the benefit of having a calendar to record their beach trips during June and July, questions were added to the diary. These questions gave the respondent the opportunity to indicate either a specific date or a general time he or she

went on a beach trip. If a respondent did not know a specific date, he or she could narrow the time down by specifying a week and whether they went to the beach on a weekday or a weekend.

SAMPLE & PRODUCTION MANAGEMENT

Initial Recruitment and Replenishment:

Sample management for the initial recruitment and replenishment portions of this study was handled according to normal procedures regarding a random digit dial (RDD) study. Each phone number released into the initial recruitment and replenishment portions of the study was called a minimum of 10 times, unless the number reached a final disposition prior to the 10th call. On every call, the interviewer was responsible for properly coding the outcome of the call. Cases are considered to have a final disposition if:

1. An interview is completed.
2. A member of the household refused to be involved in the interview and requested that we not call back.
3. A member of the household twice refused to be interviewed.
4. It was determined the number was not in service.
5. It was determined the number belonged to a business.
6. The number belonged exclusively to someone not yet 18 years of age.
7. It was determined that the selected member of the household would be absent for the entire survey period.
8. No member of the household spoke either English or Spanish.
9. The respondent terminated the call in the middle of an interview and refused a callback at another time.

Other cases were called up to 10 times. If no contact was made by the tenth call, the number was discarded.

If an unanticipated situation occurred during the call so that the interviewer did not know how to code the call, the interviewer could code the call as a “bizarre problem.” Either Allen Lunde or Nicole Thompson reviewed all problems, and the proper outcome was determined.

In late May of 2000, a meeting was held in Sacramento to discuss the progress of the study. One issue that was

discussed was an apparently high number of non-working numbers that were appearing in the panel sample on callbacks. A strategy for dealing with these not-in-service numbers was developed and later implemented by the SRC. As described further below, the strategy included financial incentives and renewed efforts to identify the best contact information for respondents.

Postcards were sent that asked sample members to indicate the best telephone number to use to reach them, as well as the best times and days to call. Each time a completed post card was returned, the best contact number and preferred call times were entered into the computer as case notes for interviewers to review before making the call.

These post cards were first included in all third diary mailers, which went out on June 12, 2000. They were also included in all fourth diary mailers, which went out on August 21, 2000.

We offered two types of financial incentives to induce greater respondent cooperation: small gifts of thanks for past participation and larger offers of payment for continued participation. On June 28, 2000 and September 18, 2000, we mailed respondents letters that included gifts of \$5 as a token of appreciation for their participation in the study. On December 11, similar letters were mailed to respondents with gifts of \$5 or \$10. The amount a respondent received as a final token of appreciation was decided by random assignment. On September 22nd, those respondents that had previously refused to continue with the study were sent a letter offering them a \$50 incentive if they completed the entire schedule of interviews. A total of 69 respondents received this mailing. Only 4 respondents qualified for this incentive.

At the time of mailing the 5th diary materials, it was decided that we needed to use stronger methods to attempt to secure continued cooperation from sample members for whom we had non-working numbers. Two different letters were drafted and mailed on October 12, 2000, and again on December 11, 2000 with post cards and return envelopes. Which of the two types of letters a sample member received was based on random assignment. Both letters indicated that we had been trying to contact them to update our information about their recent beach experiences but

had been unsuccessful. Respondents were told that they were randomly selected, and their participation was critical to maintaining the validity of the study results. They were reminded of the option of calling our 800 number as well as the hours we could be reached. Both versions of the letter offered a \$20 incentive to respondents who completed the last three diary surveys. The first version stated that we would send \$20 to everyone who completed the three remaining interviews, and we hoped they would take us up on the offer. The second version stated that, if they dropped out of the study, they would lose the \$20 token of appreciation being offered to everyone who completed the three remaining interviews.

Letters offering some financial incentive to complete outstanding diary surveys were sent out to every respondent for which a "Phone Not Available" outcome was obtained during the fourth diary production. Fifty letters were mailed out at this time. This procedure was repeated when the 6th diary materials were mailed for a total of twenty-nine letters. In total, 79 individuals received the letter. However, none of the 79 actually completed all 6 diaries.

In addition, in June of 2000, it was decided that efforts should be made to try and acquire good phone numbers for persons who could no longer be contacted. Two methods were considered and tried. One (1) was to call directory assistance for each bad phone number, using the address information we had on file for the given respondent. The second method (2) was to use on-line directory assistance programs on the Internet. The first method was found to be somewhat effective, whereas the second was found to be ineffective. For the remainder of the study, directory assistance calls were periodically made to all not-in-service numbers.

PRODUCTION STATISTICS

Initial Recruitment:

Interviews for the initial recruitment survey were conducted from November 17, 1999 to January 9, 2000. A total of 7344 numbers received at least one call, and a total of 3855^{iv} numbers were identified as working residential numbers. A total of 1848 interviews were completed, 981 potential respondents refused to complete the interview, and 22 respondents completed only part of the interview.

A total of 1034 respondents had gone to the beach during the 12 months preceding the interview, of which 887 agreed to participate in the panel.

Disregarding refusals and partials, a total of 1790 cases representing potentially qualified households did not result in a completed interview by the completion of the recruitment process. A total of 794 cases had never been answered or had ended in a busy signal. Another 305 were answering machines. A total of 518 resulted in scheduled callbacks where the survey was never completed. A total of 172 cases represented households where no one could be found who spoke English or Spanish. There were 9 cases where a survey was completed, but the data was considered unreliable because the respondent had a diminished capacity^v to complete the survey, and therefore the data was not used. The response rate for the initial recruitment was 43%^{vi}. The average length of an interview was 8.4 minutes^{vii}.

The results of the recruitment survey are summarized below in Tables 1 and 2.

Table 1. Results of Recruitment Survey
Outcome

	Number	Percent
Completed interviews		
Completed Interviews	981	53.2
Other working residential numbers		
Refusals	22	1.2
Break Off	518	27.6
Scheduled Callbacks	305	16.3
Answering Machine	172	9.3
Language Barrier		
Bad Data		
Unknown if a residential number		
Never Answered		
Non-residential numbers		
Businesses		
Nonworking numbers		
Total		
	1848	100.0

9	
794	
1232	
1463	
7344	
	25.2
	13.4
	0.3
	7.1
	4.2
	2.3
	0.1
10.8	
16.8	
19.9	
100.0	

Table 2. Results of Recruitment Survey (Breakdown by eligibility)

Outcome	Number	Percent
Numbers fielded (Getting at least one call)		
Working residential numbers		
Businesses		
Nonworking numbers		
No answer		
7344		
3855		
1232		
1463		
794		
	100.0	
	52.5	

	16.8
	19.9
}	10.8

Working residential numbers

Eligible ^{viii} (One or more beachgoers)	
Ineligible (No beachgoers)	
Refused to complete screener plus those with undetermined eligibility	
3855	
1180	
838	
1837	
	100.0
	30.6
	21.7
	47.7

Eligible households

Recruited for panel	
Not recruited	
1180	
887	
293	
	100.0
	75.2
	24.8

Replenishment:

Replenishment was conducted from August 12, 2000 to October 8, 2000. At the start of the replenishment study, a sample 3414 potential respondents was released. From those, a total of 759 completed the interview, 519 refused, and 34 were partially completed. Of the 759 respondents who completed the replenishment survey, 299 were not eligible to participate in the panel because they were not beach goers. Of the remaining 460 respondents, 369 agreed to participate in the panel; 91 refused to participate.

Disregarding refusals and partials, a total of 948 cases representing potentially qualified households did not result in a completed interview by the completion of the replenishment process. A total of 378 cases had never been answered or had ended in a busy signal. Another 207 were answering machines. A total of 306 resulted in scheduled callbacks where the survey was never completed. A total of

54 cases represented households where no one could be found who spoke English or Spanish. There were 3 cases where a survey was completed, but the data was considered unreliable because the respondent had a diminished capacity to complete the survey, and therefore the data was not used. The response rate for the replenishment survey was 36%. The average length of an interview was 7.1 minutes.

The results of the replenishment survey are summarized below in Table 3 and Table 4.

Table 3. Results of Replenishment Survey
Outcome

	Number	Percent
Completed interviews		
Completed Interviews	759	
Other working residential numbers		
Refusals	520	
Break Off	35	
Scheduled Callbacks	306	
Answering Machine	207	
Language Barrier	54	
Bad Data	3	
Unknown if a residential number		
Never Answered	378	
Non-residential numbers		
Businesses	534	
Nonworking numbers	618	
Total	3414	

22.2

	15.2
	1.0
}	9.0
	6.1
	1.6
	0.1
11.1	
15.6	
18.1	
100.0	

Table 4. Results of Replenishment Survey (Breakdown by eligibility)
Outcome

	Number
	Percent
Numbers fielded (Getting at least one call)	
Working residential numbers	3414
Businesses	1884
Nonworking numbers	534
No answer	618
	378
100.0	
55.2	
15.6	
18.1	
11.1	
Working residential numbers	
Eligible (One or more beachgoers)	
Ineligible (No beachgoers)	
Refused to complete screener plus	
those with undetermined eligibility	1884
	530
	297
	1057
100.0	
28.1	
15.8	
56.1	
Eligible households	
Recruited for panel	
Not recruited	530
	369

161
100.0
69.6
30.4

Next, the production results for each of the six diary surveys are discussed. The results for the six dairy surveys are summarized at the end of the section in Table 3.

First Diary:

Interviews for the first diary were conducted from February 15, 2000 to April 12, 2000. After un-callable numbers, bad addresses, and cases used in the pretest were removed from the file, a total of 861^{ix} potential respondents were released into the sample. Of that number, 660 completed the first diary survey. Of those who did not complete the survey, only 31 refused to participate, and 5 were partially completed interviews.

An additional 165 failed to result in the completion of the first diary. In 13 of those, cases respondents were unavailable for contact, while 73 cases were instances where the respondent could not be reached at the phone number provided. A total of 4 cases had never been answered, and 22 were answering machines. Another 47 cases resulted in scheduled callbacks, which were never reached. Another 4 cases were determined to be unusable because the person originally qualified for the survey was, in fact, not eligible. In two cases, a survey was completed, but the data was deemed to be unusable because of a communication issue with the respondent. This resulted in a response rate of 77%^x. The average length of an interview was 5 minutes^{xi}.

Second Diary:

Diary two interviews were conducted from April 11, 2000 to August 10, 2000. After clean up^{xii}, a total of 873^{xiii} potential respondents were released into the sample. Of those, 638 completed the second diary survey. Of those who did not complete the survey, 57 refused to participate. An additional 5 completed part of the study.

Disregarding refusals and partials, there were 173 cases that did not result in a completed interview. In 28 of those cases, respondents were unavailable for contact, and in 55 cases, the telephone numbers attempted were unavailable.

A total of 29 cases had been answering machines; 9 cases were never answered; and 51 cases were scheduled callbacks, which were never reached. In one case, the telephone number turned out to be a business. This resulted in a response rate of 73% for this diary. A total of 71% of the original panel was interviewed for this diary. The average length of a survey was 6 minutes.

Third Diary:

Third diary interviews were conducted from July 21, 2000 to October 11, 2000. A total of 753 potential respondents were released into the sample. Of those, 509 completed the third diary survey. There were 47 who refused to participate in the survey, and 9 who partially completed the survey.

Another 188 cases failed to result in an interview at the completion of this diary. There were a total of 50 cases where the telephone number was unavailable and 17 instances where the respondent was unavailable for contact. There were 16 cases resulting in answering machines, 8 that were never answered, and 92 scheduled callbacks, which were never reached. There were 2 numbers that went to a business phone, 1 case where the respondent was determined to not be qualified for the study, and 2 cases where an interview was completed, but the data was determined to be unreliable due to communication issues with the respondent. This resulted in a response rate of 68%^{xiv} for this diary. A total of 57% of the original panel was interviewed for this diary. The average length of an interview was 15.1 minutes.

Fourth Diary:

Data collection for the Fourth Diary was conducted from August 23, 2000 to October 16, 2000. After cleanup, 720 potential respondents were released into the sample. Of those, 529 completed the fourth diary survey and 27 refused to participate. At the completion of the diary, there were 3 partially completed interviews.

A total of 161 cases remained unresolved at the completion of diary four. Of those, 9 were answering machines, 3 had never been answered, and 64 were scheduled callbacks, which were never reached. In 35 instances, the respondent was unavailable for contact, and in 45 cases, the telephone number was incorrect. One telephone number went to a

business, and an additional 4 cases were completed where the data was determined to be unreliable because of communication issues with the respondent. This resulted in a response rate of 73% for this diary. A total of 60% of the original panel was interviewed for this diary. The average length of an interview was 6.1 minutes.

Fifth Diary:

Interviews for diary five were conducted from October 20, 2000 to December 10, 2000. After cleanup, there were 1002 potential respondents released into the sample^{xv}. From those, 731 respondents completed the fifth diary interview. There were 42 who refused to complete the interview, and 2 where the interview was partially completed.

At the completion of the diary five interview, there were 227 additional cases that did not result in an interview. Of those, 91 were scheduled callbacks, which were never reached, 9 were answering machines, 5 were never answered, and 6 were business numbers. There were a total of 80 cases where the telephone number was unavailable, and 33 where the respondent was unavailable for contact. An additional 3 cases were completed where the data was determined to be unreliable because of communication issues with the respondent. This resulted in a response rate of 73% for this diary. A total of 476, or 54%, of the original panel completed an interview for this diary. The average length of an interview was 7.6 minutes.

Sixth Diary:

Interviews for diary six were conducted from December 13, 2000 to January 26, 2001. There were 911 potential respondents released into the sample. Of those, 720 respondents completed the interview, and 21 refused to participate.

At the completion of diary six, there were 170 additional cases that did not result in a completed interview. Of those, 74 were instances where the telephone number was unavailable and 2 where the respondent was unavailable for contact. There were a total of 15 cases ending in answering machines, 2 were never answered, and 63 cases were scheduled callbacks, which were never reached. A total of 11 numbers were determined to be businesses. A single additional case was determined to belong to

someone who did not qualify for the study. An additional 2 cases were completed where the data was determined to be unreliable because of communication issues with the respondent. This resulted in a response rate of 79% for this diary. A total of 461, or 52%, of the original panel completed an interview for this diary. The average length of an interview was 3.8 minutes.

A total of 306, or 34% of the original panel, completed all six diaries.

Results for each of the six diary surveys are summarized below in the Table 3.

Table 5. Results for Diaries

	First Diary		Second Diary		Third Diary		Fourth Diary		Fifth Diary		Sixth Diary	
	n	%	n	%	n	%	n	%	n	%	n	%
Number Fielded	861	100.0	873	100.0	753	100.0	720	100.0	1002	100.0	911	100.0
Completed Diary	660	76.7	638	73.1	509	67.6	529	73.5	731	73.0	720	79.0
Refusal	31	3.6	57	6.5	47	6.2	27	3.8	42	4.2	21	2.3
Break Off	5	0.6	5	0.6	9	1.2	3	0.4	2	0.2		
Non-contact												
Respondent Unavailable	13	1.5	28	3.2	17	2.3	35	4.9	33	3.3	2	0.2
Incorrect Phone Number	73	8.5	55	6.3	50	6.6	45	6.3	80	8.0	74	8.1
Never Answered	4	0.5	9	1.0	8	1.1	3	0.4	5	0.5	2	0.2
Other Non-response												
Answering Machine	22	2.6	29	3.3	16	2.1	9	1.3	9	0.9	15	1.6
Scheduled Callback	47	5.5	51	5.8	92	12.2	64	8.9	91	9.1	63	6.9
Not Qualified	4	0.5	1	0.1	3	0.4	1	0.1	6	0.6	12	1.3
Bad Data	2	0.2			2	0.3	4	0.6	3	0.3	2	0.2
Field period for the diary	2/15/00 – 4/12/00		4/11/00 – 8/10/00		7/21/00 – 10/11/00		8/23/00 – 10/16/00		10/20/00 – 12/10/00		12/13/00 – 1/26/01	

INTERVIEWER TRAINING

All interviewers at the SRC receive a minimum of 12 hours of training before being allowed to conduct interviews on any project. This training consists of instruction in the operation of the SRC's CATI system, but also includes extensive instruction in the proper methods for conducting a scientific interview, as well as avoiding refusals.

Components of this instruction are:

1. Reading questions accurately.
2. Conducting the interview in a neutral manner.
3. Selecting respondents within households.
4. The proper use of probes in avoiding item refusals.
5. The proper use of probes in convincing potential respondents to complete the survey.
6. Methods of ensuring interviews are consistent from one respondent to another and one interviewer to another.
7. Sample management.

At the completion of training, each interviewer is monitored while conducting an mock interview with an SRC supervisor and must pass a written exam before being allowed to conduct interviews for the SRC

Each interviewer who worked on the Southern California Beach Study also received four to six hours of additional training on the project itself. Linwood Pendleton conducted the initial training, while Allen Lunde and SRC supervisors conducted additional trainings. These trainings included:

1. An overview of the purpose and goals of the project.
2. Methodology of household enumeration specific to this project.
3. Extensive training regarding the location and nature of the various beaches involved in the study.
4. Study specific probes.
5. Refusal avoidance specific to the study.

All interviewers who participated in the Southern California Beach Study were required to pass a quiz about beaches in Southern California before they were allowed to conduct interviews on this project.

At the beginning of each diary phase of the study, callers

received an additional briefing covering the particulars of that wave.

Throughout the study, interviewers were periodically monitored to ensure interviews were being conducted in a scientifically- rigorous manner.

DATA SET CONSTRUCTION

The completed interview data from CASES for each of the six waves of data collection were output into SPSS data files. Wave 3 served as a model for the other waves (except for Wave 1). After Wave 1, all interviews utilized a rostering system where each line became a "trip," which could be a single day or multiple day event and could include one or more beaches visited. In other words, the original data files allowed for multiple beaches to be listed within a single record. The goal was to transform the original data set to one in which the data for each beach visited during a particular trip would appear on a single line. The new data set used the data for an individual trip by a respondent to a single beach as the unit of analysis. The first line for each respondent within the new matrix also included summary data like the number of trips during the two-month period, as well as the data from the modules which relate directly to the respondent rather than to the individual trips to the beach.

The variable WT was created to indicate what category of trip the visit to that particular beach belonged. Again, the five trip modes were:

1. A single beach on a single day.
2. Multiple beaches on a single day.
3. A single beach over multiple days.
4. Multiple beaches over multiple days.
5. Multiple beaches gone past in a single day due to a recreational mode of transportation (i.e. bicycling past several beaches).

The plan developed to accomplish the transformation of the raw data to the desired final format is represented in the spreadsheet entitled Beach Plan. The Beach Plan spreadsheet shows precisely how the various beach trip data were reorganized. A separate, detailed document explains in detail how the data transformation proceeded.

Only a brief description is offered here.

Data from the original SPSS dataset were divided into two Excel spreadsheets (use data and module data) to compensate for the columnar limitations of Excel (256 columns). Unique identifiers were preserved for each record in both sets so an accurate merge could be accomplished after the use data had been transformed. The data from each record within the spreadsheet were examined for the existence of multiple beaches. If multiple beaches were found, a new record was created bearing the respondent ID coding, date and trip type. Data regarding that particular beach and its use were moved to the appropriate columns. If a particular record came from respondents who reported more than 5 trips within the two-month period, the data on beach usage came from the "Typical Day" data. Occasionally, there were beach names listed in this "Typical Day" data that did not correspond to any of the Beach Numbers listed earlier in the records for that respondent. These data were ignored. Also on occasion, the "Typical Day" data for a specific Beach Number was missing. The process of moving data into the appropriate columns of the matrix continued row by row until the entire matrix had been transformed.

Identical procedures were performed on the data from Waves 2, 4, 5, and 6. For the data generated by the Replenishment Survey, additional calculations were needed to calculate an approximate date for beach trips for which the respondent could not remember the exact date. Data taken from responses on questions regarding which week of the month and weekday/weekend designation were used to obtain a correct range of days for that trip. A systematic cycling through the days for a designated period was employed so that no erroneous within-week trends would be created in the approximation process.

In the first wave (Diary 1), the original data set was constructed to contain all the data for a single interview in a single record that allowed for multiple trips to the beach over the two-month period. This original record contained up to 2153 variables. Examination of the frequencies from the original SPSS data file indicated that 1633 of these variables contained no data and were eliminated from the data set. A new matrix was constructed which redefined the meaning of the individual record. The original matrix

used the respondent as the unit of analysis.

Since in this first wave (paper records) precise data indicated if a trip extended beyond a single day, special codes for the WT needed to be added to indicate that ambiguity. The additional two modes were:

6. Single beach / ? day
7. Multiple beach / ? day.

After each "use" dataset for a wave was completed, it was merged with the corresponding "module" dataset using SPSS 9.0. The resulting SPSS data file was then converted back to a single Excel file and shipped to Allen Lunde and his assistants for error checking.

APPENDIX A

OUTCOMES FOR ALL PANEL MEMBERS FOR ALL DIARIES

Table 6: Key To Codes Used In Reporting

SRC Code	SRC Category	AAPOR Category	AAPOR Code
COMP	Completed Interview	Complete Interview	I
AM	Answering Machine	Non-contact	NC
NA	Never Answered	Unknown if household/occupied HU	UH
NIS	Phone Not Available/Wrong Number	Not Eligible	NONE
PART	Partial Interview	Refusal or Break-off	R
NPER	Respondent Not/No Longer Available	Non-contact	NONE
PRE	Case Used in Pretest	Not Eligible	NONE
REF	Refusal	Refusal or Break-off	R
BDAT	Bad Data	Non-contact	NC
BUS	Business	Not Eligible	NONE
NSMP	Not in Sample	Not Eligible	NONE
SCB	Scheduled Callback	Non-contact	NC

This appendix has been deleted due to its length. It is available upon request.

APPENDIX B

PANEL
CORRESPONDENCE

March 22, 2000

«SEX» «MLNM» «MLNM1»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear : «SEX» «MLNM» «MLNM1»

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during December and January.

Meanwhile, for your convenience, we are enclosing a calendar covering the months of February and March, together with another copy of our map of the beaches in Orange and Los Angeles Counties.

If you go to the beach in San Diego, Orange, Los Angeles, Ventura or Santa Barbara Counties during February or March, please could you note the name of each beach you visit in the appropriate date block of the enclosed calendar. If you get into the water during a beach visit, please write a "W" beside the name of the beach on the calendar.

When we call soon to ask you about your beach trips in December and January, please have your December and January calendars handy.

If you have any questions about our study, please feel free to call us toll-free during regular business hours at 1-800-477-8783.

We want to thank you for your assistance with this very important research project. We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

May 24, 2000

«TITL» «FNAM» «LNAM»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear : «TITL» «LNAM»

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during February and March .

Meanwhile, for your convenience, we are enclosing a calendar covering the months of April and May, together with another copy of our map of the beaches in Orange and Los Angeles Counties.

If you go to the beach in San Diego, Orange, Los Angeles, Ventura or Santa Barbara Counties during April or May, please could you note the name of each beach you visit in the appropriate date block of the enclosed calendar. If you get into the water during a beach visit, please write a "W" beside the name of the beach on the calendar.

When we call soon to ask you about your beach trips in February and March, please have your February and March calendars handy.

If you have any questions about our study, please feel free to call us toll-free during regular business hours at 1-800-477-8783.

We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

June 12, 2000

Dear <titl> <lnam>:

We want to thank you for your continued help with the Southern California Beach Survey. We would like to give you an update on our progress.

You are one of approximately 900 residents of Southern California who are participating in this survey. This small group was selected using scientific random sampling to obtain a statistically valid sample of the much larger population of beach users in Southern California. The data that you and others provide will be used to analyze how beach recreation contributes to the Southern California economy and to estimate the economic value that residents of Southern California get from the beaches there. The information that you provide on your beach use is very important to the success of the survey and the accuracy of the results.

What we have learned so far is that about 40% of the panel members went to the beach at least once in February or March. Of those who went to the beach, about 18% went to more than 5 beaches during this period. We are now compiling a list of the most popular beaches in the survey.

We are currently contacting the panel to ask about beach use in April and May. However, we have found that some panel members have moved or have been assigned new area codes or telephone numbers, and we have been unable to reach them.

Our study will remain valid only if we can continue to reach all panel members. We are asking you to provide us your current phone number and address. We have enclosed a self-addressed, stamped envelope for this purpose. One of our interviewers will call you at the time you suggest a few days after we hear from you.

This is an important study, and it has received a lot of publicity. You may have seen articles about this study in the newspapers. For your information, we have enclosed copies of articles which appeared in the *Los Angeles Times* and *Long Beach Press Telegram*. They provide more information on the survey and its importance to Southern California.

The agencies funding the study have asked that we continue collecting data on beach use through November. This will provide data for a full year of beach use. We are grateful for the help you have given us so far, and we sincerely hope that you will continue to help us.

If you have questions about the survey, please call Allen Lunde, Nicole Thompson, or me toll-free at 1-800-477-8783. You can also get information about the study, including extra copies of the maps and calendars, at our web site: <http://www.jps.net/katsii> .

Sincerely,

James E. Fletcher, Ph.D., Director

5 enclosures: post card, map, calendar, and 2 newspaper clippings.

June 28, 2000

Dear «TITL» «FNAM» «LNAME»:

We want to thank you for your continued help with the Southern California Beach Survey. We would like to give you an update on our progress.

You are one of approximately 900 residents of Southern California who are participating in this survey. This small group was selected using scientific random sampling to obtain a statistically valid sample of the much larger population of beach users in Southern California. The data that you and others provide will be used to analyze how beach recreation contributes to the Southern California economy and to estimate the economic value that residents of Southern California get from the beaches there. The information that you provide on your beach use is very important to the success of the survey and the accuracy of the results.

What we have learned so far is that about 40% of the panel members went to the beach at least once in February or March. Of those who went to the beach, about 18% went to more than 5 beaches during this period. We are now compiling a list of the most popular beaches in the survey.

We are currently contacting the panel to ask about beach use in April and May. However, we have found that some panel members have moved or have been assigned new area codes or telephone numbers and we have been unable to reach them.

Our study will remain valid only if we can continue to reach all panel members. We are asking you to provide us your current phone number and address. We have enclosed a self-addressed, stamped envelope for this purpose. One of our interviewers will call you at the time you suggest a few days after we hear from you.

This is an important study and it has received a lot of publicity. You may have seen articles about this study in the newspapers. For your information, we have enclosed copies of articles which appeared in the *Los Angeles Times* and *Long Beach Press Telegram*. They provide more information on the survey and its importance to Southern California.

The agencies funding the study have asked that we continue collecting data on beach use through November. This will provide data for a full year of beach use. We are grateful for the help you have given us so far and we sincerely hope that you will continue to help us. We are enclosing a \$5.00 bill as a token of our gratitude to you for helping us by participating in this important study.

If you have questions about the survey, please call Allen Lunde, Nicole Thompson, or me toll-free at 1-800-477-8783. You can also get information about the study, including extra copies of the maps and calendars, at our web site: <http://www.jps.net/katsii>.

Sincerely,

James E. Fletcher, Ph.D., Director

5 enclosures: post card, map, calendar, and 2 newspaper clippings.

September 18, 2000

«TITL» «FNAM» «LNAM»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear : «TITL» «LNAM»

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during June and July.

Meanwhile, for your convenience, we are enclosing a calendar covering the months of August and September, together with another copy of our map of the beaches in Orange and Los Angeles Counties.

If you go to the beach in San Diego, Orange, Los Angeles, Ventura or Santa Barbara Counties during April or May, please could you note the name of each beach you visit in the appropriate date block of the enclosed calendar. If you get into the water during a beach visit, please write a "W" beside the name of the beach on the calendar.

When we call soon to ask you about your beach trips in June and July, please have your June and July calendars handy.

We are enclosing a \$5.00 bill as a token of our gratitude to you for helping us by participating in this important study.

If you have any questions about our study, or have changed your contact phone number, please call us toll-free during regular business hours at 1-800-477-8783. For your convenience, you may also provide us a phone number, as well as the best days and times to reach you by filling out the enclosed card and mailing it in the postage paid envelope provided in this package. You can also get information about the study, including extra copies of the maps and calendars, at our web site: <http://www.jps.net/katsii> .

We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

October 12, 2000

«TITL» «FNAM» «LNAM»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear : «TITL» «LNAM»

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during August and September .

Meanwhile, for your convenience, we are enclosing a calendar covering the months of October and November, together with another copy of our map of the beaches in Orange and Los Angeles Counties.

If you go to the beach in San Diego, Orange, Los Angeles, Ventura or Santa Barbara Counties during October or November, please could you note the name of each beach you visit in the appropriate date block of the enclosed calendar. If you get into the water during a beach visit, please write a “W” beside the name of the beach on the calendar.

When we call soon to ask you about your beach trips in August and September, please have your August and September calendars handy.

If you have any questions about our study, please feel free to call us toll-free during regular business hours at 1-800-477-8783.

We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

October 12, 2000

<titl> <fnam> <lnam>
<add1>
<add2>, <add3> <add4>

Dear : <titl> <lnam>

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during October and November. We are pleased to announce that this next interview marks the completion of the data collection portion of this project.

The data you have provided us has proven extremely valuable. We are enclosing a \$10.00 bill as a token of our gratitude to you for helping us by participating in this important study.

If you have any questions about our study, or have changed your contact phone number, please call us toll-free during regular business hours at 1-800-477-8783. You can also get information about the study, including extra copies of the maps and calendars, at our web site: <http://www.jps.net/katsii> .

We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

October 12, 2000

«TITL» «FNAM» «LNAM»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear : «TITL» «LNAM»

Thank you for your continuing participation in the Southern California Beach Study. We will be telephoning you in the next few weeks to ask about your use of the beaches in Southern California during October and November. We are pleased to announce that this next interview marks the completion of the data collection portion of this project.

The data you have provided us has proven extremely valuable. We are enclosing a \$5.00 bill as a token of our gratitude to you for helping us by participating in this important study.

If you have any questions about our study, or have changed your contact phone number, please call us toll-free during regular business hours at 1-800-477-8783. You can also get information about the study, including extra copies of the maps and calendars, at our web site: <http://www.jps.net/katsii> .

We look forward to talking with you in the near future.

Sincerely,

James E. Fletcher, Ph.D.
Professor and Director

Enclosures: 2

October 12, 2000

«TITL» «FNAM» «LNAM»
«ADD1»
«ADD2», «ADD3» «ADD4»

Dear «TITL» «LNAM»:

My colleagues and I are very grateful to you for the help you have given us so far on the Southern California Beach Study. Right now we are collecting information from panel participants for the fourth time. We realize the study has imposed quite a burden on you and the other participants, and we are enclosing \$5 as a token of our appreciation for your help to date.

Over the next few months, we will be completing the study. As you may know, the study has already attracted a lot of attention from the press around the state. We are enclosing a few clippings that illustrate the publicity the study has received so far. You can be sure that state officials and policy makers will take the final results of the study very seriously. We believe that the Southern California Beach Study will be a landmark investigation, shaping policy and research on the beaches of Southern California and elsewhere for years to come.

The scientific value of a study like ours depends on the continuing cooperation of the study participants. That is why I'm writing to you. We understand that you may be reluctant to continue to take part in the study. We have tried hard to keep our questions brief and to the point, but we realize how time-consuming they must seem to you. Because it is important for us to track use of Southern California's beaches throughout the year, we would like you to stick with us by completing the final three diaries. To underscore the importance of your continued participation, we are offering an incentive of \$50 to participants who complete the entire schedule of interviews. We will mail you a check for that amount at the conclusion of the study. If you choose not to continue with the study, please return the enclosed sheet.

If you have any questions about the study, please don't hesitate to call us at our toll-free number (1-800-477-8783). Please help us provide the best information to California's decision-makers.

Sincerely,

Michael Hanemann
Principal Investigator
Southern California Beach Study

Southern California Beach Survey

Name: _____ **Case ID:**

I decline to continue with the Southern California Beach Study.
Please do not call me anymore.

December 11, 2000

«titl» «fnam» «lnam»
«add1»
«add2», «add3» «add4»

Dear «titl» «lnam» :

Thank you for taking part in our study of the Southern California beaches. We have been trying to call you to interview you to update our information about your recent beach experiences. However, we have not been able to reach you.

We need your help to finish our study. You are one of a randomly selected group of beach users, and your participation is critical to maintaining the validity of the results. We realize we are asking a lot of you, and we appreciate your cooperation to date. As a token of our appreciation, we will SEND \$20 to everyone who completes the three remaining interviews. We hope you take us up on this. Would you please take a moment to provide us your current phone number as well as the best days and times to reach you? Just fill out the enclosed card and mail it to us in the postage paid envelope provided in this package.

You may also call us at 1-800-477-8783 to complete the survey at your convenience any time Monday through Friday between 6 PM-9 PM, Saturday 10 AM-1 PM and 2 PM-5 PM, or Sunday 2 PM-5 PM and 6 PM-9 PM. We greatly appreciate your assistance and look forward to receiving your input.

Sincerely,

James E. Fletcher, Ph.D., Director

Enclosure: Post Card

December 11, 2000

«titl» «fnam» «lnam»
«add1»
«add2», «add3» «add4»

Dear «titl» «lnam» :

Thank you for taking part in our study of the Southern California beaches. We have been trying to call you to interview you to update our information about your recent beach experiences. However, we have not been able to reach you.

We need your help to finish our study. You are one of a randomly selected group of beach users, and your participation is critical to maintaining the validity of the results. We realize we are asking a lot of you, and we appreciate your cooperation to date. However, if you drop out of the study now, you will lose the \$20 token of appreciation we will give to everyone who completes the three remaining interviews. It would be a shame for you to miss out on this. Would you please take a moment to provide us your current phone number as well as the best days and times to reach you? Just fill out the enclosed card and mail it to us in the postage paid envelope provided in this package.

You may also call us at 1-800-477-8783 to complete the survey at your convenience any time Monday through Friday between 6 PM-9 PM, Saturday 10 AM-1 PM and 2 PM-5 PM, or Sunday 2 PM-5 PM and 6 PM-9 PM. We greatly appreciate your assistance and look forward to receiving your input.

Sincerely,

James E. Fletcher, Ph.D., Director
Enclosure: Post Card

Southern California Beach Survey

Name: _____ Case ID: _____

The best telephone number to call to reach me is:

(____) _____
(Area Code & Phone Number)

The best days and times to reach me are:

MON	TUES	WED	THUR	FRI	SAT	SUN
AM	AM	AM	AM	AM	AM	AM
AM	AM	AM	AM	AM	AM	AM
AM	AM	AM	AM	AM	AM	AM

PM	PM	PM	PM	PM	PM	PM
PM	PM	PM	PM	PM	PM	PM
PM	PM	PM	PM	PM	PM	PM

**PLEASE MAIL THIS BACK TO US IN THE POSTAGE PAID ENVELOPE
PROVIDED. THANK YOU.**

APPENDIX C

EXAMPLES OF THE WEB PAGES


APRIL 2000 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites History

Address G:\STUDIES\Search Data\New 2002\Web Page\apr1000.htm Go Links

APRIL 2000

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	Survey Research Center California State University, Chico 800-477-8783		

Done LocalIntranet

Start Web Page APRIL 2000 - Microsoft Word Microsoft Word 1:25 PM

APRIL 2000 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

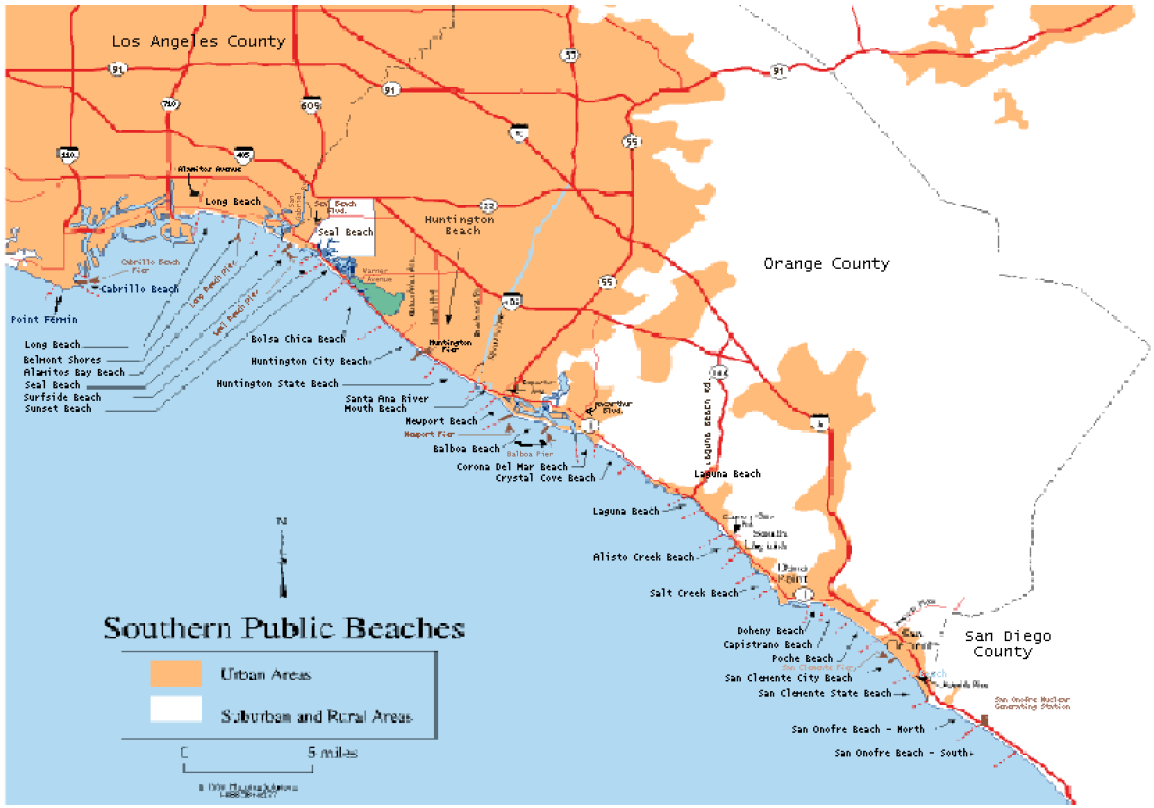
Back Forward Stop Search Favorites History

Address G:\STUDIES\Search Data\New 2002\Web Page\apr1000.htm Go Links

Done LocalIntranet

Start Web Page APRIL 2000 - Microsoft Word Documents - Microsoft Word 1:25 PM





APPENDIX D

EXPLANATION OF RESPONSE RATES

The diary surveys use RR1. According to AAPOR, "**Response Rate 1 (RR1)**, or the minimum response rate, is the number of completed interviews divided by the number of interviews (complete plus partial) plus the number of non-interviews (refusals plus non-contact^{xvi}) plus all cases of unknown eligibility^{xvii}". Cases deemed not eligible^{xviii} are not part of the equation

$$\frac{I}{I + P + NC + UH}$$

For the column designated "First Diary" the equation would be:

$$\frac{660}{660 + 5 + (31+75) + (13+73+4)} = 77\%$$

According to AAPOR, "**Response Rate 3 (RR3)**, estimates what proportion of unknown eligibility cases are actually eligible. In estimating e, one must be guided by the best available scientific information on what share eligible cases make up among the unknown cases and one must not select a proportion in order to boost the response rate". For the purpose of this survey, the SRC uses the proportion of eligible to ineligible cases to calculate e^{xix}. Cases where eligibility is not determined are excluded from the calculation. In the Initial Recruitment directory there were 6550 cases released where eligibility was determined. Of that there were 3855 eligible cases. This means that 58.9% of the cases released into the Initial Recruitment directory sample, where eligibility could be determined, were eligible. As a result e is determined to be 58.9%. The entire equation for RR3 is:

$$\frac{I}{I + P + NC + eUH}$$

For the Initial Recruitment the equation would be:

$$1848 + 22 + \frac{1848}{(981+518+305+172+9) + (.589*794)} = 43\%$$

End Notes

APPENDIX 2. Detailed list of errors and changes.

The format of this reporting of errors and changes is as follows:

Identifying information: Explains which records are being addressed.

Problem: Explains the error finding.

Action: Explains the correction and the reasoning behind correction if any action has been taken.

Two members of the research team—Chris Busch and Craig Mohn—have had primary responsibility for the team’s data checking. Each compiled lists of errors and corrections. Chris Busch’s findings are detailed first., followed by those of Craig Mohn.

****Chris Busch’s Data Checking Findings Begin Here****

Identifying information: Wave 1

- 10744, 186 (format is: CSID, RECID)
- 10744, 187 (format is: CSID, RECID)
- 12382, 465 (format is: CSID, RECID)
- 12382, 466 (format is: CSID, RECID)
- 13679, 769 (format is: CSID, RECID)
- 18860, 1442 (format is: CSID, RECID)
- 18891, 1447 (format is: CSID, RECID)
- 19244, 1522 (format is: CSID, RECID)
- 19244, 1523 (format is: CSID, RECID)

Problem: Start and end dates show a one day trip (i.e. they are the same day) and the WT variable shows a multi day trip (WT = 2).

Action: WT changed to =1 in every case.

Note SRC comment: All cases double checked and found same start and end date. These situations arose from mismatch between data entered in Q4MT and Q6 (month start) and Q4 and Q6A (day start) located in different parts of the matrix. They were supposed to carry identical data but in some cases did not. Also there were cases listed in Q4A (multiday trip?) as yes but had same start and finish dates. When WT was assigned Q4A was used to make the designation. Corrections to the dataset as above

Identifying information: Wave 1: CSID 15997 (RECID 1152-1168) and CSID 19943 (RECID 1620-1641).

Problem: Problem with parking costs, variable Q4I in data delivered by Chico SRC. There are \$75 values that fall outside the range of acceptable answers. (Up to \$40 was allowed.)

Action: Change parking values from 75 to 0 (zero). Chico explanation: “ These two cases reflect persons who paid \$75/yr for parking. Examination of the next several waves indicates that they list no fees for parking.”

Identifying information: wave 1, RECID 1655 (CSID 19986)

Problem: Data show (using Chico variable labels with explanation)

Q4MT (start month)	Q4 (start day)	Q6B (start month)	Q6DY (start day)
1	29	30	

Return date (month—Q6B, and day—Q6DY) obviously wrong

Action: Make change as follows

Q4MT (start month)	Q4 (start day)	Q6B (start month)	Q6DY (start day)
1	29	1	30

Correction based on research team suggestion. SRC agreement (“suspect data entry error, dataset corrected to reflect 2day trip using your suggestion.”)

Identifying information: Wave 2, RECID 454, CSID 12197

Problem: Trip date error.

Q4MT (start month)	Q4 (start day)	Q6B (start month)	Q6DY (start day)
2	18	1	18

This shows trip ending before it started. Reversing start and end dates is also problematic. The vast majority of multi day trips are only two days and longest are not much more than a week. A month long trip does not seem plausible.

Action: WT changed to one (single day) and date changed to 2/18. Chico: “original data (q6,q6a,q6b,q6dy) show 1/18 - 1/18 (out of proper date ranges), likely 2/18-2/18. changed dates to 2/18-2/18 and WT changed to 1.”

Identifying information: Wave 2, RECID 1004, CSID 15502

Problem:

Q4MT	Q4	Q6B	Q6DY
3	11	11	11

This shows an 8 month trip.

Action. Change end date of the trip to unknown. Following Chico: “that is how the orig. data reads. suggest making the ending date 88888/88888.”

Identifying information: Wave 2, CSID 17161, RECID 1198.

Problem: Logical contradiction.. Shows a trip of more than one month and WT value = 1 (single day trip). Changing month of return makes the trip 4 days, but this still contradicts WT value of 1.

Action. Delete use data based on Chico recommendation. Chico: “A bogus trip! Terminated by interviewer with beach Code 999 contains no use data. Response in Q1AA=5 (no trips) Dates and WT removed from dataset.”

Identifying information: Wave 2, CSID 16098, RECID 1090.

Problem. Scrutiny of outlier trip. Research Team observation: “End date indicates trip length of one month and one day, and this end date falls outside of diary period. Perhaps the end date month should be changed to March (3) from April (4). This would make the trip an overnighter.” Chico: “that is what orig data say! I agree w/ you dates changed to 3/12-3/13.”

Action. Change return date (month) to March from April.

Identifying information. Wave 2. CSID and RECID as indicated below.

Problem. Cases where start and end dates show a one day trip (they are the same) and the WT variable shows a multi day trip.

Action. Based on Chico recommendations...

- CSID 12197 (RECID 454). WT changed to 1
- CSID 12888 (RECID 621). WT changed to 1
- CSID 12371 (RECID 479). WT changed to 1
- CSID 12888 (RECID 622). WT changed to 1
- CSID 14297 (RECID 843). WT changed to 1
- CSID 14312 (RECID 856). WT changed to 1
- CSID 15539 (RECID 1012). WT changed to 1
- CSID 15550 (RECID 1013). WT changed to 1
- CSID 14638 (RECID 893). **Date changed to 3/17-3/18.**

Identifying information. Wave 2, CSID 17035 RECID 1179

Problem. WT shows multi beach trip but no second beach listed.

Action. WT Changed to 2 to indicate single beach trip

Identifying information. Wave 2, CSID 13654 RECID 755

Problem. Chico SRC identifies problem with dates for trip.

Action. Following Chico SRC: “Dates for multibeach/multiday trip changed from 3/4-4/5 to 3/4-3/5 based on info from other trips for that respondent.”

Identifying information. Wave 2, CSID 17035 RECID 1180

Problem. Identified by Chico.

Action. Following Chico SRC: “record dropped for lack of data.”

Identifying information. Wave 2, CSID 19082, RECID 1289-1292.

Problem. These four records show four 1 day, 1 beach trips all occurring on the same day (indicated by WT = 1 for all four rows). This seems to imply that these four trips should be recoded as a 1 day, multi beach trip.

Action. WT changed to 3 for these records. Chico agrees with intuition, “Chico: Roster indicates separate trips but I agree with your WT changed to 3 for all 4 trips based on hours arrived at beach and time spent there.”

Identifying information. Wave 2, CSID 15815, RECID 1033.

Problem. Trip date indicates unreasonably long trip. Research Team observation: Q6B (return date month) given as 6, but looks like it should be 2.

Action. Change return month to February. Chico comment, “Yes, agree with your suggested change.”

Identifying information. Wave 2. CSID 12734, RECID 552.

Problem. RE: parking cost, Q4I. The 32 value looks like an error. This is a 4 person, one day trip to Oceanside city beach. RECID 551 is for the same respondent and for a trip with a similar profile, but the cost of parking was only \$2 for RECID 551. Probably the same for RECID 552.

Action. Change value to \$2. Chico comment: “agree! Q4I changed to \$2”

Identifying information. Wave 3. CSID 18387, RECID 855-856.

Problem. These records show same start and end date but WT =2.

Action. WT changed to 1 for both. Chico: agree that WT should be changed to =1.

Identifying information. Wave 3, all cases.

Problem. There are no values for Q6B, Q6DY = no end dates, implies = no multi day trips.

Action. “Chico: new dataset should clear that up. several start/finish dates in set seemed unreasonable. All corresponded to data as entered. Mo. adjusted to make trip last only several days.(e.g. 5/6-7/7 changed to 5/6-5/7)” (Research team note: We didn’t observe the initial form of errors mentioned in Chico’s response, and so no quantitative measure is possible.)

Identifying information. Wave 4. CSID 14898 (RECID 757).

Problem. Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.

Action. Following Chico SRC: “date correction, changed to 7/5-7/6 as per original data.”

Identifying information. Wave 4. CSID 15522 (RECID 836).

Problem: Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.

Action. Following Chico SRC: “date correction, changed to 7/1-7/4 as per original data”

Identifying information. Wave 4. CSID 12350 (RECID373).

Problem. End date listed is one day before start date.

Action. Following Chico SRC: “date correction, changed to 7/20-7/22 as per original data.”

Identifying information. Wave 5. CSID 10119, RECID 13,-(missing),14, , rows 18-20 in original data delivered by Chico.

Problem. Data show. (format is month/day of departure/return)

Q4MT	Q4	Q6B	Q6DY
8	30	9	4
8	30	9	4
9	1	9	4

The first two rows are for a multi beach, multi day trip.

The third row is for a multi beach, single day trip.

The problem is that these data indicate that these two apparently distinct trips overlap.

Action. Following Chico SRC: “Since the Beach code is the same for row 19 & 20, no info is lost by dropping row 20. Resolve by dropping row 20.”

Identifying information. Wave 5. CSID 31755 (RECID 1355).

Problem. Shown to be a trip date problem. Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a

multi-day trip.

Action. Following Chico SRC: “date changed to 8/24-8/25.”

Identifying information. Wave 5. CSID 16258 (RECID 806).

Problem: Shown to be a trip date problem. Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.

Action. Following Chico SRC: “date changed to 8/4-8/6.”

Identifying information. Wave 5. CSID 10515 (RECID 70).

Problem. Shown to be a trip date problem. Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.

Action. Following Chico SRC: “date changed to 8/12-8/13.”

Identifying information: Wave 5. CSID 30638 (RECID 1116).

Problem. End date listed is one day before start date.

Action. Following Chico SRC: “date changed to 8/9-8/10.”

Identifying information. Wave 5. CSID 30748, RECID 1152.

Problem. Shows one day trip (WT=1) but data on hours at beach (Q4G) reads 24, which implies an overnight trip.

Action. Following Chico SRC: “assume 24 hrs is a typo replaced with 88888”

Identifying information. Wave 5. CSID 19986 (PREID 1050).

Problem: Research team previous observation: “End date listed is in month before start date.” Chico, initial response: date changed to 8/9-8/17.

However, There’s still a problem here. The change means that the trip detailed in this row is in fact contained in the trip detailed in the two rows above. Suggest we change WT and dates so that this is a beach visited on the larger trip outlined in the two rows above.

Action. Change WT to indicate multi day, multi beach trip with agreement from Chico SRC.

Identifying information. Wave 5.

Problem. Beach names should be listed in Chico variables q18, q22. Q22c indicates respondent remarks should be recorded (e.g. spelled out), but data show only 0s? Respondent remarks were also missing for Q23a, Q24, Q25, Q26, and Q27.

Action. Chico SRC supplied these upon request.

Identifying information. Wave 6.

Problem. There is no data in Q4MT, which shows the month component of the trip start date.

Action. Chico SRC found and added this data. (Not quantified.)

Identifying information. Wave 6. CSID 18407 (RECID 650).

Problem. Return month, Q6 in this case, shows May (5), which is beyond diary time period and a likely error.

Action. Following Chico SRC: “changed to 11 based on starting mo.”

Identifying information: Wave 6. CSID 14884

Problem. Can’t tell how long the trip is because of missing month, but return date is January 1, and interviewer includes comment in BC variable, “2 mo.! Don’t think so.”

Action. Following Chico SRC: “changed ending date to 88888/88888.”

Identifying information: Wave 6. CSID 10805 (RECID 88-89).

Problem. Record shows two 1 day, 1 beach trips that occur on the same day. This seems to imply WT should be recoded so that these are both part of the same 1 day, multi beach trip.

Action. Following Chico SRC: “rostered as separate trips, but examination of times indicate should be a Type 3 trip, WT changed to 3, both coded as part of trip 5 for that case.”

Identifying information: Wave 6. CSID 19027 (PREID 742).

Problem. All dates missing. Interviewer note in BC column indicates part of same trip as PREID 741.

Action. Following Chico SRC: “corrected using fill-down from previous record.”

Identifying information: Wave 6. Following cases.

- 35475, 1197 (format is CSID, PREID)
- 32223, 1152
- 32223, 1153
- 32223, 1155
- 32223, 1156
- 31507, 1059
- 31507, 1060

Problem. Start day is missing.

Action. Following Chico SRC: “Chico: all corrected using fill-down from previous record.”

Identifying information: Wave 6. CSID 19867, RECID 777.

Problem. Q6 (return month) is listed as 12, outside the two-month acceptable range that we have defined, though not logically impossible. Interviewer writes “3 or 33 days.”

Action. Following Chico SRC: “Recommend coding as 3 days (return month should be 11).”

Identifying information. All waves.

Problem. There is a problem with the definition of the GOAL variable (which is the total number of trips for the two-month time period).² The GOAL variable reads zero

² The variable GOAL was not part of the original survey; it was created by Chico in the course of extracting the CATI data.

when in a number of cases there have been multiple trips to the beach. GOAL seems to be calculated as the sum of Q2A and Q3A, the number of trips in the first and second months of the survey time period respectively. The problem seems to be that when Q2A and Q3A are missing, GOAL is defined as zero, rather than as missing. I tried to note the cases where this happened for each diary survey in my line-by-line walk through the 6 datasets. I have this information jotted down (that is, other case id numbers where GOAL equals zero and there were some beach trips taken.)

Initial Chico response. “I would argue that GOAL is properly defined. If the respondent indicated that they went to no beaches, then they were not asked Q2 or Q3 and Goal=0 indicates that the respondent went to 0 beaches during that two-month period.”

Research Team reaction: There are a number of cases in each dataset where the GOAL variable shows a zero (0) value but the data otherwise indicate some beach trips have been taken (i.e. beaches visited or activities undertaken at the beach are listed for the respondent). Here are some specific instances where this is the case:

Wave 1: CSID 15096, CSID 16198, CSID 17327; Wave 2: CSID 11499, CSID 12047, CSID 14304, CSID 15949, CSID 17293, CSID 19761; Wave 3: CSID 15047. Wave 4: CSID 12888, CSID 14685; Wave 5: CSID 15251, CSID 15949, CSID 16314, CSID 30431; Wave 6: CSID 14844.

Next Chico SRC Response from Allen Lunde: “From my initial look I would say we are talking a computer anomaly. The data make it clear that these folks have gone to the beach. The data indicate that they went in either December or January, but not how many times. It may have to do with changing the number of times they remembered going. Either way, goal was originally a control variable and not a variable that I intended to be analyzed. We could find the anomalies and make them jive with the data if you would like to use the variable for analysis.”

Action. We told Chico to take no action at this time to fix the GOAL variable. Eventually, we asked them to create indexing variables (Chico SRC called them TRIP and BTRIP) to track the number of beach trips and beaches visited on each trip.

Identifying information. Waves 2, 3, 4, variables inwaterX and howdeepX³

Problem. Logical inconsistency on Q4d – Q4e skip pattern results.

Water-Based Recreation Results	Diary 2	Diary 3	Diary 4
inwaterX. Went in Water? # Saying Yes/Total Responses (% Yes)	192/914 (21.0%)	96/454 (21.1%)	179/689 (26.0%)
Q4E. howdeepX? # Saying “Got Head Wet”/Total Going in Water (% That Got Head Wet of Those Going in Water)	33/204* (16.2%)	40/153* (26.1%)	83/243* (34.2%)

*It is strange that for each of the diary surveys the number of responses to howdeepX exceeds the number of “yes” responses to inwaterX. The respondent is only asked howdeepX if they answer “yes” to inwaterX. There is a logical inconsistency here.

Action. First consider a quantification measure of the number of errors (extra responses) as a percentage of total observation in howdeepX. In Wave 2, 12 errors out of 204

³ These are variables Q4D and Q4E in the survey text and Chico coding.

observations (5.9%). In Wave 3, 57 errors out of 153 observations (36%). In Wave 4, 64 errors out of 243 observations (26%). Though these are somewhat high percentages, since we can trace back and eliminate the “no” answers from inwater X that shouldn’t have been asked howdeepX and solve such problems with a high degree of confidence. This—extra responses causing untidy skip pattern results—seemed to be common when we look at branching (e.g. skip pattern) results, as detailed in Research Team Memos. Eventually, Chico SRC stopped making corrections even when they agreed there was a problem. Some of these branching structure results have been flagged but not yet corrected. Other waves need to be examined for this. Quantification for dataset as a whole not done at this time.

Identifying information. Chico variables Q2 and Q2a (whether went to the beach and how many times in first month of the wave) for all waves. (Stands in for problems with Q3 and Q3a as well.)

Problem. logical inconsistency in skip pattern results —yes answers to q2 should equal number of responses in total to q2a.

Errors in Q2A variable across diaries

Diary	Q2, yes responses	Q2a response	# errors	% errors
1	189	264	75	28
2	174	273	99	36
3	134	210	76	36
4	187	287	100	35
5	261	327	66	20
6	135	203	68	33

Action. We are not using these variables for analysis and so chose not to pursue their correction. Further, as in the immediately preceding problem, the issue is always extra responses that could be eliminated. This is better than missing responses. Here was Chico SRC’s response on this: “As for problems with q2/q2a and q3/q3a my only theory is that problems occurred when the respondent changed their mind and required the interviewer to attempt to leave the normal skip pattern. The actual dates and lines of data should indicate the number of times they went in a given month. The program did not rely on the respondent to get the number of trips right on the first try, but allowed the respondent to change there mind if they realized that they went to the beach more or less than they originally remembered. We could add a variable or two based on the actual trip data that would be more accurate than q2a or q3a, which only indicate what the respondent thought when first asked.”

Identifying information. Waves 1-5.

Problem. Missing Case ID (CSID) and Record ID (RECID) numbers occur as follows:

Diary 1. Format Problem. Match to other diaries, e.g. Case ID for each row.

Diary 2. RECID missing =103 (6.8%),
CSID missing =14 (0.9%).

Diary 3. RECID missing =1(0.1%),
CSID missing =1 (0.1%)

Diary 4. RECID missing =80 (5.7%),
CSID missing =12 (0.9%).

Diary 5. RECID missing =47 (2.9%),
 CSID missing =47 (2.9%).

Diary 6. No problems.

Some of cases of missing CSID numbers do not appear very troubling, for example when the missing value is bracketed above and below by the same number, but others are more mysterious and potentially important. Also, please modify Diary 1 so that it is in the same format as the others? While other diary surveys list CSID in each row, Diary 1 only lists CSID once for each respondent.

Action. Chico SRC retrieved the missing values from the raw data.

Identifying information. Replenishment dataset.

Problem. There are some cases added via replenishment that appear in Diaries 5 and 6 are missing from the Replenishment Dataset!

Action. The SRC found these and provided them to the Research Team. Here is the Chico SRC explanation: “Here is what happened. Looks like there were about 44 cases that were recruited in a pretest directory. Since nothing was modified it was decided to use these cases, but by the time we output the data, I had forgotten about the recruits that came from the pretest. So I need to output all this data and have David restructure it before it comes to you.”

Identifying information. Replenishment dataset.

Problem. In January 2002, Chico SRC informed Research Team that it had found 99 cases (non-beach users) that belong in the Replenishment dataset.

Action. The SRC provided these to the Research Team.

Identifying information. CV bid data from Wave 5 module.

Problem. Initially missing.

Action. Chico SRC provided these upon request.

Identifying information. Wave 3 perceptions module data.

Problem. Inconsistencies as detailed in last memorandum to Chico (full text provided in next Appendix).

Action. Pending.

Identifying information. Wave 2 health module data.

Problem. Inconsistencies in skip pattern results as detailed in last memorandum to Chico (full text provided in next Appendix).

Action. Pending.

Identifying information. Wave 5 diary survey data.

Problem. Inconsistencies in skip pattern results as detailed in last memorandum to Chico (full text provided in next Appendix).

Action. Pending.

****Craig Mohn’s Data Checking Findings Begin Here****

Identifying Information: wave 1, recid 57
Problem: data incompatible with a trip
Action: flag as non-trip (set destid = missing)

Identifying Information: wave 1, recid 128
Problem: data supports classification as trip
Action: flag as trip. leave tripid = 1, set destid = 1

Identifying Information: wave 2, recids 44,58,128,217,228,251,270,340,379,478,
665,691,902,976,1042,1058,1071,1220,1404
Problem: spurious record inserted after last trip
Action: delete these records

Identifying Information: wave 4, preid 838
Problem: missing trip number
Action: replace tripid = 3, destid = 2

Identifying Information: wave 4, recid 1187
Problem: only data is month
Action: delete record

Identifying Information: wave 5, recid 14
Problem: empty record, labeled "ignore this row"
Action: delete record

Identifying Information: wave 2, recid 72
Problem: no details, month = 99 = "no more trips"
Action: delete record

Identifying Information: wave 1, recid 1652-1659
Problem: algorithm which calculates trip id missed changes in personid for records 1652
and 1653
Action: code 1652 as non-trip, renumber tripid for 1653-1659 starting at 1

Identifying Information: wave 1, recid 259
Problem: incorrect trip numbering
Action: tripid = 2

Identifying Information: wave 2, preid 661-663
Problem: incorrect trip numbering
Action: renumber tripids = 21-23 as per chico response

Identifying Information: wave 2, preid 1170-1172
Problem: incorrect numbering of trips/destinations
Action: 1170 becomes tripid 2, destid 2
1171 becomes tripid 3, destid 2

1172 becomes tripid 4, destid 2

Identifying Information: wave 4, preid 819-824

Problem: bad trip numbers

Action: drop preid 824, renumber tripid from 2-6 for preid 819-823

Identifying Information: wave 4, preid 1129-1131

Problem: bad trip numbers

Action: drop

Identifying Information: wave 4, preid 1227

Problem: bad trip numbers

Action: renumber tripid = 7 for preid 1227

Identifying Information: wave 5, preid 1161-1163

Problem: bad trip numbers

Action: renumber tripid = 18-20 for these preids

Identifying Information: wave 6, preid 587-589

Problem: renumber trips, chico says keep these (from recid 566)

Action: tripids become 1-3 and destids = 2

Identifying Information: wave 6, preid 940

Problem: trip number bad

Action: tripid = 7

Identifying Information: wave 6, preid 1176

Problem: apparent duplicate record 1143, insufficient details to recover

Action: drop preid 1176

Identifying Information: wave 1, preid 120 and 1588

Identifying Information: wave 3, preid 826

Identifying Information: wave 4, preid 571

Problem: apparent nontrips with beach info supplied

Action: reclassify as trips

Identifying Information: wave 1, recids 110, 437

Problem: destid sequence doesn't start at 1

Action: destid = 1

Identifying Information: wave1 recid 1381

Problem: new person, trip sequence starts wrong

Action: tripid = 1, destid = 1

Identifying Information: wave 1, recid 1495-1500

Problem: same destinations on multiday trip

Action: drop 1495-1498, tripid = 1, destid = 7,8 for 1499,1500 respectively

Identifying Information: wave 2, recid 1299

Problem: tripid is 22 instead of 2

Action: replace tripid = 2

Identifying Information: wave 1, recids 1382, 1383

Problem: didn't start trip numbering at 1

Action: tripid = 1,2 for 1382,1383 respectively

Identifying Information: wave 1 recid = 1585

Problem: first trip missing destid, appears as nontrip

Action: destid = 1

Identifying Information: wave 1, preid 828, 846

Problem: claim no trips and trips

Action: drop these records

Identifying Information: wave 1, preid 120

Problem: tripid/destid flags say non-trip, but has dates

Action: replace destid = 1

Identifying Information: wave 2, recid 468

Problem: tripid/destid imply nontrip, partial date, no beach

Action: set date and month to missing as per chico

Identifying Information: wave 2, recid 1143

Problem: tripid/destid imply nontrip, partial date, no beach

Action: set date and month to missing as per chico

Identifying Information: wave 3, recid 818

Problem: tripid/destid imply nontrip, partial date, no beach

Action: set date and month to missing as per chico

Identifying Information: wave 4, recid 249

Problem: tripid/destid imply nontrip, no beach

Action: set date and month to missing as per chico

Identifying Information: wave 5, preid 1053

Problem: tripid/destid imply nontrip, incomplete

Action: set date and month to missing as per chico

Identifying Information: wave 6, recid 1093

Problem: tripid/destid imply nontrip, partial date, no beach

Action: set date and month to missing as per chico

Identifying Information: wave 2, preid 363,1038
wave 3, preid 206,774
wave 4, preid 974
wave 5, preid 1619
wave 6, preid 128,172,175,248,356,599,657,735

Problem: destid set, person claimed to take no trips
Action: code as non-trip (destid = missing) as per chico

Identifying Information: wave 4, personid 14685, tripid 2
Problem: this appears to have the wrong personid associated with it
Action: drop trip and renumber higher trips for this person

Identifying Information: wave 4, personid 14908, trips 3,4
Problem: trips 3 and 4 overlap, single multideestination trip
Action: change tripid 4, destid 1 into tripid 3, destid 2 and decrement tripid for higher-numbered trips for this person in this wave

Identifying Information: wave 1, personid 13927
Problem: tripid 3 and 4 appear to be duplicates
Action: drop tripid 4 and decrement higher numbered trips

Identifying Information: wave 1, personid 15370
Problem: tripid 24 should begin on 24th rather than 20th
Action: change start date to 24th

Identifying Information: wave 1, personid 19239
Problem: tripids 6 and 11 begin on same day, no time given
Action: could be true, ignore

Identifying Information: wave 2, personid 10124
Problem: trips 8 and 9 to same beach, same date, same time
Action: drop tripid 9 and decrement higher trip numbers

Identifying Information: wave 2, personid 11698
Problem: trips 1 and 2 identical except duration
Action: drop tripid 2

Identifying Information: wave 2, personid 12000
Problem: trips 1 and 2, different times but trip 2 has zero time spent
Action: ignore, time spent may be unreliable and is frequently missing anyway

Identifying Information: wave 2, personid 12047
Problem: trips 3 and 4 exact duplicates
Action: drop tripid 4 and decrement higher tripids

Identifying Information: wave 2, personid 12774

Problem: trips 8,9,24,25,28,29 suspicious - 9 and 25 are duplicates but 29 is not
Action: drop tripid 9 and tripid 25, decrement higher tripids appropriately

Identifying Information: wave 2, personid 13896

Problem: trips 16 and 17 - same day, same time, different beaches

Action: make this a multideestination trip, tripid 17 goes to 16, destid goes to 2, decrement high tripids appropriately

Identifying Information: wave 2, personid 17293

Problem: tripids 4 and 5 are exact duplicates

Action: drop tripid 5 and decrement higher tripids appropriately

Identifying Information: wave 2, personid 19082

Problem: trips 1 to 4 are a single multideestination trip

Action: change trips to 1, and destid to 1,2,3,4

Identifying Information: wave 2, personid 19986

Problem: trip 2 identical to trip1 except for beach, beach unusual for panelist

Action: drop tripid 2, decrement higher tips appropriately

Identifying Information: wave 3, personid 11923

Problem: trips 8 and 9 exact duplicates

Action: drop tripid 9, decrement higher tripids appropriately

Identifying Information: wave 3, personid 14637

Problem: trips 7 and 8 exact duplicates

Action: drop tripid 8, no higher trips to decrement

Identifying Information: wave 3, personid 15512

Problem: trips 1 and 2 overlap, different beaches

Action: change to multideestination trip, tripid 2 becomes 1 destid becomes 2
no higher trips to decrement

Identifying Information: wave 4, personid 11923

Problem: trips 14 and 20 duplicates of 13 and 13

Action: drop tripids 14 and 20, appropriately decrement higher tripids

Identifying Information: wave 4, personid 14908

Problem: trips 3 and 4 overlap

Action: make trip 3 a multideestination trip (tripid 4 becomes 3 with destid 2)

Identifying Information: wave 4, personid 18925

Problem: trip 3 appears to have miskeyed start month

Action: change month to 6 (from 7)

Identifying Information: wave 4, personid 19566

Problem: trip 7 a duplicate of trip 6
Action: drop tripid 7, decrement others appropriately

Identifying Information: wave 5, personid 10805
Problem: trips 3 and 4 are one multidest trip, as are 5 and 6
Action: change destid to 2 for trips 4 and 6, decrement tripids appropriately
(tripid 4 becomes 3, 5 and 6 become 4, and 7 becomes 5)

Identifying Information: wave 5, personid 11456
Problem: trip 14 destination 2 is on a different day than destination 1
Action: change tripid 14 destid 2 to tripid 15 destid 1, increment higher tripids appropriately

Identifying Information: wave 5, personid 11456
Problem: trip 16 and 17 two adjacent trips on same day to same beach
Action: drop tripid 17, adjust beginning time and duration to reflect merger
5.5 hour trip beginning at 2:30 instead of 2.5 hour trip at 2:30
plus 3 hour trip at 5

Identifying Information: wave 5, personid 13183
Problem: trip 10 a duplicate of trip 9
Action: drop tripid 10, decrement higher tripids appropriately

Identifying Information: wave 5, personid 13943
Problem: trip 20 a duplicate of trip 19
Action: drop tripid 20, decrement higher tripids appropriately

Identifying Information: wave 5, personid 30125
Problem: trip 8 a duplicate of trip 7
Action: drop tripid 8, decrement higher tripids appropriately

Identifying Information: wave 5, personid 30638
Problem: trip 3 a duplicate of trip 2
Action: drop tripid 3, decrement higher tripids appropriately

Identifying Information: wave 5, personid 35765
Problem: trip 1 and 2 same date, same time. both multidestination, but trip 1 is 3 destination and trip 2 is only 2 destination. likely data entry error, but unrecoverable
Action: drop tripid 1, decrement higher tripids appropriately

Identifying Information: wave 6, personid 14685
Problem: trip 14 a duplicate of trip 13, probably a miskeyed date on 13, but unrecoverable
Action: drop tripid 14, decrement higher tripids appropriately

Identifying Information: wave 6, personid 18941

Problem: trips 8,12,9,10,20,21,29,30 all day trips to venice, dates for trips 10 and 12 and 21 are duplicates. long run of adjacent days on these trips so correction is pretty obvious for 2 of 3

Action: tripid 10 date = 13, tripid 12 date = missing tripid 21 = 31

Identifying Information: wave 1, personid 19082

Problem: a mess. a 3 day trip with much duplication and date confusion

Action: drop preids 1498, 1499, 1500, 1501 (duplicates)

preid 1502 becomes tripid 2 destid 1

preid 1503 becomes tripid 1 destid 7, start date = 29

Identifying Information: wave 4, preids 450,454,457,459,461,588,777

Problem: keys shifted one to right on multiday, date and month. all have multiday entered as 4 (3 means single day) and nonsense dates which either conflict or have obviously wrong month

Action: all become single day. preid 450 starts 4/8, preid 454 starts 4/21, preid 457 starts 4/30, preid 459 starts 5/1, preid 461 starts 5/13, preid 588 starts 4/9, preid 777 starts 4/30

Identifying Information: wave 4, personid 13280

Problem: trips 2,5,6,9 and 12 are entered as multideestination but each is on a different day

Action: split these trips into single destination trips, incrementing higher tripids appropriately

Identifying Information: wave 4, personid 13565

Problem: trip 1 is entered as multideestination but each is on a different day

Action: split this trips into single destination trips, incrementing higher tripids appropriately

Identifying Information: wave 4, personid 10225

Problem: trip 3 is entered as multideestination but each is on a different day

Action: split this trips into single destination trips, incrementing higher tripids appropriately

Identifying Information: wave 4, personid 11796

Problem: trip 1 is entered as multideestination but each is on a different day

Action: split this trips into single destination trips, no other trip

Identifying Information: wave 4, personid 14884

Problem: trip 28 has multiple dates - July 30,31,32,33, and 34

this does NOT suggest that other multidest trips have similar date issue, since trip 27 has two identical dates

Action: change dates to July 30 for all destids

Identifying Information: wave 4, personid 15802

Problem: trip 3 entered as multkdestination but each on different date

Action: split trip into 3 separate trips, increment other tripids

Identifying Information: wave 1, personid 10316, tripid 6

wave 1, personid 15949, tripid 3

wave 1, personid 19239, tripid 5

wave 1, personid 19986, tripid 1,3

Problem: multiday not set, end date not same as start date

Action: set multiday

Identifying Information: wave 1, personid 13952, tripid 3

Problem: multiday missing, end date distinct from start date, another trip takes place in middle of this trip, this trip reconstructed from paper records.

Action: drop end date

Identifying Information: wave 4, personid 12350, tripid 9

Problem: single day, end date distinct from start date

Action: drop end date

Identifying Information: wave 6, personid 31187, tripid 1

Problem: single day, end date distinct from start date

Action: drop end date

Identifying Information: wave 6, personid 31187, tripid 2

Problem: bad dates

Action: set begin date 10/27 end date 10.29 as per chico

Identifying Information: wave 1, personid 15659

Problem: trips 1-26 could be either december or january

Action: not fixable, leave missing

Identifying Information: wave 1, personid 12309

Problem: trips 1-4 could be either december or january

Action: not fixable, leave missing

Identifying Information: wave 2, preid 21

Problem: one trip date out of sequence

Action: change start date from 9 to 19

Identifying Information: wave 1, personid 14752, tripid 3

Problem: missing start month

Action: set to january based on other dates, as per chico

Identifying Information: wave 1, personid 15060, tripid 1

Problem: missing start month
Action: set to december, as per chico

Identifying Information: all trips in replenishment dataset
Problem: which records (possibly incomplete) refer to trips
Action: trips have at either a destination or a start date specified, or are part of a set of multiple records per person

Identifying Information: Replenishment dataset and personid 35988
Problem: record for first trip contains no information other than recid. probably a trip, but maybe just a spurious entry
Action: drop this record. decrement all higher tripids for this respondent.

Identifying Information: 1075 records in replenishment dataset
Problem: contain a variable called revdate which agrees with start date where start date is specified, and often contains a date when start date not specified
Action: use revdate when start date not useable

Identifying Information: Replenishment dataset, personid 31564
Problem: tripid 1 has 7 destids = 1,1,1,2,2,3,3 - dates are not the same
Action: use dates to renumber:
 preid 524 as tripid 16, destid 1
 preid 527 as tripid 16, destid 2
 preid 528 as tripid 17, destid 1
 preid 525 as tripid 18, destid 1
 preid 529 as tripid 19, destid 1
 preid 526 as tripid 20, destid 1
 preid 530 as tripid 21, destid 1
Problem: trip 10 has unusual beach (21) all others are 52, 122, or 123. would like to change beach to 122 or 121
Action: ignore

Identifying Information: Replenishment dataset, personid 31619
Problem: two records with tripid 1, destid 1
Action: renumber preid 551 as tripid 8, based on dates

Identifying Information: Replenishment dataset, personid 32229
Problem: three records with tripid 1, destid 1
Action: renumber based on dates
 preid 849 as tripid 2, destid 1
 preid 852 as tripid 4, destid 1
 preid 855 as tripid 5, destid 2
 preid 854 as tripid 5, destid 1
 preid 856 as tripid 6, destid 1
 preid 857 as tripid 7, destid 1

preid 858 as tripid 8, destid 1
preid 851 as tripid 9, destid 1
preid 859 as tripid 10, destid 1
preid 860 as tripid 11, destid 1

Identifying Information: Replenishment dataset, personid 35256

Problem: five records with tripid 1, destid 1

Action: renumber based on dates

preid 983 as tripid 10, destid 1
preid 984 as tripid 11, destid 1
preid 985 as tripid 12, destid 1
preid 986 as tripid 13, destid 1

Identifying Information: Replenishment dataset, personid 30795

Problem: tripid 1 has 2 destids with differing dates

Action: second destination is really trip 10, set tripid 10, preid 1
for preid 78

Identifying Information: Replenishment dataset, personid 31015

Problem: tripid 1 has 2 destids with differing dates

Action: second destination is same date as tripid 12,
so set tripid 10, preid 2

Identifying Information: Replenishment dataset, personid 31271

Problem: bad date on second destid for tripid 1, recid is same as destid 1

Action: set date to 6/13

Identifying Information: Replenishment dataset, personid 31293

Problem: bad month on second destid for tripid 2, recid is same as destid 1

Action: set date to 6/17

Identifying Information: Replenishment dataset, personid 30899

Problem: two multiday trips overlap.

Action: set duration of first trip (preid 160) to one day. set single day.

Errors in and changes to screener data.

Most of these errors are for people where either 20000 or 10000 was added to the age and gender was missing. I speculate that the first digit (1 or 2) is related to the gender of the respondent. this error only occurred for the first person in the household (the respondent), and for single respondent cases the age given in diaries corresponded to the age here with the leading 100 or 200 omitted. corrections are based on these assumptions. gender corrections could not be verified because diary gender responses fluctuate since they are recorded based on surveyors' perceptions.

I did a whole bunch of wave 4-specific corrections for gender - I looked at gender and age across waves. this effort was not comprehensive, will be superceded by a more general attempt to correct ages and genders across waves.

We dispense with the previous format. CSID (= personid) uniquely identifies the record of interest. Problem found is in the next row and is followed by the action taken.

csid 11978
gender not specified, age=20888 for respondent
set age = 888

csid 12469
gender not specified, age=20067 for respondent
set age = 67

csid 12562
gender not specified, age=10023 for respondent
set age = 23

csid 12699
gender not specified, age=20056 for respondent
set age = 56

csid 14315
gender not specified, age=20040 for respondent
set age = 40

csid 15049
gender not specified, age=20035 for respondent
set age = 35

csid 16027
gender male, age=4 for respondent
set age = 43 based on diary data

csid 19251
gender not specified, age=20031 for respondent
set age = 31

csid 31186
gender not specified, age=20037 for respondent
set age = 37

csid 31594
gender not specified, age=20070 for respondent
set age = 70

csid 31757
gender not specified, age=20072 for respondent
set age = 72

csid 35452
gender not specified, age=10019 for respondent
set age = 19

csid 35503
gender not specified, age=20076 for respondent
set age = 76

csid 35850
gender not specified, age=20075 for respondent
set age = 75

csid 35958
gender not specified, age=20072 for respondent
set age = 72

Zip code errors and corrections.

zip code errors were complicated to track down. in general, a person's zip code was propagated through waves until changed by a response in the diaries. pc-miler was used to track down invalid zip codes, and a trial-and-error approach was used to correct some erroneous zip codes. I examined the addresses when a zip code indicated a possible move to determine whether it was a new move or a correction. in the below, wave 0 is recruitment phase

csid 15244 new zipcode waves 0,1,2 90262

csid 10641 zipcode bad in recruitment, participates never, leave bad

csid 15111, 12533 zipcode good, but not in California, correct these later

csid 11805 replace 92677 with 92766

csid 13216 replace wave 1,2 zips 90220 (bad correction in diary 1)

csid 30688 replace wave 4 zip with correct one from wave 5

csid 30974 replace wave 4 zip with correct one from wave 5

NOTE The checking and cleaning of zip codes went involved a number of iterations. The notes given above cover just Craig Mohn's work. Chris Busch, Ted Davis (a

graduate students at UC Davis who was also part of the research team) and Allen Lunde also worked to clean the zip codes. We do not have details of Lunde – Chico SRC work, and so have not conducted a quantification of the errors here, since this would be only a partial measure.

Appendix 3. Memoranda and Communications Produced During Data Checking.

List of Contents in Appendix 3.

1. Chico Response to December 31, 2001 Memo from Research Team.
 2. Chico Response to September 13, 2001 Memo from Research Team.
 3. Research Team Memo to Chico SRC, September 13, 2001.
 4. Research Team Memo to Chico SRC, September 8, 2001.
 5. Research Team Memo to Funding Partners, September 2001.
 6. Research Team Memo to the File on Corrections, July 2001.
 7. Memorandum from Chico SRC in response to June 22 memo.
 8. Memorandum from Research Team to Chico SRC, June 22, 2001.
 9. Email Traffic Week of June 18, 2001.
 10. Memorandum to the File on Scrutinizing Outliers, June 8, 2001.
 11. Memorandum from Research Team to Funding Partners, May 29, 2001.
-
-

Note. As the identifying information at the start of this memo indicates, the base material for this memo was originally developed as a memo with questions for Chico SRC. Chico SRC's responses are embedded in the text of the memo.

MEMORANDUM

TO: Allen, Chico SRC
FROM: Research Team
DATE: 31 December 2001
RE: Questions on Beach Data and Documentation

This memo includes questions and issues that fall into six categories.

1. Replenishment Survey Dataset
2. Productions Report
3. Perceptions Module Data (Wave 3) – Appendix illustrates some of these questions.
4. Health Module Data (Wave 2)
5. Wave 5 Data
6. San Onofre, Wave 6 Module Data

1. Replenishment Survey: *Completed*

- Q1AA—taken trips since Memorial Day-- should provide a distinction between types of records, e.g. people for whom there should be data on beach trips vs. not. However, there are 59 people who answer this question "yes" but have only a single

record and no beach selection indicated. All told there are 146 observations where the variable Q1AA indicates trips were taken, yet there is no value for the beach selected. Many of these are cases where several apparently valid trips are specified, and the last trip has no destination beach. We need to know what is going on with these apparent trips without destinations.

Chico made corrections to the dataset and provide a new version to Research Team.

- Codebook shows same definition for Q2M and Q1AA, but the reading script shows Q2M asks about beach trips taken ON memorial day. Please confirm that the script is right, and that the codebook is wrong.
- Over the summer we asked you about the Q1A2 variable, that is why there was almost no data in the variable. Your response was, “What we wanted to do here is get diary data for those who went to the beach. The program asks no information of those who were non-beach users. I notice that we put variables into the codebook that were not actually asked. I actually think that in the beginning they were asked but we switched, for time sake, after a few days. So you should be looking at "cat" to determine if someone was a beach user or not. The codebook hadn't listed the Spanish codes for these, I've updated the codebook and taken out the variables that we don't have data for. I've included the questionnaire, but taken out the part that was skipped by the program.”

A couple of questions about this. When you say some people were asked questions and others not, that is only for variables that are no longer in the dataset (and codebook), right? Was everyone asked the same questions on their beach trips, or was their a hierarchy of questions based on the type and number of trips, as in the diary surveys?

Lastly, you say that no information was asked of non-beach users, but that's not what we want and not what's in the dataset. There are non-beach users in the survey, and as in the recruitment, we wanted demographic data on these people. That's what was done, yes?

- RE: CAT variable. Definitions carries over from recruitment--beach users means has gone to beach in last 12 months. Correct?
- Note Codebook problem:
Q1AA codebook lists values of 1,3,5,d,r
Dataset contains values 1,3,5,88888

2. Production Report: *Completed*

These questions and comments on the Production Report are a collection of those from Project Funders and the Research Team.

There are continuing small numerical disagreements between numbers in the report and between numbers given in the report and datasets. We focus on these since we will be unable to resolve such questions after Chico's contract has been fulfilled. (We can continue to work on the text thereafter.) It is important that when a number is broken down, the sum of the parts equals the total. Also, there should be exact correspondence between numbers reported in tables and those given in the text.

- RE: Production Statistics, Initial Recruitment p.8.

Report reads, "A total of 1850 interviews were completed." But 2 of these were no good, right? So change to 1848. (That way the production report matches the recruitment dataset.)

Report reads, "A total of 1035 respondents had gone to the beach during the 12 months preceding the interview of which 887 agreed to participate in the panel." Our work on the recruitment dataset shows 1034 beach users and 814 non-users. (We agree with the 887 figure.)

Also make corresponding changes in Table 1.

"Of the 1980 working residential numbers that did not result in an interview, a total of 965 potential respondents refused to complete the survey and an additional 20 respondents completed only part of an interview.

A total of 1556 cases remained unresolved at the completion of the recruitment process..."

What's the difference between working residential numbers not leading to a completed interview and unresolved cases? Is the 1556 a subset of the 1980? What happened with the 1980-965-20 cases not explained? Please clarify.

- RE: Production Statistics, Replenishment

Problem with replenishment table (table #2). For number fielded, parts (3429) do not equal sum (3414).

We're confused on the question of eligible households and number of completed interviews...

pg. 9, "Of the 760 respondents who completed the replenishment survey..." If this is so, why are there 662 cases in the replenishment dataset. Of these 662, 458 were eligible (e.g. beach users) and 204 were not eligible. Why does Table 2 show 665 eligible households. Please reconcile production report with replenishment dataset, which shows (662 total cases/ 458 eligible/ 204 not eligible).

“Of the remaining 466 respondents, 341 agreed to participate on the panel and receive the mailing, 26 agreed to participate via the web, and 100 refused to participate in the panel.” 466 not equal to 467. The 367 number implied by the text (341+21) matches the replenishment dataset, which is good, but the dataset shows 91 cases declining to join the panel (not 100).

“At the completion of the replenishment survey, there were 2191 unresolved cases. Of those, 1206 were instances where the telephone number was unavailable, 259 where the respondent was unavailable for contact, and 173 were scheduled callbacks, which were never reached, 180 were answering machines, and 371 were never answered. The response rate for the replenishment survey was 43%.” 2191 not equal to 2189.

- pg. 11, second diary, “There were 188 total unresolved cases at the completion of this diary. In 21 of those cases, respondents were unavailable for contact, and in 97 cases the telephone numbers attempted were unavailable. A total of 22 cases had been answering machines, 2 cases were never answered, and 44 cases were scheduled callbacks, which were never reached.” 188 not equal to 186.
- End note 6, added in response to Roger’s request for definition of “clean up” says: “The clean-up refers to the fact that a small number of cases were never released into the diary sample because they were defiantly lost or refused to be called back.” Presumably defiantly is supposed to be definitely. Still, what does “definitely lost” mean. The numbers were misplaced? People were no longer reachable at that number?
- Number added to Roger’s Table 3 (diary outcomes). Diary 3, 753 total does not equal 751 in parts. Diary 5 percentages add to 100.5, too large to be due to rounding (one of other columns adds to 99.8 – we can add a footnote on numbers not adding to 100 exactly due to rounding). Note research team will add “total” row to Table 3, as suggested by Funders.
- Chico has added text on option of participating by web, but haven’t yet added example of web page that we requested. Allen writes, *I’ve added the material but am still trying to get an example for the appendix. The page is no longer online. I am attempting to contact the web master to get the source code.* Yes, please add if at all possible.
- pg. 1, in discussion of software used, please list specific versions.
- pg. 3, “This arrangement meant that there was a total of ten different modes in which a given beach trip could be recorded and an almost infinite number of combinations that could occur over a two month period.” What are the ten different modes? Aren’t there just six? The five kinds of beach trips plus typical day data.

- A bit more on financial incentives (requested by Roger and the funders): how many people were offered financial incentives of each type, and how many accepted these offers?
- We need to reconcile numbers in each wave given in production statistics section of the production report with the datasets that were delivered to the research team.

	Wave 1	2	3	4	5	6
Production Report	656	638	506	547	721	720
Datasets We Received	660	638	509	529	731	720

You can see that, in 2 waves, the two numbers are in agreement, but otherwise not (I have bolded the waves where the numbers disagree). Please can you resolve these anomalies and get these numbers to agree. To find the number of respondents in datasets received, we used unique CSID numbers. Would you suggest something else?

Note: In the process of revising the production report it was discovered that a small number of cases were left out of some of the data files. These cases need to be sent to the research team. They are:

Diary 1: 12257, 15004, 15551, 16447

Diary 3: 10119, 18752, 19508

Diary 4: 19633

Diary 5: 10579, 11136, 12039, 12611, 12709, 13654, 13671, 13862, 15512, 19858, 30187, 35381

3. Perceptions Module Data (Wave 3)

- The 6 randomly selected beaches that the respondent is to be asked about in detail (re: perceptions on sand, water, parking) are listed in the variables: nmor – nmo6. We combined these variables, and this aggregate variable should give for each beach the number of responses we should have on perceptions (e.g. rating of sand, water, parking). There is a problem of a lack of good correspondence between the beaches listed in the nmor-nmo6 variables and the number of responses on perceptions for beaches.

Many of the differences between the two numbers (number of times listed in nmor-nmo6 and number of responses on perceptions) are not very large. But we will need an explanation to give to people to explain differences that can not be resolved. Where the number of times listed in nmor-nmo6 is a little larger than the number of responses on perceptions, a candidate explanation would be that the person dropped out in the course of the module. But just as often, there are the number of times listed in nmor-nmo6 is a little larger than the number of responses on perceptions. How can this be?

The program contained a switch for each of the beaches picked in nmor-nmo6. The switch would be kicked on when the caller entered a beach number in nmor-nmo6. If a caller entered a number incorrectly, than changed the number, the switch would remain turned on. Therefore when the program approached the questions about the extra beach, it would read the switch as saying that those questions should be asked. Therefore, there will be cases where only 6 beaches are listed in nmor-nmo6, but data for additional beaches appears. A solution for this problem could be a recode that would eliminate data for beaches not listed in nmor-nmo6. An example of this problem is case 11993. This case refers to San Onofre South.

Of special concern is where there is a large difference between the two numbers. The most egregious example is San Clemente State, where there are 75 observations in nmor-nmo6, but there are only 49 observations for that beach on perceptions (e.g. in nb1c, and etc.).

Looking at the data for nb1c, the situation may actuality be worse than it first appears. There seems to be a disconnect with those that said they went to San Clemente and those who answered the follow-up I have no explanation for this. I cannot find a problem in the program and the coding sheet used by the callers is correct. There may be a programming issue that I cannot find. David Philhour is looking into the matter further

Of special concern is where there is a large difference between the two numbers. The most egregious example is San Clemente State, where there are 75 observations in nmor-nmo6, but there are only 49 observations for that beach on perceptions (e.g. in nb1c, and etc.). Some other examples are Huntington State (135 not equal to 141), Long beach (133 not equal to 143), and Seal Beach (118 not equal to 125).

The big problems occur where the observations are fewer than expected. San Clemente is one example. I believe the only other example is Salt Creek. I have not actually looked at the data for Salt Creek. However, I believe the data for these cases is unreliable.

In those cases where there are more observations than expected, I believe this happened because the caller "turned the switch on". The more observations for a given beach, the more difference could be expected. These extra observations could be cleaned with a recode.

See the Appendix for the full story on this. It gives tabulated results for nmor-nmo6 and for perceptions questions for each beach.

- We're having trouble understanding the flow of respondents through questions on visitation and familiarity. Recall we first ask people if they visited a beach, and if they hadn't visited, we ask they are familiar with the beach. There is a problem of either too many or not enough respondents in the follow up on familiarity. (Usually there are too many responses). Consider the sequence on San Onofre South. The question "Theo" asks, Ever been to San Onofre State Beach North?

354 say no, these should go to q15a on familiarity, but there are 357 responses to q15a, e.g. 3 too many. Similar story on next question on San O. South—400 “no” responses on having visited, but 407 total responses on the question of familiarity. Please help us understand why this is so.

There are two reasons this happens.

First, the respondent changes their mind after the caller has coded the follow-up question and says that they actually had gone to beach x. The caller jumps back and changes the answer to the "ever been" question, but the data remains in the follow-up. This can be cleaned with a re-code that says if the "ever-been" question is yes <1>; the follow-up question is missing data.

Second, the respondent said "Don't Know" to the "ever been" question. This would have been coded <d>. The person who output the original data did not tell SPSS to accept string variables, so the <d> was ignored. These can be output, and the more common "don't know" code of <88888> can be recoded into the proper variables.

- A follow up on the questions on having visited. This is the first question about a beach that each respondent is asked about. The full sample that is answering the module should respond to each of these. How should we respond if we are asked why the sample size varies from question to question? Most of the time, 484 responses are recorded on the visitation questions, but almost half the time there are a smaller number of responses. (A quick survey shows 481, 475, 482, and as small as 469 for the sample size for the first few of these questions.)

This is another area where the missing values should indicate that the respondent said "don't know". This can be fixed with a recode.

- We are puzzled by the questions n101 – n108 and strt – st4d. The plan was to ask about only 6 beaches from the set of 51 for which we asked visited/familiarity questions. These questions n101 – n108 and strt – st4d allow respondents to list beaches from outside of the 51 and to rate 4 more beaches on water/sand/parking quality. So most people answer questions on 6 beaches, but if they say that they are “familiar with conditions at any other beaches” then they answer questions on up to 10 beaches. Why was this done? Was there a miscommunication between the Research Team and Chico?

The only reason I would have put the programming in is that I was asked to. Below is the text directly from the "perceptions module" I received.

Are you familiar with any of the other beaches in San Diego, Ventura or Santa Barbara Counties?"

If Yes, Q8

Which ones?

Then ask Q3-5 for that beach. (This assumes that the pretest indicates that people will mention only up to 4 additional beaches. Else we will run up against time constraints.)

- Why are there only 28 observations on Q15 (asks respondent if he/she wants to complete the module now)? Shouldn't there be 484 responses on this? And 19 people say "no, not now," in response to Q15, which sends the respondent to Q15w. So why are there 21 responses to Q15w (should equal number of no responses to Q15 =19)?

Only persons who said they had more than 14 beach trips were asked q15 initially. Considering this was a control variable, I believe the pattern changed at different points in the study and should not be used in analysis.

- N23C. In dataset but not in codebook. What is this variable? Assume we should ignore it/delete it? Here are the observations included for the variable in the dataset.

N23C	Freq.	Percent	Cum.
2	1	33.33	33.33
88888	2	66.67	100.00
Total	3	100.00	

How would you rate the water at Poche Beach.

The following are not questions. They are observations for the record on fixes that will have to be made.

- Note codebook error. "nmor" defined as, "caller chooses 6 beaches to ask more about." This should say interviewer chooses 6 beaches at random from visited or familiar list.
- Note codebook error. Codebook codes for ratings at Newport: nb10c, nb10f, nb10g. Variables in dataset are n10c, n10f, and n10g.

Will you make these changes be made by you or us?

4. Health Module Data (Wave 2)

There seem to be some problems with respondents following the skip pattern as would be expected based on responses. Many of these are off by only 1. (Should we conclude that one

interviewer asked follow up questions no matter what, or how should we interpret these?) Some other skips have larger discrepancies, and we list these first.

- RE: Q15LC. There should be 130 answers to this (based on “yes” answer to Q15L). Why are there 142?

This is a skip pattern issue. Those cases where q15l is no should be cleared with a recode.

- RE: q15f. It has 25 “yes” responses that should be send to the follow up questions, e.g. 15fa and 15fb. Why are there 27 responses to these follow up questions? Especially, why are there are 22 responses to 15fd?

There are 2 cases where q15f is no, but the follow-up questions are answered. Based on notes it is clear that for case 10784, the answer should be no. However the woman had a skin problem that had been happening for some years, so the respondent originally and answered yes and answered the follow-ups before the caller found out it was not related. For case 16390 I suspect the caller went down the wrong path and tried to correct it by coding everything as no.

- q15c has 54 “yes” responses that should be send to the follow up questions, e.g. 15cb and onward. Why are there 30 or 31 responses to these follow up questions?

There are 3 inappropriate responses to 15cb where the answer to q15c is no. In all three cases (10784, 13654, 15522) q15c was changed to no after it was originally coded as yes. The discrepancy in data can be fixed with a recode.

- RE: q15d. It has 27 “yes” responses that should be send to the follow up questions, e.g. 15da and onward. Why are there 28 responses to these follow up questions?

Again 10784 had changed his/her response to q15d from yes to no. Apparently because the condition was pre-existing. Evidently this person had multiple health problems not related to Beach use, and that confused this portion of the interview.

- Only females who answer 15ha should answer 15hd. But these two have the same number of responses. Were all respondents in 15ha female, or was this a mistake?

At the time question 15hd is asked, the computer does not know if the respondent is male or female. The instruction in all caps is an instruction to the interviewer. If the respondent is a male we don't actually have to ask him if he is pregnant. We simply enter "no" and move on.

- Reading script problem – it says “>15hd<ASK ONLY IF RESPONDENT IS FEMALE. IF THE RESPONDENT IS MALE SKIP TO Q15j(i). IF YOU

ALREADY KNOW THE ANSWER TO THE FOLLOWING QUESTION, JUST CODE THE ANSWER WITHOUT ASKING THE QUESTION AND GO ON TO Q15(i).” There is no Q15j(i). What should this say? Also, the instructions say to code the answer to the next question if the answer is known. How would the interviewer know if the person is pregnant?

Checking on the skip pattern, it is correct in the program. I believe the skip pattern indicated is a remnant from an earlier version. The computer simply takes the caller to the next question. A woman may have stated earlier in the call that she was pregnant, for example if asked if she went into the water on a beach trip. The instruction is simply saying that, if the woman already indicated she was pregnant, we don't have to ask at that point, simply code and move on. The instruction should probably be removed from the codebook to avoid confusion.

- >q15i< Have you had diarrhea during the past 7 days?
<1> Yes [goto 15ia]

There are 21 answers “yes” to this, but 22 responses to the follow up. (Other follow ups okay.)

- >q15j< During the past seven days have you had stomach pain or cramps?
<1> Yes [goto 15ja]

There are 69 answers “yes” to this, but 70 responses to the follow up questions.

The extra case is once again 10784.

- Question regarding reading script. It says:

>15ba< **Where** you very hot, did you have a fast pulse that caused you discomfort?
<1> Yes
<3> No

<88888> Don't know
<99999> Refused

Do you agree that we should replace “where” with “when”?

Actually it should be "were".

- Note codebook error for the record. Q15 – remove, “<5> man,” as response option. This is not a gender specific question.

You can do this or ask us to do this.

5. Wave 5 Data

- Based on response to Q1AA, there should be 332 responses to Q2, but there are 335 responses. Why might this be? Reading script shows that Q2a should have the same number of responses (327) as “yes” answers to Q2 (261). Why are there 327 as opposed to 261.

The computer automatically stores a <0> response in q2a for those cases where the respondent said no to q2. Three cases (12966, 14632, 30001) had their answer changed from yes to no on q1aa. The follow-up data should be removed with a recode.

- Based on response to Q1AA, there should be 332 responses to Q3, but there are 333 responses. Why might this be? Reading script shows that Q3a should have the same number of responses (327) as “yes” answers to Q3 (209). Why are there 327 as opposed to 209.

The computer automatically stores a <0> response in q3a for those cases where the respondent said no to q3. Case 30001 had erroneous data in q3a because of the changed data in q1aa. The data should be removed with a recode.

- For question q15b, there should be 209 responses (based on answers to q15), but we have 198 responses. These 11 (209-198) responses are also missing in the subsequent question q15c. Why is this?

There are too many responses in q15. Case 12382 has 8 responses. It should only have 1. Case 18935 has 3 responses, it should only have 1

- For question Q15, there are more responses than for other questions in the module (731>720). Why did this happen?

Same as above.

- Problem with reading script for diary surveys. Skip pattern for Q2 and Q3 is incomplete. Need to add [go to q3] and [go to calC] to these questions for responses other than yes, as shown in bold below. Currently shows all answers going to questions immediately following, which is incorrect.

The reading scripts really weren't designed to be used in the analysis. Therefore I don't think there is any problem adding the changes below if it makes the reading script easier to read. I will put the actual section from the program below.

>q2< Let's begin with the calendar for February. Did you go to a beach in any of these counties, at least once during February?

<1> Yes [goto q2a]

<3> No

<d> Don't know

<r> Refused

====> **[go to q3]**

>q2a< [allow 2] During February, how many times did you go to the beach?

<1-29>

====> [goto q3]

>q3< How about March. Did you go to a beach in any of these counties at least once during March?

<1> Yes [goto q3a]

<3> No

<d> Don't know

<r> Refused

====> **[goto calc]**

>q3a< [allow 2] During March , how many times did you go to the beach?

<1-31>

====>[goto calc]

Also, presumably “====>[goto calc]” means figure out which sequence of questions on beach trips the respondent should get and send them there? Please confirm this

The Actual program reads:

>q1aa< [#loc 300/1]

Did you go to a beach in Southern California, That is San Diego, Orange, Los Angeles, Ventura or Santa Barbara counties, at least once in August or September

<1> Yes

<3> Not Sure

<5> No [goto q15]

<d> Don't know

<r> Refused

====> [goto q2]

>q1aa.2<

Usted fue a alguna playa en el sur de California, Esto es en los condados de San Diego, Orange, Los Angeles, Ventura o Santa Barbara, por lo menos una vez en Agosto o Septiembre

<1> Si
<3> No esta Seguro
<5> No

<d> No Sabe
<r> No Responde

====>

>fxq2< [allow 2] What is the correct number of days they
went to the beach in August

<1-29>

====>

>fxq2.2< Cuantos dias fueron ellos a la playa en Agosto

<1-29>

====>

>q2fx< [store <> in q2a][store fxq2 in q2a]
[store <> in godc][store <> in gojn][store <0> in goal]
[goto calc]

>q2< Let's begin with the calendar for August. Did you go to a beach
in any of these counties, at least once during August?

<1> Yes [goto q2a]
<3> No

<d> Don't know
<r> Refused

====>

>q2.2<Empecemos con el calendario de Agosto. Usted fue a la playa en
alguno de estos condados, por lo menos una vez durante Agosto?

<1> Si
<3> No

<d> No Sabe
<r> No Responde

====>

>tof1< [if q2 eq <3>][store <0> in q2a]
[else]
[if q2 eq <d>][store <0> in q2a]
[else]
[if q2 eq <r>][store <0> in q2a]

```

    [else]
    [goto q2a]
    [endif][endif][endif]
    [goto q3]

>q2a< [allow 2] During August, how many times did you
go to the beach?

    <1-29>

    ==> [goto q3]

>q2a.2< Durante Agosto, cuantas veces fue usted a la playa ?

    <1-29>

    ==>

>fxq3< [allow 2]What is the correct number of days they went
to the beach in September

    <1-31>

    ==>

>fxq3.2< [allow 2]Cuantos dias fueron ellos a la playa en Septiembre

    <1-31>

    ==>

>q3fx< [store <> in q3a][store fxq3 in q3a]
    [store <> in godc][store <> in gojn][store <0> in goal]
    [goto calc]

>q3< How about September. Did you go to a beach in any of these
counties at least once during September?

    <1> Yes [goto q3a]
    <3> No

    <d> Don't know
    <r> Refused

    ==>

>q3.2< Y en Septiembre. Usted fue a la playa en alguno de estos
condados, por lo menos una vez en Septiembre?

```

```

    <1> Si
    <3> No

    <d> No Sabe
    <r> No Responde

    ===>

>tofx< [if q3 eq <3>][store <0> in q3a]
      [else]
        [if q3 eq <d>][store <0> in q3a]
      [else]
        [if q3 eq <r>][store <0> in q3a]
      [else]
        [goto q3a]
      [endif][endif][endif]
      [goto calc]

>q3a< [allow 2] During September, how many times did you go to the
beach?

    <1-31>

    ===>[goto calc]

>q3a.2< [allow 2]Durante Septiembre, cuantas veces fue usted
a la playa?

    <1-31>

    ===>

>calc<
  [store q2a in godc]
  [store q3a in gojn]
  [store q2a in goal]
  [add q3a to goal]
  [if goal eq <0>][goto q15]
  [else]
  [goto ninc]
  [endif]

>ninc< [allow 3][store <1> in ninc]

>gtdn< [allow 1][store <0> in gtdn]

```

- The dataset includes a variable v38 that is not in the codebook. What is this variable, or can we ignore it?

This is a control variable. I cannot imagine why it is in the data. Please ignore it.

6. San Onofre Questions (Wave 6 Module)

For the record:

The number of “no” answers to Q17: *Gone to San Onofre Beach in past 2 yrs?* is 598. An equal number of responses should have been found for Q17b: *Why didn't go to San Onofre Beach in past 2 yrs?* Instead, there are a total of 599 responses for Q17b. This discrepancy is due to an extraneous response to Q17b by respondent csid = 15522. Since respondent csid = 15522 answers “yes” to Q17, there should not be a response to Q17b.

Understood

Appendix – Supports Questions on Perceptions Module

Appendix

The following are tabulated results for the nmor-nmor6 series of variables.

This gives all beaches selected at random from visited-familiar list for each respondent.

```
. tab all_nmor
```

all_nmor	Freq.	Percent	Cum.
1	24	0.89	0.89
2	75	2.77	3.66
3	36	1.33	4.99
4	86	3.18	8.17
5	62	2.29	10.46
6	31	1.15	11.60
7	43	1.59	13.19
8	125	4.62	17.81
9	58	2.14	19.96
10	90	3.33	23.28
11	151	5.58	28.86
12	135	4.99	33.85
13	125	4.62	38.47
14	69	2.55	41.02
15	57	2.11	43.13
16	31	1.15	44.27
17	118	4.36	48.63
18	18	0.67	49.30
19	68	2.51	51.81
20	133	4.92	56.73
21	69	2.55	59.28
22	23	0.85	60.13
23	14	0.52	60.64
24	20	0.74	61.38
25	2	0.07	61.46
26	49	1.81	63.27
27	132	4.88	68.14
28	92	3.40	71.54
29	88	3.25	74.80
30	52	1.92	76.72
31	26	0.96	77.68
32	99	3.66	81.34
33	95	3.51	84.85
34	119	4.40	89.25
35	30	1.11	90.35
36	11	0.41	90.76
37	56	2.07	92.83
38	43	1.59	94.42
39	3	0.11	94.53
40	8	0.30	94.83
41	28	1.03	95.86
42	5	0.18	96.05

43		2	0.07	96.12
44		4	0.15	96.27
45		2	0.07	96.34
46		14	0.52	96.86
47		12	0.44	97.30
48		10	0.37	97.67
50		21	0.78	98.45
51		42	1.55	100.00

Total		2706	100.00	

Now, I got through and tabulate for each beach responses to the question on rating of water quality at each beach. The number of responses for each beach should equal the number of times the beach appears in the above list. (A check shows the same number of responses for water, sand, and parking ratings. So I list only water quality.) The order of beaches for below results follows the codebook.

The name of the beach is given in CAPS, followed in parenthesis by number of times listed under nmor variables (should equal total number of tabulated responses). Where not equal to expected number, I highlight in **bold**. This shows problems with 33/51 of the beaches. Many of these are only off by one or two, but others are off by 5, 10, or more.

. tab q15b SAN ONOFRE N (43)

Q15B		Freq.	Percent	Cum.
1		4	9.30	9.30
2		19	44.19	53.49
3		8	18.60	72.09
4		1	2.33	74.42
5		1	2.33	76.74
88888		9	20.93	97.67
99999		1	2.33	100.00

Total		43	100.00	

. tab n37c SAN ONOFRE SOUTH (24)

N37C		Freq.	Percent	Cum.
1		2	8.00	8.00
2		10	40.00	48.00
3		9	36.00	84.00
4		1	4.00	88.00
88888		3	12.00	100.00

Total		25	100.00	

. tab nb1c SAN CLEMENTE STATE (75)**

NB1C		Freq.	Percent	Cum.
1		6	12.24	12.24

2		26	53.06	65.31
3		11	22.45	87.76
88888		6	12.24	100.00

Total		49	100.00	

. tab nb2c SAN CLEMENTE CITY (36)

NB2C		Freq.	Percent	Cum.

1		3	8.11	8.11
2		24	64.86	72.97
3		4	10.81	83.78
5		1	2.70	86.49
88888		5	13.51	100.00

Total		37	100.00	

. tab n23c POCHE (2)

N23C		Freq.	Percent	Cum.

2		1	33.33	33.33
88888		2	66.67	100.00

Total		3	100.00	

. tab nb3c CAPISTRANO (86)

NB3C		Freq.	Percent	Cum.

1		9	10.34	10.34
2		51	58.62	68.97
3		11	12.64	81.61
4		1	1.15	82.76
5		1	1.15	83.91
88888		14	16.09	100.00

Total		87	100.00	

. tab nb4c DOHENY (62)

NB4C		Freq.	Percent	Cum.

1		5	8.06	8.06
2		32	51.61	59.68
3		12	19.35	79.03
4		3	4.84	83.87
88888		10	16.13	100.00

Total		62	100.00	

. tab nb5c SALT CREEK (31)**

NB5C		Freq.	Percent	Cum.

2		8	80.00	80.00

3		1	10.00	90.00
4		1	10.00	100.00

Total		10	100.00	

. tab nb6c ALISO CREEK (43)

NB6C		Freq.	Percent	Cum.

1		5	11.63	11.63
2		17	39.53	51.16
3		12	27.91	79.07
4		1	2.33	81.40
5		1	2.33	83.72
88888		7	16.28	100.00

Total		43	100.00	

. tab nb7c LAGUNA (125)

NB7C		Freq.	Percent	Cum.

1		11	8.46	8.46
2		72	55.38	63.85
3		29	22.31	86.15
4		6	4.62	90.77
5		1	0.77	91.54
88888		11	8.46	100.00

Total		130	100.00	

. tab nb8c CRYSTAL COVE (58)

NB8C		Freq.	Percent	Cum.

1		11	18.33	18.33
2		31	51.67	70.00
3		8	13.33	83.33
4		1	1.67	85.00
88888		9	15.00	100.00

Total		60	100.00	

. tab nb9c CORONA DEL MAR (90)

NB9C		Freq.	Percent	Cum.

1		8	8.60	8.60
2		52	55.91	64.52
3		22	23.66	88.17
4		2	2.15	90.32
5		1	1.08	91.40
88888		8	8.60	100.00

Total		93	100.00	

. tab nl0c NEWPORT (151)

NL0C	Freq.	Percent	Cum.
1	15	9.26	9.26
2	73	45.06	54.32
3	51	31.48	85.80
4	15	9.26	95.06
5	1	0.62	95.68
88888	7	4.32	100.00
Total	162	100.00	

. tab n11c HUNT STATE (135)

N11C	Freq.	Percent	Cum.
1	5	3.55	3.55
2	55	39.01	42.55
3	46	32.62	75.18
4	21	14.89	90.07
5	5	3.55	93.62
88888	9	6.38	100.00
Total	141	100.00	

. tab n12c HUNT CITY (125)

N12C	Freq.	Percent	Cum.
1	4	3.08	3.08
2	34	26.15	29.23
3	44	33.85	63.08
4	29	22.31	85.38
5	7	5.38	90.77
88888	12	9.23	100.00
Total	130	100.00	

. tab n13c BOLSA CHICA (69)

N13C	Freq.	Percent	Cum.
2	28	40.00	40.00
3	27	38.57	78.57
4	6	8.57	87.14
88888	8	11.43	98.57
99999	1	1.43	100.00
Total	70	100.00	

. tab n14c SUNSET (57)

N14C	Freq.	Percent	Cum.
1	2	3.51	3.51
2	20	35.09	38.60

3		28	49.12	87.72
4		1	1.75	89.47
88888		6	10.53	100.00

Total		57	100.00	

. tab n15c SURFSIDE (31)

N15C		Freq.	Percent	Cum.

1		2	6.25	6.25
2		10	31.25	37.50
3		9	28.12	65.62
5		2	6.25	71.88
88888		9	28.12	100.00

Total		32	100.00	

. tab n16c SEAL (118)

N16C		Freq.	Percent	Cum.

1		5	4.00	4.00
2		52	41.60	45.60
3		44	35.20	80.80
4		12	9.60	90.40
5		1	0.80	91.20
88888		11	8.80	100.00

Total		125	100.00	

. tab n17c ALAMITOS (18)

N17C		Freq.	Percent	Cum.

2		6	31.58	31.58
3		5	26.32	57.89
4		4	21.05	78.95
88888		4	21.05	100.00

Total		19	100.00	

. tab n18c BELMONT SHORES (68)

N18C		Freq.	Percent	Cum.

1		1	1.54	1.54
2		29	44.62	46.15
3		17	26.15	72.31
4		6	9.23	81.54
88888		12	18.46	100.00

Total		65	100.00	

. tab n19c LONG BEACH (133)

N19C		Freq.	Percent	Cum.
------	--	-------	---------	------

1	3	2.10	2.10
2	33	23.08	25.17
3	40	27.97	53.15
4	41	28.67	81.82
5	12	8.39	90.21
88888	13	9.09	99.30
99999	1	0.70	100.00
Total	143	100.00	

. tab n20c CABRILLO (69)

N20C	Freq.	Percent	Cum.
1	3	4.23	4.23
2	30	42.25	46.48
3	14	19.72	66.20
4	13	18.31	84.51
5	2	2.82	87.32
88888	9	12.68	100.00
Total	71	100.00	

. tab n21c POINT FERMIN (23)

N21C	Freq.	Percent	Cum.
1	2	8.70	8.70
2	9	39.13	47.83
3	6	26.09	73.91
4	2	8.70	82.61
88888	4	17.39	100.00
Total	23	100.00	

. tab n22c ROYAL PALMS (14)

N22C	Freq.	Percent	Cum.
1	2	14.29	14.29
2	7	50.00	64.29
3	3	21.43	85.71
4	1	7.14	92.86
88888	1	7.14	100.00
Total	14	100.00	

. tab n24c ABALONE COVE (20)

N24C	Freq.	Percent	Cum.
1	2	10.00	10.00
2	11	55.00	65.00
3	6	30.00	95.00
88888	1	5.00	100.00

Total	20	100.00
-------	----	--------

. tab m24c MALAGA COVE (21)

M24C	Freq.	Percent	Cum.
1	1	4.76	4.76
2	9	42.86	47.62
3	5	23.81	71.43
4	1	4.76	76.19
88888	5	23.81	100.00
Total	21	100.00	

. tab n25c TORRANCE (49)

N25C	Freq.	Percent	Cum.
1	1	1.96	1.96
2	22	43.14	45.10
3	12	23.53	68.63
4	5	9.80	78.43
88888	10	19.61	98.04
99999	1	1.96	100.00
Total	51	100.00	

. tab n26c REDONDO (132)

N26C	Freq.	Percent	Cum.
1	4	2.82	2.82
2	55	38.73	41.55
3	42	29.58	71.13
4	23	16.20	87.32
5	4	2.82	90.14
88888	14	9.86	100.00
Total	142	100.00	

. tab n27c HERMOSA (92)

N27C	Freq.	Percent	Cum.
1	4	4.35	4.35
2	41	44.57	48.91
3	26	28.26	77.17
4	7	7.61	84.78
88888	14	15.22	100.00
Total	92	100.00	

. tab n28c MANHATTAN (88)

N28C	Freq.	Percent	Cum.
1	5	5.56	5.56
2	38	42.22	47.78

3		30	33.33	81.11
4		10	11.11	92.22
88888		7	7.78	100.00

Total		90	100.00	

. tab n29c EL SEGUNDO (50)

N29C		Freq.	Percent	Cum.

2		15	28.85	28.85
3		14	26.92	55.77
4		6	11.54	67.31
5		2	3.85	71.15
88888		15	28.85	100.00

Total		52	100.00	

. tab n30c DOCKWEILER (26)

N30C		Freq.	Percent	Cum.

2		9	34.62	34.62
3		7	26.92	61.54
4		9	34.62	96.15
88888		1	3.85	100.00

Total		26	100.00	

. tab n31c MARINA DEL REY (99)

N31C		Freq.	Percent	Cum.

1		4	3.92	3.92
2		51	50.00	53.92
3		24	23.53	77.45
4		11	10.78	88.24
5		2	1.96	90.20
88888		10	9.80	100.00

Total		102	100.00	

. tab n32c VENICE (95)

N32C		Freq.	Percent	Cum.

1		2	2.00	2.00
2		24	24.00	26.00
3		36	36.00	62.00
4		22	22.00	84.00
5		1	1.00	85.00
88888		14	14.00	99.00
99999		1	1.00	100.00

Total		100	100.00	

. tab n33c SANTA MONICA (119)

N33C		Freq.	Percent	Cum.
------	--	-------	---------	------

1	6	4.44	4.44
2	37	27.41	31.85
3	43	31.85	63.70
4	34	25.19	88.89
5	3	2.22	91.11
88888	12	8.89	100.00
Total	135	100.00	

. tab n51c WILL ROGERS (42)

N51C	Freq.	Percent	Cum.
1	2	4.76	4.76
2	14	33.33	38.10
3	15	35.71	73.81
4	5	11.90	85.71
88888	6	14.29	100.00
Total	42	100.00	

. tab n34c TOPANGA (30)

N34C	Freq.	Percent	Cum.
1	5	17.24	17.24
2	10	34.48	51.72
3	10	34.48	86.21
88888	4	13.79	100.00
Total	29	100.00	

. tab n35c LAS TUNAS (11)

N35C	Freq.	Percent	Cum.
2	6	54.55	54.55
3	2	18.18	72.73
88888	3	27.27	100.00
Total	11	100.00	

. tab n36c MALIBU LAGOON (56)

N36C	Freq.	Percent	Cum.
1	4	6.90	6.90
2	30	51.72	58.62
3	12	20.69	79.31
4	6	10.34	89.66
88888	6	10.34	100.00
Total	58	100.00	

. tab n38c DAN BLOCKER (3)

N38C	Freq.	Percent	Cum.
2	2	66.67	66.67
3	1	33.33	100.00
Total	3	100.00	

. tab n39c POINT DUME (8)

N39C	Freq.	Percent	Cum.
1	2	25.00	25.00
2	5	62.50	87.50
88888	1	12.50	100.00
Total	8	100.00	

. tab n40c FREE ZUMA (ZUMA CO) (10)

N40C	Freq.	Percent	Cum.
1	1	9.09	9.09
2	6	54.55	63.64
3	2	18.18	81.82
4	1	9.09	90.91
88888	1	9.09	100.00
Total	11	100.00	

. tab n47c ZUMA (28)

N47C	Freq.	Percent	Cum.
1	8	27.59	27.59
2	17	58.62	86.21
3	2	6.90	93.10
4	1	3.45	96.55
88888	1	3.45	100.00
Total	29	100.00	

. tab n41c EL MATADOR (4)

N41C	Freq.	Percent	Cum.
1	1	33.33	33.33
2	1	33.33	66.67
4	1	33.33	100.00
Total	3	100.00	

. tab n42c LA PIEDRA (2)

N42C	Freq.	Percent	Cum.
2	2	66.67	66.67

4	1	33.33	100.00

Total	3	100.00	

. tab n43c EL PESCADOR (5)

N43C	Freq.	Percent	Cum.
2	3	75.00	75.00
88888	1	25.00	100.00

Total	4	100.00	

. tab n44c NICHOLAS CANYON (2)

N44C	Freq.	Percent	Cum.
1	1	50.00	50.00
2	1	50.00	100.00

Total	2	100.00	

. tab n45c LEO CARRILLO (14)

N45C	Freq.	Percent	Cum.
1	2	11.76	11.76
2	8	47.06	58.82
3	2	11.76	70.59
4	1	5.88	76.47
88888	3	17.65	94.12
99999	1	5.88	100.00

Total	17	100.00	

. tab n46c COUNTY LINE (12)

N46C	Freq.	Percent	Cum.
1	3	25.00	25.00
2	5	41.67	66.67
3	1	8.33	75.00
88888	2	16.67	91.67
99999	1	8.33	100.00

Total	12	100.00	

Chico Response to September 13 Memo

Chico Responses to Data questions (includes corrections after *****).

1) A check that every record was either a trip (TRIP and BTRIP both positive integers) or else a response that no trip (TRIP = 1, BTRIP missing) was taken and the following errors and presumed corrective actions emerged:

Wave 1

- RECID 57 refers to a valid CSID and is presumed a zero response; propose setting TRIP to 1. *****AGREE
- RECID 128 is a trip; propose setting TRIP and BTRIP to 1 *****AGREE

Wave 2

- wave 2 has a lot of observations which have no content except a 99 in the month field after a set of trips for a given CSID (data entry error?); propose deleting RECID 44, 58, 128, 217, 228, 251, 270, 340, 379, 478, 665, 691, 902, 976, 1042, 1058, 1071, 1220, 1404 *****AGREE

Wave 4

- wave 4 PREID 838 has a missing TRIP, propose setting TRIP to 4
****In that it shares RECID 775 with the previous case and is listed as a type 3 trip, I would set TRIP = 3 and BTRIP=2
- wave 4 RECID 1187 contains only MONTH=7, propose deleting it *****AGREE

Another Wave 1 question

- wave 1 RECID 1643-1663 must have errors in TRIP, BTRIP wrong (see diagram below);
propose setting RECID 1652 BTRIP to missing, and
propose setting RECID 1653 - 59 TRIP from 1 to 7 *****AGREE

(columns are PREID RECID CSID TRIP BTRIP GOAL)

1640	1637.00	19943	18	1	
1641	1638.00	19943	19	1	
1642	1639.00	19943	20	1	
1643	1640.00	19943	21	1	
1644	1641.00	19943	22	1	
1645	1642.00	19953	1		0
1648	1643.00	19964	1	1	6
1647	1644.00	19964	2	1	
1648	1645.00	19964	3	1	
1649	1646.00	19964	4	1	
1650	1647.00	19964	5	1	
1651	1648.00	19964	6	1	
1652	1649.00	19964	7	1	
1653	1650.00	19964	8	1	
1654	1651.00	19964	9	1	
1655	1652.00	19964	10	1	0
1656	1653.00	19966	11	1	9
1657	1654.00	19966	12	1	
1658	1655.00	19966	13	1	
1659	1656.00	19966	14	1	
1660	1657.00	19966	15	1	
1661	1658.00	19966	16	1	
1662	1659.00	19966	17	1	
1663	1660.00	19967	1		0
1664	1661.00	19998	1	1	5
1665	1662.00	19998	2	1	
1666	1663.00	19998	3	1	
1667	1664.00	19998	4	1	
1668	1665.00	19998	5	1	

2) A check that TRIP sequence has to be increasing from 1 to N revealed the following errors and presumed corrective actions:

Wave 1

- RECID 259; propose changing TRIP to 2 *****AGREE
- RECID 1382-83; propose changing TRIP 1-2 *****AGREE
- RECID 1585; propose changing BTRIP to 1 ****AGREE

Wave 2

- PREID 661, 662, 663; propose deletion for these because they have a duplicate RECID to prior, and unusual beaches for this CSID
 ***These beaches came from the "any that we missed??" section at the end of the questionnaire . The partial data came from the "usual trip" section. THEY were part of the original Record #645 and therefore have the same RECID. I would keep them.

- PREID 1170-72; propose dropping these – they are the same day as a prior trip, are santamonica and venice a likely single-trip pair?

***These cases came in from the "any we forgot?" section also. Instead of dropping them why not include them as BTRIP = 2 for the appropriate dates under the Venice Trip. Change triptype (WT) to reflect multiple beach trip Like such:

1170	1111	16209	2	2
1171	1111	16209	3	2
1172	1111	16209	3	2
1173	1112	16209	2	1
1174	1113	16209	3	1
1175	1114	16209	4	1
1176	1115	16209	5	1
1177	1116	16209	6	1

- RECID 1299; propose to change TRIP to 2 from 22 ***** AGREE

Wave 4

- PREID 819-24; propose changing TRIP to 2-7 (from duplicate RECID 757) *****AGREE ...but should drop PREID 824 because it lacks any real data
- PREID 1129-31; propose dropping since appear to replications of RECID 1054 *****AGREE
- PREID 1227; propose changing TRIP to 7 , duplicate RECID problem *****AGREE

Wave 5

wave 5 RECID 1161-63; propose changing TRIP to 18-20 because of duplicate RECID 1130 *** you refer to PREID 1161-63, These are some of the "any we left out" group. I AGREE with your solution.

Wave 6

- PREID 587-89; propose dropping these since they are duplicates of RECID 566 *****These are from the "any other??" group, they appear to be second beaches on the same days that they went to santa monica. I propose marking those with the appropriate trip # and changing BTRIP to 2 and changing the trip type (WT =3) see below

PREID, RECID,CSID, TRIP,BTRIP

586	566.00	15987	1	1
587	566.00	15987	1	2
588	566.00	15987	2	2
589	566.00	15987	3	2
590	567.00	15987	2	1
591	568.00	15987	3	1
592	569.00	15987	4	1
593	570.00	15987	5	1
594	571.00	15987	6	1
595	572.00	15987	7	1

- PREID 940; propose to change TRIP to 7, duplicate RECID 915 ***AGREE
- PREID 1176; propose dropping this, duplicate RECID 1143, no date, no details

*****AGREE

3) *A check that the destination sequence has to run from 1 to M revealed the following problems and possible resolutions:*

Wave 1

- RECID 110; propose setting BTRIP to 1 *****AGREE
- RECID 437; propose setting BTRIP to 1 *****AGREE
- RECID 1831; propose setting TRIP to 1 and BTRIP to 1
- *****NO RECID1831 exists in data (perhaps a typo?)
- RECID 1495-1500 are scrambled. proposal: delete RECID 1495-98, change 1499-1500 to TRIP 1 BTRIP 7-8. ***** WOW What a Mess!! After a bit of figuring I AGREE with your change.

4) *A check that no CSID have both BTRIP missing (indicating no trips that wave) and present (indicating trips) revealed the following problem and possible resolution:*

- **Wave 1:** PREID 828 and 846 are partial duplicates of other observations; propose deleting them ***** AGREE

5) *A look at non-trip (BTRIP missing) records which have a start date revealed:
(Note that several of these have no obvious resolution)*

Wave 1

- PREID 120; propose setting BTRIP to 1 *****AGREE

Wave 2

- RECID 468 is confusing - has partial date but no beach
***** Note: Q1AA=5 (no trips) Recommend dropping all trip data after Q1AA for this case and just keeping the demographic & module part (q141 and beyond)
- RECID 1143 is confusing - has date but no beach
- *****same as previous case

Wave 3

- RECID 818 is confusing - has partial date but no beach
*****same as previous case

Wave 4

- RECID 249 is confusing - has date but no beach
*****This is strange! They indicate no trips (Q1AA=5), but should have data for the module (but that is missing). I would drop everything after Col G but keep them in the database since that CSID shows up in other waves.

You may need it for cross-wave analysis.

Wave 5

- RECID(preid) 1053 is incomplete, propose calling it a no trip **AGREE
Leave TRIP = 1, BTRIP = (blank) delete Data between Q1AA and Q15 (save the demographics and module data at the end)

*******NOTE***** In the Wave 5 merged set Cols AK,AL,AM were included to assure an accurate merge. They should be deleted before analysis.**

Wave 6

- wave 6 RECID 1093 has only a partial date *** drop trip data since they indicated no trips (Q1AA=5) but keep everything from Q16 and beyond

6) In summary, these changes leave:

8463 records (beach destinations)

1034 participants

2179 person-wave reports involving zero trips

1582 person-wave reports involving 5917 trips

5560 of these trips appear to refer to a valid beach and have a valid date

4383 of these trips also have generally good demographic info

1766 of the zero-trip reports have generally good demographic info

End of Memorandum

Chico Response to August 31 Memo

Chico Responses to Data questions (includes corrections after *****).

1) A check that every record was either a trip (TRIP and BTRIP both positive integers) or else a response that no trip (TRIP = 1, BTRIP missing) was taken and the following errors and presumed corrective actions emerged:

Wave 1

- RECID 57 refers to a valid CSID and is presumed a zero response; propose setting TRIP to 1. *****AGREE
- RECID 128 is a trip; propose setting TRIP and BTRIP to 1 *****AGREE

Wave 2

- wave 2 has a lot of observations which have no content except a 99 in the month field after a set of trips for a given CSID (data entry error?); propose deleting RECIDs 44, 58, 128, 217, 228, 251, 270, 340, 379, 478, 665, 691, 902, 976, 1042, 1058, 1071, 1220, 1404 *****AGREE

Wave 4

- wave 4 PREID 838 has a missing TRIP, propose setting TRIP to 4
****In that it shares RECID 775 with the previous case and is listed as a type 3 trip, I would set TRIP = 3 and BTRIP=2
- wave 4 RECID 1187 contains only MONTH=7, propose deleting it *****AGREE

(continued next page)

Another Wave 1 question

- wave 1 RECID 1643-1663 must have errors in TRIP, BTRIP wrong (see diagram below);
propose setting RECID 1652 BTRIP to missing, and
propose setting RECID 1653 - 59 TRIP from 1 to 7 *****AGREE

(columns are PREID RECID CSID TRIP BTRIP GOAL)

1640	1637.00	19943	18	1	
1641	1638.00	19943	19	1	
1642	1639.00	19943	20	1	
1643	1640.00	19943	21	1	
1644	1641.00	19943	22	1	
1645	1642.00	19953	1		0
1648	1643.00	19964	1	1	6
1647	1644.00	19964	2	1	
1648	1645.00	19964	3	1	
1649	1646.00	19964	4	1	
1650	1647.00	19964	5	1	
1651	1648.00	19964	6	1	
1652	1649.00	19964	7	1	
1653	1650.00	19964	8	1	
1654	1651.00	19964	9	1	
1655	1652.00	19964	10	1	0
1656	1653.00	19966	11	1	9
1657	1654.00	19966	12	1	
1658	1655.00	19966	13	1	
1659	1656.00	19966	14	1	
1660	1657.00	19966	15	1	
1661	1658.00	19966	16	1	
1662	1659.00	19966	17	1	
1663	1660.00	19967	1		0
1664	1661.00	19998	1	1	5
1665	1662.00	19998	2	1	
1666	1663.00	19998	3	1	
1667	1664.00	19998	4	1	
1668	1665.00	19998	5	1	

2) A check that TRIP sequence has to be increasing from 1 to N revealed the following errors and presumed corrective actions:

Wave 1

- RECID 259; propose changing TRIP to 2 *****AGREE
- RECID 1382-83; propose changing TRIP 1-2 *****AGREE
- RECID 1585; propose changing BTRIP to 1 ****AGREE

Wave 2

- PREID 661, 662, 663; propose deletion for these because they have a duplicate RECID to prior, and unusual beaches for this CSID
 ***These beaches came from the "any that we missed??" section at the end of the questionnaire . The partial data came from the "usual trip" section. THEY were part of the original Record #645 and therefore have the same RECID. I would keep them.

- PREID 1170-72; propose dropping these – they are the same day as a prior trip, are santamonica and venice a likely single-trip pair?

***These cases came in from the "any we forgot?" section also. Instead of dropping them why not include them as BTRIP = 2 for the appropriate dates under the Venice Trip. Change triptype (WT) to reflect multiple beach trip Like such:

1170	1111	16209	2	2
1171	1111	16209	3	2
1172	1111	16209	3	2
1173	1112	16209	2	1
1174	1113	16209	3	1
1175	1114	16209	4	1
1176	1115	16209	5	1
1177	1116	16209	6	1

- RECID 1299; propose to change TRIP to 2 from 22 ***** AGREE

Wave 4

- PREID 819-24; propose changing TRIP to 2-7 (from duplicate RECID 757)
***AGREE ...but should drop PREID 824 because it lacks any real data
- PREID 1129-31; propose dropping since appear to replications of RECID 1054
****AGREE
- PREID 1227; propose changing TRIP to 7 , duplicate RECID problem
*****AGREE

Wave 5

wave 5 RECID 1161-63; propose changing TRIP to 18-20 because of duplicate RECID 1130 *** you refer to PREID 1161-63, These are some of the "any we left out" group. I AGREE with your solution.

Wave 6

- PREID 587-89; propose dropping these since they are duplicates of RECID 566
*****These are from the "any other??" group, they appear to be second beaches on the same days that they went to santa monica. I propose marking those with the appropriate trip # and changing BTRIP to 2 and changing the trip type (WT =3) see below

PREID, RECID,CSID, TRIP,BTRIP

586	566.00	15987	1	1
587	566.00	15987	1	2
588	566.00	15987	2	2
589	566.00	15987	3	2
590	567.00	15987	2	1
591	568.00	15987	3	1
592	569.00	15987	4	1
593	570.00	15987	5	1
594	571.00	15987	6	1
595	572.00	15987	7	1

- PREID 940; propose to change TRIP to 7, duplicate RECID 915 ****AGREE
- PREID 1176; propose dropping this, duplicate RECID 1143, no date, no details *****AGREE

3) *A check that the destination sequence has to run from 1 to M revealed the following problems and possible resolutions:*

Wave 1

- RECID 110; propose setting BTRIP to 1 *****AGREE
- RECID 437; propose setting BTRIP to 1 *****AGREE
- RECID 1831; propose setting TRIP to 1 and BTRIP to 1
- *****NO RECID1831 exists in data (perhaps a typo?)
- RECID 1495-1500 are scrambled. proposal: delete RECIDs 1495-98, change 1499-1500 to TRIP 1 BTRIP 7-8. **** WOW What a Mess!! After a bit of figuring I AGREE with your change.

4) *A check that no CSID have both BTRIP missing (indicating no trips that wave) and present (indicating trips) revealed the following problem and possible resolution:*

- **Wave 1:** PREID 828 and 846 are partial duplicates of other observations; propose deleting them ***** AGREE

5) *A look at non-trip (BTRIP missing) records which have a start date revealed: (Note that several of these have no obvious resolution)*

Wave 1

- PREID 120; propose setting BTRIP to 1 *****AGREE

Wave 2

- RECID 468 is confusing - has partial date but no beach
***** Note: Q1AA=5 (no trips) Recommend dropping all trip data after Q1AA for this case and just keeping the demographic & module part (q141 and beyond)
- RECID 1143 is confusing - has date but no beach
- *****same as previous case

Wave 3

- RECID 818 is confusing - has partial date but no beach
*****same as previous case

Wave 4

- RECID 249 is confusing - has date but no beach

****This is strange! They indicate no trips (Q1AA=5), but should have data for the module(but that is missing). I would drop everything after Col G but keep them in the database since that CSID shows up in other waves. You may need it for cross-wave analysis.

Wave 5

- RECID(preid) 1053 is incomplete, propose calling it a no trip **AGREE
Leave TRIP = 1, BTRIP = (blank) delete Data between Q1AA and Q15 (save the demographics and module data at the end)

*****NOTE***** **In the Wave 5 merged set Cols AK,AL,AM were included to assure an accurate merge. They should be deleted before analysis.**

Wave 6

- wave 6 RECID 1093 has only a partial date *** drop trip data since they indicated no trips (Q1AA=5) but keep everything from Q16 and beyond

6) *In summary, these changes leave:*

8463 records (beach destinations)

1034 participants

2179 person-wave reports involving zero trips

1582 person-wave reports involving 5917 trips

5560 of these trips appear to refer to a valid beach and have a valid date

4383 of these trips also have generally good demographic info

1766 of the zero-trip reports have generally good demographic info

End of Memorandum

MEMORANDUM

TO: Allen Lunde and David Philhour, Chico State SRC

FROM: Research Team

DATE: 13 September 2001

RE: Just About the Last Data Checking Questions

We have noted the following inconsistencies in the data. In some cases the remedy is obvious, and we ask Chico to state their agreement or alternative solution. In other cases the remedy is less clear and we would appreciate Chico's suggestions.

Agree with proposed remedy?

Wave 1 RECID 1381 – this is a new CSID, set TRIP = 1, BTRIP = 1

Wave 4 CSID 16198 TRIPS 1 & 2 overlap (6/18-7/18 and 7/8-7/9) – clearly need to set Q6BR = 6 for TRIP 1, which is single-day

Wave 4 CSID 18925 TRIPS 3 & 4 overlap (6/12-7/12 and 6/15) – clearly need to set Q6BR = 6 for TRIP 3

Wave 5 RECID 14 – empty, labeled “ignore this row” – delete this observation

Wave 5 CSID 12606 TRIPS 3 & 4 overlap – clearly need to set Q6BR = 8 for TRIP 3

Suggestions?

There is no clear resolution to the following. Unless Chico has evidence of data-entry error, we suggest leaving these as is and noting that there are some data inconsistencies in the dataset, since there are undoubtedly other minor errors in coding which will be discovered as time goes by.

Wave 4 CSID 10728 TRIP 4 (7/31-8/3) overlaps Wave 5 CSID 10728 TRIP 1 (8/1-8/3)

Wave 5 CSID 12960 TRIP 1 (9/27-10/2) overlaps Wave 6 CSID 12960 TRIP 1 (10/1-10/2)

Wave 6 CSID 14231 TRIPS 1 & 2 overlap (10/14-10/16 and 10/15)

Wave 1 CSID 19875 TRIP 1 (1/31-2/5) overlaps Wave 2 CSID 19875 TRIP 1 (2/1-2/4)

MEMORANDUM

TO: Allen Lunde and David Philhour, Chico State SRC

FROM: Research Team

DATE: 8 September 2001

RE: More Data Checking Questions

Here is a list of further discrepancies we have found in the data. We suspect that some of the discrepancies may be difficult to resolve, however resolution will result in a larger dataset. Note that for some types of analysis that the Research Team will conduct, an

invalid observation requires that a larger number of related observations be discarded. We would like to avoid this if possible, so we would greatly appreciate your help in resolving the issue.

Discrepancies

- 1) **A check that the Q4A multiday variable was set for every valid trip with a distinct start date and end date yielded the following discrepancies.**

Chico should verify that the following observations are truly multiday trips although Q4A is not set to 1:

- wave 1 CSID 10316 TRIP 6 - end date next day
- wave 1 CSID 15949 TRIP 3 - dates 12/14 - 12/15
- wave 1 CSID 19239 TRIP 5 - dates 12/15 - 12/17
- wave 1 CSID 19986 TRIP 1
- wave 1 CSID 19986 TRIP 3
- wave 4 CSID 12350 TRIP 9 - only trip with end date, all others indicated as nonmultiday

The following trips have problems:

- wave 6 CSID 31187 TRIP 1+2 – trip 2 overlaps 3,4,5, and 6 (see below)
- no obvious solution, drop this person from the wave

961.00	31187	1	1	5	1	1	2	1	3	10	1	10	1
961.00	31187	1	2							10	1		
962.00	31187	2	1							10	17	10	29
963.00	31187	3	1							11	19		
964.00	31187	4	1							11	26		
965.00	31187	5	1							11	27		
966.00	31187	6	1							11	27		
967.00	31187	7	1										
968.00	31204	1	1	2	1	1	1	1	1	10	7		
969.00	31204	2	1							11	11		
970.00	31216	1		0	5								

wave 1 CSID 13952 TRIP 3 - dates 12/17 to 1/3 - another trip is in middle -
"this observation was reconstructed from paper records" in comment - possibly
this is a 4th trip on 1/3 or some data entry error
resolution - drop the end date or discard this CSID in this wave

2) A check that observations with a valid-appearing beginning date (Q4) have a valid month (Q4MT) revealed the following:

Note that there are no obvious resolutions to the ambiguity here, so these observations will be lost for any analysis requiring dates. There are a lot (34) of trips here.

wave 1 CSID 14752 TRIP 3 could be either dec or january so keep it coded as bad
wave 1 CSID 15060 TRIP 1 could be either dec or january so keep it coded as bad
wave 1 CSID 15659 TRIP 1-26 all could be either dec or jan
wave 2 CSID 12309 TRIP 1-4 could be either feb or march

3) Miscellaneous errors

wave 1 preid 988
wave 1 preid 996
wave 2 preid 347
wave 2 preid 21 - change date (Q4) from 9 to 19, sequence looks more reasonable???
wave 6, CSID 19867, TRIP 2, (Recid 778) Q4= 99 not allowed, no WT.
Delete this record?
wave 6, CSID 30037 (Recid 807), WT missing, change it to 1?

4) Trips which start on the same day.

Many of these trips appeared to be possible duplicate records. However, it is equally likely that the date was entered incorrectly and the "duplicate" observation(s) refer to distinct trips.

Duplicates should be dropped, otherwise the dates should be corrected.

Note that it is possible for multiple trips to occur on the same day, so we examined starting time and duration where available.

wave 1 CSID 13927 TRIP 3 & 4 appear to be duplicates to same beach
wave 1 CSID 15370 TRIP 24 should probably be on the 24th rather than the 20th, look at
data
wave 1 CSID 19239 TRIP 6 & 11 - both on same day- no time given...possibly true...
wave 2 CSID 10124 TRIPS 8 & 9 - same day same beach, same time
wave 2 CSID 11698 TRIPS 1 & 2 - everything same except duration
wave 2 CSID 12000 TRIPS 1 & 2 - different beaches, different times, but trip 2 has zero
time spent
wave 2 CSID 12047 TRIPS 3 & 4 - exact duplicate entries

wave 2 CSID 12774 TRIPS 8 & 9 & 24 & 25 & 28 & 29 - 9 and 25 are dups, 28 & 29 appear distinct

wave 2 CSID 13896 TRIPS 16 & 17 - appear distinct, yet same time, same day, different beaches - no clue what to do

wave 2 CSID 14304 TRIPS 4 & 5 - different beaches, different times

wave 2 CSID 14884 TRIPS 11 & 12 - different beaches, different times

wave 2 CSID 17293 TRIPS 4 & 5 - exact duplicates

wave 2 CSID 19082 TRIPS 1 & 2 & 3 & 4 - distinct TRIPS, although perhaps should be a single multidest trip

wave 2 CSID 19986 TRIPS 1 & 2 - everything same except beaches, second beach seems unusual, drop??

wave 3 CSID 11923 TRIPS 8 & 9 - exact duplicates

wave 3 CSID 14637 TRIPS 7 & 8 - exact duplicates

wave 3 CSID 15512 TRIPS 1 & 2 - different beaches, different times, but overlap

wave 4 CSID 10297 TRIPS 2 & 3 - different beaches, different times

wave 4 CSID 11923 TRIPS 13 & 14 & 19 & 20 - pairs of dups

wave 4 CSID 14685 TRIPS 1 & 3 - distinct nonoverlapping times

wave 4 CSID 14908 TRIPS 3 & 4 & 6 & 7 & 8 - distinct, although 3 & 4 overlap

wave 4 CSID 18925 TRIPS 3 & 11 - appear to be duplicates

wave 4 CSID 19566 TRIPS 6 & 7 - duplicates, but other than date matches many other TRIPS - can't fix so drop

wave 5 CSID 10805 TRIPS 3 & 4 & 5 & 6 - distinct TRIPS, pairs of adjacent - probably should be multidest single TRIPS

wave 5 CSID 11456 TRIPS 16 & 17 - 2 adjacent TRIPS on same day to same beach

wave 5 CSID 12774 TRIPS 12 & 13 & 20 & 21 & 29 & 30 - same beach, same 3 days, but distinct nonoverlapping TRIPS

wave 5 CSID 13183 TRIPS 9 & 10 - duplicates

wave 5 CSID 19943 TRIPS 19 & 20 - duplicates

wave 5 CSID 30125 TRIPS 7 & 8 - duplicates

wave 5 CSID 30638 TRIPS 2 & 3 - duplicates, no data on time & duration

wave 5 CSID 35765 TRIPS 1 & 2 - duplicates, trip 1 is 3-dest trip 2 is 2-dest suspect that here, as in others, date was misentered, but probably nonrecoverable SO DROP

wave 6 CSID 12499 TRIPS 2 & 3 - first trip is 9am-3pm, second is a 3-day trip, probably okay as is

wave 6 CSID 14685 trps 13 & 14 - exact duplicates, probably miskeyed date on first, but drop

wave 6 CSID 18941 TRIPS 8 & 12 & 9 & 10 & 20 & 21 & 29 & 30 - allday TRIPS usually to venice, sometimes others
 TRIPS 29 & 30 don't overlap
 TRIP 10 probably supposed to be 10/13 since long run of otherwise adjacent days
 TRIP 12 is on either the 15th, 16th, or 17th
 TRIP 20 or 21 is probably miskeyed

wave 6 CSID 30431 TRIPS 2 & 3 & 4 - no overlap although 2 are adjacent in time and trip 4 is first

5. Q4d-Q43

The problem with Q4d – Q4e that was raised early on seems to still be unresolved.

Here is how our question to you on this was formulated the first time.

Water-Based Recreation Results

	Diary 2	Diary 3	Diary 4
Q4D. Went in Water? # Saying Yes/Total Responses (% Yes)	192/914 (21.0%)	96/454 (21.1%)	179/689 (26.0%)
Q4E. How Deep? # Saying “Got Head Wet”/Total Going in Water (% That Got Head Wet of Those Going in Water)	33/204* (16.2%)	40/153* (26.1%)	83/243* (34.2%)

It is strange that for each of the diary surveys the number of responses to Q4E exceeds the number of “yes” responses to Q4D. The respondent is only asked Q4E if they answer “yes” to Q4D. There is a logical inconsistency here. More people indicate that they went in the water in Q4E than say “yes” when asked in Q4D if they went in the water at all.

Here was the Chico response.

I have solved the problem with q4d and q4e responses. There was a skip problem in the program. If q4c-sum3 (or their equivalents) indicated that the respondent participated in an activity that naturally would have taken them into the water (i.e. swimming) they were not supposed to be asked q4d or q4e. However in some cases there was a programming error that took the questionnaire to q4e when it should have gone to q4f. Because this problem only occurred in a few instances, we did not discover it in our data checks. It is easily cleaned up with a macro program. I'll get to that today.

Problem. The data on q4d and q4e remain the same as in the original datasets.

MEMORANDUM

TO: Beach Project Funders
FROM: Craig Mohn
RE: Data Verification and Modeling Approaches
DATE: September 17, 2001

The data provided by Chico contain errors, as do all survey data. Mistakes in data entry and transformation are inevitable, and at acceptably low levels they need not cause serious problems for economic analysis.

My overall impression of the Chico data is favorable. The other dataset I have encountered that is comparable in size is a 13-month panel diary survey covering outdoor recreation by of about 1,000 residents of Montana, conducted by RTI in 1992-93 with a survey budget of about \$1 million. Compared to the RTI survey data, I find that the Chico data has fewer internal contradictions and seems to be of relatively higher quality.

The “rostered” representation of the data, used by Chico from Wave 2 onwards, imposes additional data-checking burdens, so that correct trips and destinations are assigned to the correct individuals. In utility-based choice models, small numbers of trips assigned to the wrong person or destination can lead to disproportionate changes in measures of consumer welfare, so I have taken particular care to ensure that the critical indexing variables seem accurate.

This memo summarizes the efforts I have made, between July and now, to ensure the validity of the key sequencing variables in the data provided by Chico, and to quantify the number of observations containing valid responses for key variables. It is important to note that many economic models can use partial responses, so this memo quantifies how many observations remain when criteria are applied which are relevant to several types of models.

The data provided from Chico (including the first rounds of corrections in response to Chris Busch’s questions) contained 8503 records from 6 waves of diary surveys. Each record corresponds either to a panel member reporting that there were no trips to the beach during that wave or to a panel member reporting a particular trip to a particular beach. There may be multiple records corresponding to a single trip by a panel member if that beach outing included visits to multiple beaches on the same trip (which in general was quite infrequent).

The first round of consistency-checking centered on errors or anomalies in the indexing variables. This resulted in the deletion of 32 records and the modification of 55 other records. The 8471 records remaining contain data on 2179 respondents who reported taking no trips during a wave, and on 5918 trips to 6292 beach destinations undertaken by 1582 respondents over the six waves of the survey.

The second round of consistency-checking involved reviewing the multiple records which correspond to multideestination trips and inspecting the reported start dates and end dates of the trips. This resulted in the deletion of 20 records and the modification of 27 other record. In total, about 1.6 percent of the records were deleted or modified.

A few remaining records contain inconsistencies about dates or trip start times which will probably remain forever ambiguous. However, the partial information contained in them remains useful for some economic modeling approaches. Among these minor

inconsistencies are errors involving the beginning and/or end dates of multiday trips. We cannot employ these observations in site choice models using covariates that change on a daily basis or for models which predict the duration of beach trips. However, there is no reason to doubt the specified choice of beach, and thus these observations are perfectly fine for choice models with covariates which do not vary on a daily basis. Similar comments apply to observations where multiple trips are indicated to occur on the same day but there is overlap, with the start of one beach visit appearing to occur before an earlier visit to another beach has ended.

Panel members were generally very good at revealing key demographic data, with income and availability of a car being the least-revealed among key demographic variables such as education, employment, gender, and ethnicity. People who gave useful responses to these key variables are considered to have provided “good” demographic information in the table below.

In the table below, the column labeled “As Is” gives the number of observations useable with the current zip code data. We are still revising the zip code data, which is essential for accurate modeling of travel cost modeling. We have good zip codes for almost all the recruitment survey and we expect to correct the few remaining gaps by using PC Miler; this has not yet been completed because we have waited three weeks for ALK Technologies to send us the registration key for our batch version of PC Miler. Also, we have just recently received from Chico zip codes for some of the replenishment survey records, which remain to be incorporated. The column “Fix Zip Codes” indicates the number of observations after we complete these revisions to the zip code data.

Table 1 - Number of useable observations in all waves

Model Type	“As Is”	Fix Zip Codes
Simple Choice Model	4288	4333
Separate Count Model	2909 person-waves	
Repeated Choice (FIML)	4185	4228
Repeated Choice(LIML)	4288 & 4312	4333 & 4356

Table 2 - Number of useable observations in wave 4

Model Type	“As Is”	Fix Zip Codes
Simple Choice Model	857	869
Separate Count Model	439 cases	
Repeated Choice (FIML)	819	830
Repeated Choice (LIML)	857 & 847	869 & 858

The assumptions underlying the above table are as follows.

The basic site choice model is assumed to require the CSID (the respondent ID) and zip code, whether the person owns a car and the person's income, as well as the beach destination for that trip.

The total number of beach trips made during a wave – the number of choice occasions – can be modeled in at least two ways: through a count model, in which the dependent variable is the number of trips in that wave, or through a repeated binary choice model in which the respondent is represented as deciding each week whether to take no trips, one trip, two trips, or more than two trips.

The count model is assumed to require a count of trips per wave and good demographic data. There are two ways to estimate a repeated binary choice model. A full-information maximum likelihood (FIML) repeated choice model, which simultaneously uses the destination and date information, would require the same information as the count model and the basic site choice model together with accurate information on the start date for each beach trip. A limited-information maximum likelihood (LIML) repeated choice model can use trips without good start dates for the choice part of the model and dated trips without accurate destinations for the yes/no choice as to whether to take a beach trip.

Any further corrections to the data and modeling approach decisions will affect the numbers in the table.

The research team is continuing to refine the specific criteria for inclusion or exclusion in the models that we estimate, and will finalize these as the updated data files from Chico are integrated to form a final data set.

Appendix contains three memos that I sent to Chico, with Chico's responses to my queries interleaved with my questions to them. In every case, Chico just indicated their approval of what I had proposed or selected one of the alternatives that I offered.

Appendix – Three Memos on Data Issues Sent to Chico Together with Chico’s Reply.

MEMORANDUM SENT TO CHICO ON 31 August 2001

Introduction

Craig has spent the past several weeks doing an independent review of the data to check for errors and anomalies. This memo reflects his progress so far on the ongoing task of ensuring data reliability. There will likely be at least one more set of questions which we will get to Chico very shortly.

A general comment is that one is virtually guaranteed to find some errors when scrutinizing a brand new dataset. Trips containing only partial information, such as date without destination or destination without date, are valuable even when some of the information contained in a record is found to be erroneous – many economic models can incorporate this partial data, including the models we will be using in this project. Even the knowledge that a beach trip was taken to an unknown destination at an unknown time during the wave can be useful, so we must try to preserve as many records as possible in the data-cleaning process.

A key concern that has emerged concerns the trip sequencing variables TRIP and BTRIP. These variables are important to translating the supplied data into a list of trips attributable to individuals useful for econometric analysis. Errors in these will result in incorrect allocation of trips to users, omission of trips, spurious trips and possible other miscounting of demand. Since these counts and choices are key to economic analysis, and directly drive all results of interest, these data-structure variables were the first thing I verified. The rest of this memo will be of a list of inconsistencies and (in most cases) a suggested remedy. Other small problems may well be uncovered in these sequencing variables as analysis progresses and we format and combine the data differently.

In the following list of concerns, errors are identified by wave and either RECID or PREID. The following generally uses the variable RECID where it uniquely identifies an observation, otherwise the variable PREID is used. RECID is not always one-to-one with the records, although it generally is unique. PREID is always unique for each record.

We request that Chico review the findings reported below and comment on proposed resolutions.

Data questions

1) A check that every record was either a trip (TRIP and BTRIP both positive integers) or else a response that no trip (TRIP = 1, BTRIP missing) was taken and the following errors and presumed corrective actions emerged:

Wave 1

- RECID 57 refers to a valid CSID and is presumed a zero response; propose setting TRIP to 1. *****AGREE
- RECID 128 is a trip; propose setting TRIP and BTRIP to 1 *****AGREE

Wave 2

- wave 2 has a lot of observations which have no content except a 99 in the month field after a set of trips for a given CSID (data entry error?); propose deleting RECID 44, 58, 128, 217, 228, 251, 270, 340, 379, 478, 665, 691, 902, 976, 1042, 1058, 1071, 1220, 1404 *****AGREE

Wave 4

- wave 4 PREID 838 has a missing TRIP, propose setting TRIP to 4
****In that it shares RECID 775 with the previous case and is listed as a type 3 trip, I would set TRIP = 3 and BTRIP=2
- wave 4 RECID 1187 contains only MONTH=7, propose deleting it *****AGREE

Another Wave 1 question

- wave 1 RECID 1643-1663 must have errors in TRIP, BTRIP wrong (see diagram below);
propose setting RECID 1652 BTRIP to missing, and
propose setting RECID 1653 - 59 TRIP from 1 to 7 *****AGREE

(columns are PREID RECID CSID TRIP BTRIP)

1640	1637.00	19943	18	1	
1641	1638.00	19943	19	1	
1642	1639.00	19943	20	1	
1643	1640.00	19943	21	1	
1644	1641.00	19943	22	1	
1645	1642.00	19953	1		0
1646	1643.00	19964	1	1	6
1647	1644.00	19964	2	1	
1648	1645.00	19964	3	1	
1649	1646.00	19964	4	1	
1650	1647.00	19964	5	1	
1651	1648.00	19964	6	1	
1652	1649.00	19964	7	1	
1653	1650.00	19964	8	1	
1654	1651.00	19964	9	1	
1655	1652.00	19964	10	1	0
1656	1653.00	19966	11	1	9
1657	1654.00	19966	12	1	
1658	1655.00	19966	13	1	
1659	1656.00	19966	14	1	
1660	1657.00	19966	15	1	
1661	1658.00	19966	16	1	
1662	1659.00	19966	17	1	
1663	1660.00	19967	1		0
1664	1661.00	19968	1	1	5
1665	1662.00	19968	2	1	
1666	1663.00	19968	3	1	
1667	1664.00	19968	4	1	
1668	1665.00	19968	5	1	

GOAL)

2) A check that TRIP sequence has to be increasing from 1 to N revealed the following errors and presumed corrective actions:

Wave 1

- RECID 259; propose changing TRIP to 2 *****AGREE
- RECID 1382-83; propose changing TRIP 1-2 *****AGREE
- RECID 1585; propose changing BTRIP to 1 ****AGREE

Wave 2

- PREID 661, 662, 663; propose deletion for these because they have a duplicate RECID to prior, and unusual beaches for this CSID

***These beaches came from the "any that we missed??" section at the end of the questionnaire . The partial data came from the "usual trip" section. THEY were part of the original Record #645 and therefore have the same RECID. I would keep them.

- PREID 1170-72; propose dropping these – they are the same day as a prior trip, are santamonica and venice a likely single-trip pair?

***These cases came in from the "any we forgot?" section also. Instead of dropping them why not include them as BTRIP = 2 for the appropriate dates under the Venice Trip. Change triptype (WT) to reflect multiple beach trip Like such:

1170	1111	16209	2	2
1171	1111	16209	3	2
1172	1111	16209	3	2
1173	1112	16209	2	1
1174	1113	16209	3	1
1175	1114	16209	4	1
1176	1115	16209	5	1
1177	1116	16209	6	1

- RECID 1299; propose to change TRIP to 2 from 22 ***** AGREE

Wave 4

- PREID 819-24; propose changing TRIP to 2-7 (from duplicate RECID 757)
***AGREE ...but should drop PREID 824 because it lacks any real data
- PREID 1129-31; propose dropping since appear to replications of RECID 1054
****AGREE
- PREID 1227; propose changing TRIP to 7 , duplicate RECID problem
*****AGREE

Wave 5

wave 5 RECID 1161-63; propose changing TRIP to 18-20 because of duplicate RECID 1130 *** you refer to PREID 1161-63, These are some of the "any we left out" group. I AGREE with your solution.

Wave 6

- PREID 587-89; propose dropping these since they are duplicates of RECID 566
*****These are from the "any other??" group, they appear to be second beaches on the same days that they went to santa monica. I propose marking those with the appropriate trip # and changing BTRIP to 2 and changing the trip type (WT =3) see below

PREID, RECID, CSID, TRIP, BTRIP

586	566.00	15987	1	1
587	566.00	15987	1	2
588	566.00	15987	2	2
589	566.00	15987	3	2
590	567.00	15987	2	1
591	568.00	15987	3	1
592	569.00	15987	4	1

593	570.00	15987	5	1
594	571.00	15987	6	1
595	572.00	15987	7	1

- PREID 940; propose to change TRIP to 7, duplicate RECID 915 ***AGREE
- PREID 1176; propose dropping this, duplicate RECID 1143, no date, no details *****AGREE

3) A check that the destination sequence has to run from 1 to M revealed the following problems and possible resolutions:

Wave 1

- RECID 110; propose setting BTRIP to 1 *****AGREE
- RECID 437; propose setting BTRIP to 1 *****AGREE
- RECID 1831; propose setting TRIP to 1 and BTRIP to 1
- *****NO RECID1831 exists in data (perhaps a typo?)
- RECID 1495-1500 are scrambled. proposal: delete RECID 1495-98, change 1499-1500 to TRIP 1 BTRIP 7-8. **** WOW What a Mess!! After a bit of figuring I AGREE with your change.

4) A check that no CSID have both BTRIP missing (indicating no trips that wave) and present (indicating trips) revealed the following problem and possible resolution:

- **Wave 1:** PREID 828 and 846 are partial duplicates of other observations; propose deleting them ***** AGREE

5) A look at non-trip (BTRIP missing) records which have a start date revealed: (Note that several of these have no obvious resolution)

Wave 1

- PREID 120; propose setting BTRIP to 1 *****AGREE

Wave 2

- RECID 468 is confusing - has partial date but no beach
***** Note: Q1AA=5 (no trips) Recommend dropping all trip data after Q1AA for this case and just keeping the demographic & module part (q141 and beyond)
- RECID 1143 is confusing - has date but no beach
- *****same as previous case

Wave 3

- RECID 818 is confusing - has partial date but no beach
*****same as previous case

Wave 4

- RECID 249 is confusing - has date but no beach

****This is strange! They indicate no trips (Q1AA=5), but should have data for the module (but that is missing). I would drop everything after Col G but keep them in the database since that CSID shows up in other waves. You may need it for cross-wave analysis.

Wave 5

- RECID(preid) 1053 is incomplete, propose calling it a no trip **AGREE
Leave TRIP = 1, BTRIP = (blank) delete Data between Q1AA and Q15 (save the demographics and module data at the end)

*******NOTE***** In the Wave 5 merged set Cols AK,AL,AM were included to assure an accurate merge. They should be deleted before analysis.**

Wave 6

- wave 6 RECID 1093 has only a partial date *** drop trip data since they indicated no trips (Q1AA=5) but keep everything from Q16 and beyond

6) In summary, these changes leave:

8463 records (beach destinations)

1034 participants

2179 person-wave reports involving zero trips

1582 person-wave reports involving 5917 trips

5560 of these trips appear to refer to a valid beach and have a valid date

4383 of these trips also have generally good demographic info

1766 of the zero-trip reports have generally good demographic info

End of Memorandum

MEMORANDUM SENT TO CHICO ON 8 September 2001

Here is a list of further discrepancies we have found in the data. We suspect that some of the discrepancies may be difficult to resolve, however resolution will result in a larger dataset. Note that for some types of analysis that the Research Team will conduct, an invalid observation requires that a larger number of related observations be discarded. We would like to avoid this if possible, so we would greatly appreciate your help in resolving the issue.

Discrepancies

1) A check that the Q4A multiday variable was set for every valid trip with a distinct start date and end date yielded the following discrepancies.

******Note: Q4A did not exist for paper records and therefore is blank for many records in the first wave (merged)**

Chico should verify that the following observations are truly multiday trips although Q4A is not set to 1:

wave 1 CSID 10316 TRIP 6 - end date next day ***OK

wave 1 CSID 15949 TRIP 3 - dates 12/14 - 12/15****OK

wave 1 CSID 19239 TRIP 5 - dates 12/15 - 12/17 *****OK

wave 1 CSID 19986 TRIP 1 *****OK

wave 1 CSID 19986 TRIP 3 *****OK

wave 4 CSID 12350 TRIP 9 - only trip with end date, all others indicated as nonmultiday

****raw data indicate a multiday trip but suspect error since record includes time data, suggest DROP end-date

The following trips have problems:

wave 6 CSID 31187 TRIP 1+2 – trip 2 overlaps 3,4,5, and 6 (see below)

no obvious solution, drop this person from the wave *** TRIP1: Delete end-date from Trip1 Btrip1 , keep as-is Trip1Btrip2 (from "any other?" section) Change

*** TRIP2: Delete end-date

*** TRIP 3 change dates to start: 10/27 end 10/29 Change Q4B = 26 Change WT=2

*** TRIP 4 OK ***** TRIP5 OK ***** DROP Trip6 & Trip7

You sent me:

961.00	31187	1	1	5	1	1	2	1	3	10	1	10	1
961.00	31187	1	2							10	1		
962.00	31187	2	1							10	17	10	29
963.00	31187	3	1							11	19		
964.00	31187	4	1							11	26		
965.00	31187	5	1							11	27		
966.00	31187	6	1							11	27		
967.00	31187	7	1										
968.00	31204	1	1	2	1	1	1	1	1	10	7		
969.00	31204	2	1							11	11		
970.00	31216	1		0	5								

I say change to:

31187	1	1	5	1	1	2	1	3	10	1	10	1	3	177	1
31187	1	2							10	1			3	20	1
31187	2	1							10	17			3	177	1
31187	3	1							10	27	10	29	3	26	2
31187	4	1							11	26			3	31	1
31187	5	1							11	27			3	62	1
31187	6	1							11	27			3	62	1
31187	7	1													
31204	1	1	2	1	1	1	1	1	10	7			3	79	1
31204	2	1							11	11			3	94	1

wave 1 CSID 13952 TRIP 3 - dates 12/17 to 1/3 - another trip is in middle -
 "this observation was reconstructed from paper records" in comment - possibly
 this is a 4th trip on 1/3 or some data entry error
 resolution - drop the end date or discard this CSID in this wave
 ****I Would drop the end date and classify as a single day trip on 12/17

2) A check that observations with a valid-appearing beginning date (Q4) have a valid month (Q4MT) revealed the following:

Note that there are no obvious resolutions to the ambiguity here, so these observations will be lost for any analysis requiring dates. There are a lot (34) of trips here.

wave 1 CSID 14752 TRIP 3 could be either dec or january so keep it coded as bad
 **Check with original data, should be JAN
 wave 1 CSID 15060 TRIP 1 could be either dec or january so keep it coded as bad
 ***Checked with raw data should be DEC.
 wave 1 CSID 15659 TRIP 1-26 all could be either dec or jan
 ***check of original paper record shows Month missing for all records -- keep as missing
 wave 2 CSID 12309 TRIP 1-4 could be either feb or march

3) Miscellaneous errors

wave 1 preid 988 *****missing paper record DROP
wave 1 preid 996 *****no data DROP
wave 2 preid 347 *****DROP
wave 2 preid 21 - change date (Q4) from 9 to 19, sequence looks more reasonable???
***AGREE
wave 6, CSID 19867, TRIP 2, (Recid 778) Q4= 99 not allowed, no WT.
Delete this record? *****AGREE
wave 6, CSID 30037 (Recid 807), WT missing, change it to 1? *****AGREE

4) Trips which start on the same day.

Many of these trips appeared to be possible duplicate records. However, it is equally likely that the date was entered incorrectly and the "duplicate" observation(s) refer to distinct trips.

Duplicates should be dropped, otherwise the dates should be corrected.

Note that it is possible for multiple trips to occur on the same day, so we examined starting time and duration where available.

wave 1 CSID 13927 TRIP 3 & 4 appear to be duplicates to same beach *****Agree drop Trip 4 and adjust trip #s on later trips
wave 1 CSID 15370 TRIP 24 should probably be on the 24th rather than the 20th, look at data *****Agree, typo in entry of paper data
wave 1 CSID 19239 TRIP 6 & 11 - both on same day- no time given...possibly true...
*****Check of original SPSS data shows paper records indicate both those trips are on 1/1 I would code as multibeach trip given same TRIP designation and BTRIP 1 & 2 resequentially.

wave 2 CSID 10124 TRIPS 8 & 9 - same day same beach, same time***DROP trip9
wave 2 CSID 11698 TRIPS 1 & 2 - everything same except duration ***DROP trip2
wave 2 CSID 12000 TRIPS 1 & 2 - different beaches, different times, but trip 2 has zero time spent *** ASSUME separate trips; change time spent on trip 2 to MISSING (or bad data)
wave 2 CSID 12047 TRIPS 3 & 4 - exact duplicate entries ***DROP trip4
wave 2 CSID 12774 TRIPS 8 & 9 & 24 & 25 & 28 & 29 - 9 and 25 are dups, 28 & 29 appear distinct
***Drop Trip9, 25, Keep 28 & 29 as separate trips
wave 2 CSID 13896 TRIPS 16 & 17 - appear distinct, yet same time, same day, different beaches - no clue what to do*** Since they occupy same time frame, Recode as a multibeach trip; Recode TRIP 17 to 16 with BTRIP = 2
wave 2 CSID 14304 TRIPS 4 & 5 - different beaches, different times *** Leave as-is consider them to be separate trips
wave 2 CSID 14884 TRIPS 11 & 12 - different beaches, different times *** Keep as separate trips
wave 2 CSID 17293 TRIPS 4 & 5 - exact duplicates ***** DROP Trip5

wave 2 CSID 19082 TRIPS 1 & 2 & 3 & 4 - distinct TRIPS, although perhaps should be a single multidest trip *****AGREE

wave 2 CSID 19986 TRIPS 1 & 2 - everything same except beaches, second beach seems unusual, drop?? *****AGREE

wave 3 CSID 11923 TRIPS 8 & 9 - exact duplicates ***DROP Trip9

wave 3 CSID 14637 TRIPS 7 & 8 - exact duplicates *****DROP Trip8

wave 3 CSID 15512 TRIPS 1 & 2 - different beaches, different times, but overlap **CHANGE to multibeach trip

wave 4 CSID 10297 TRIPS 2 & 3 - different beaches, different times ***Keep as separate trips (one in the morning and one at night)

wave 4 CSID 11923 TRIPS 13 & 14 & 19 & 20 - pairs of dups *****DROP 14 & 20

wave 4 CSID 14685 TRIPS 1 & 3 - distinct nonoverlapping times ** KEEP as separate trips

wave 4 CSID 14908 TRIPS 3 & 4 & 6 & 7 & 8 - distinct, although 3 & 4 overlap*****Change 3&4 to multibeach trip, Keep 6,7,8 as separate trips

wave 4 CSID 18925 TRIPS 3 & 11 - appear to be duplicates ***** Change month on trip3 to 6 (typo on entry)

wave 4 CSID 19566 TRIPS 6 & 7 - duplicates, but other than date matches many other TRIPS - can't fix so drop *****AGREE

wave 5 CSID 10805 TRIPS 3 & 4 & 5 & 6 - distinct TRIPS, pairs of adjacent - probably should be multidest single TRIPS ***AGREE

wave 5 CSID 11456 TRIPS 16 & 17 - 2 adjacent TRIPS on same day to same beach ***** DROP Trip 17 adjust start time on 16 2:30pm) and time spent (5hr30min)

wave 5 CSID 12774 TRIPS 12 & 13 & 20 & 21 & 29 & 30 - same beach, same 3 days, but distinct nonoverlapping TRIPS *** KEEP all as separate trips

wave 5 CSID 13183 TRIPS 9 & 10 - duplicates ***** DROP 10

wave 5 CSID 19943 TRIPS 19 & 20 - duplicates *****DROP 20

wave 5 CSID 30125 TRIPS 7 & 8 - duplicates ***** DROP 8

wave 5 CSID 30638 TRIPS 2 & 3 - duplicates, no data on time & duration ***DROP 3

wave 5 CSID 35765 TRIPS 1 & 2 - duplicates, trip 1 is 3-dest trip 2 is 2-dest suspect that here, as in others, date was misentered, but probably nonrecoverable SO DROP ***I say DROP Trip1; KEEP Trip2

wave 6 CSID 12499 TRIPS 2 & 3 - first trip is 9am-3pm, second is a 3-day trip, probably okay as is ***AGREE

wave 6 CSID 14685 trps 13 & 14 - exact duplicates, probably miskeyed date on first, but drop *** AGREE DROP Trip14

wave 6 CSID 18941 TRIPS 8 & 12 & 9 & 10 & 20 & 21 & 29 & 30 - allday TRIPS usually to venice, sometimes others

TRIPS 29 & 30 don't overlap *****Leave as-is

TRIP 10 probably supposed to be 10/13 since long run of otherwise adjacent days *****AGREE

TRIP 12 is on either the 15th, 16th, or 17th *****AGREE

TRIP 20 or 21 is probably miskeyed *** to follow pattern change, Change date on Trip21 to 10/31

wave 6 CSID 30431 TRIPS 2 & 3 & 4 - no overlap although 2 are adjacent in time and trip 4 is first ****I would consider coding 23&4 as multibeach trips
WT=3

5. Q4d-Q43

The problem with Q4d – Q4e that was raised early on seems to still be unresolved.

Here is how our question to you on this was formulated the first time.

Water-Based Recreation Results

	Diary 2	Diary 3	Diary 4
Q4D. Went in Water? # Saying Yes/Total Responses (% Yes)	192/914 (21.0%)	96/454 (21.1%)	179/689 (26.0%)
Q4E. How Deep? # Saying “Got Head Wet”/Total Going in Water (% That Got Head Wet of Those Going in Water)	33/204* (16.2%)	40/153* (26.1%)	83/243* (34.2%)

It is strange that for each of the diary surveys the number of responses to Q4E exceeds the number of “yes” responses to Q4D. The respondent is only asked Q4E if they answer “yes” to Q4D. There is a logical inconsistency here. More people indicate that they went in the water in Q4E than say “yes” when asked in Q4D if they went in the water at all.

Here was the Chico response.

I have solved the problem with q4d and q4e responses. There was a skip problem in the program. If q4c-sum3 (or their equivalents) indicated that the respondent participated in an activity that naturally would have taken them into the water (i.e. swimming) they were not supposed to be asked q4d or q4e. However in some cases there was a programming error that took the questionnaire to q4e when it should have gone to q4f. Because this problem only occurred in a few instances, we did not discover it in our data checks. It is easily cleaned up with a macro program. I'll get to that today.

Problem. The data on q4d and q4e remain the same as in the original datasets.

**** This is a problem Allen was working on. My only suggestion for reconciling the problem is to search out those cases where q4e was answered but q4d was not and insert (by hand) "1" as the response for q4d as the only logical alternative.

END OF MEMORANDUM

MEMORANDUM SENT TO CHICO ON 13 September 2001

We have noted the following inconsistencies in the data. In some cases the remedy is obvious, and we ask Chico to state their agreement or alternative solution. In other cases the remedy is less clear and we would appreciate Chico's suggestions.

Agree with proposed remedy?

Wave 1 RECID 1381 – this is a new CSID, set TRIP = 1, BTRIP = 1 ****AGREE

Wave 4 CSID 16198 TRIPS 1 & 2 overlap (6/18-7/18 and 7/8-7/9) – clearly need to set Q6BR = 6 for TRIP 1, which is single-day *****AGREE

Wave 4 CSID 18925 TRIPS 3 & 4 overlap (6/12-7/12 and 6/15) – clearly need to set Q6BR = 6 for TRIP 3***** My data set does not indicate Trip3 as a multi-day. I think that the 7/12 date is a typo considering it appears down in Trip11. I would change the date on Trip3 to 6/12. See below for how it appears in my dataset:

18925	1	1	14	1	1	7	1	7	6	6
18925	2	1							6	8
18925	3	1							*7	12
18925	4	1							6	15
18925	5	1							6	18
18925	6	1							6	22
18925	7	1							6	7
18925	8	1							7	2
18925	9	1							7	4
18925	10	1							7	6
18925	11	1							7	12
18925	12	1							7	18

*change to 6

Wave 5 RECID 14 – empty, labeled “ignore this row” – delete this observation
***AGREE

Wave 5 CSID 12606 TRIPS 3 & 4 overlap – clearly need to set Q6BR = 8 for TRIP 3

Suggestions?

There is no clear resolution to the following. Unless Chico has evidence of data-entry error, we suggest leaving these as is and noting that there are some data inconsistencies in the dataset, since there are undoubtedly other minor errors in coding which will be discovered as time goes by.

Wave 4 CSID 10728 TRIP 4 (7/31-8/3) overlaps Wave 5 CSID 10728 TRIP 1 (8/1-8/3)
**Several solutions are possible: A) for Wave 4 CSID 10728 TRIP 4 list as 1 day trip, (7/31) and keep Wave 5 CSID 10728 TRIP 1 (8/1-8/3) as-is or B) keep Wave 4 CSID 10728 TRIP 4 (7/31-8/3) as-is and delete Wave 5 CSID 10728 TRIP 1

Wave 5 CSID 12960 TRIP 1 (9/27-10/2) overlaps Wave 6 CSID 12960 TRIP 1 (10/1-10/2) . **Several solutions are possible: A) for Wave 5 CSID 12960 TRIP 1 (9/27-10/2) change ending date to 9/30 and keep Wave 6 CSID 12960 TRIP 1 (10/1-10/2) as-is, or B) keep Wave 5 CSID 12960 TRIP 1 (9/27-10/2) as is and delete Wave 6 CSID 12960 TRIP 1.

Wave 6 CSID 14231 TRIPS 1 & 2 overlap (10/14-10/16 and 10/15) ***DROP Trip2

Wave 1 CSID 19875 TRIP 1 (1/31-2/5) overlaps Wave 2 CSID 19875 TRIP 1 (2/1-2/4) ***same as first case.

END OF MEMORANDUM

Corrections July 2001.

MEMORANDUM

TO: The File
FROM: Chris Busch (Research Team)
DATE: 22 June 2001
RE: Additional Questions on Data Checking (Revised)
CC: Allen Lunde (Chico State SRC), Michael Hanemann (Research Team)

Last corrections prior to draft final dataset production in July 2001.

Diary 2

CSID 15815, RECID 1033. Q6B given as 6, but looks like it should be 2. Agree?
(Note: Question submitted to David Philhour. **AGREES**)

Diary 3

2. CSID 18387, RECID 855-856. These show same start and end date but WT =2. I suspect WT should be changed to =1? *****the conflict is between q4a where they answered that this is a multi-day trip and then give dates showing same start and finish. I feel that our policy should be to call these Type 1 trips**

Diary 6

CSID 19867, RECID 777: Q6 (return month) is listed as 12, outside the two-month acceptable range that we have defined, though not logically impossible. Interviewer

writes “3 or 33 days.” What do you think? (Note: Question submitted to David Philhour.)

****change dates to 11/1-11/3

MEMORANDUM

TO: David Philhour (Chico State SRC)
FROM: Chris Busch (Research Team)
DATE: 22 June 2001
RE: Additional Questions on Data Checking (Revised)
CC: Allen Lunde (Chico State SRC), Michael Hanemann (Research Team)

Introduction

Thanks to the Chico State SRC for their response to previous questions from the research team. This memo raises just a few more questions that follow from a recently completed exhaustive tabulation of all variables in the latest version of the diary survey datasets. Below we also list a few corrections agreed to by both Chico and the research team (as detailed in a memo from Chico dated 19 June 2001) that have not been implemented in the most recently delivered datasets. Please make these changes. We make this request so that all changes to the datasets are done in one place, which will reduce the chance of confusion about which changes have been made and which are the most up to date versions of the datasets. For the same reason, we ask you to make a few revisions to codebooks. The specific issues for each diary survey are detailed under the relevant heading.

Diary 1

1. Please add WT=6 and WT=7 to codebook for Diary 1.
2. Also, for all codebooks, please remove superfluous <5> value from Q4A question.
3. Regarding the \$75 parking fee reported for CSID 15997 and CSID 19943: Please change these to zero (0). That is, unless the \$75 per year parking fee refers to beach parking. I’m assuming this is residential parking. If the \$75 is for beach parking, the research team will need to consult. In this case, we may wish to divide up the fee among trips.

Diary 2

CSID 15815, RECID 1033. Q6B given as 6, but looks like it should be 2. Agree?
(Note: Question submitted to David Philhour.)

Diary 3

1. Variables 15b and beyond (pages 34-47 of the codebook) are missing!?!?
2. CSID 18387, RECID 855-856. These show same start and end date but WT =2. I suspect WT should be changed to =1? (Note: Question submitted to David Philhour.)

Diary 5

1. Tabulated data for VAR q18 are:

Q18	Freq.	Percent	Cum.
1	377	74.80	74.80
88888	127	25.20	100.00
Total	504	100.00	

Why don't we have a list of beach codes here? That is what the codebook says there should be.

2. Dataset does not contain vars Q18a-Q18d. (They are listed in the codebook.)

These are follow up questions to Q18. I can understand why there wouldn't be large amounts of data and even why a couple of these might be empty. But seems one would expect at least some "don't know" or "no more" type answers to Q18a?

3. Tabulated data for VAR q22 are:

Q22	Freq.	Percent	Cum.
0	140	100.00	100.00
Total	140	100.00	

Again, why don't we have a list of beach codes here? That is what the codebook says there should be.

4. Vars Q22a1 and Q22b1 are not included in the dataset. (They are listed in the codebook.)

As above, seems one would expect at least some "don't know" or "no more" type answers here?

5. Q22c indicates respondent remarks should be recorded (e.g. spelled out), but data show only 0s? Same for Q22c and Q23a.
6. Please remove row 20 (CSID 10119, RECID 14) as we have agreed would be appropriate.
7. Regarding this case-- CSID 19986 (PREID 1050). End date listed is in month before start date. *****date changed to 8/9-8/17. There's still a problem here. The change means that the trip detailed in this row is in fact contained in the trip detailed in the two rows above. Suggest we change WT and dates so that this is a beach visited on the larger trip outlined in the two rows above. (Note: Question submitted to David Philhour.)

Diary 6

CSID 19867, RECID 777: Q6 (return month) is listed as 12, outside the two-month acceptable range that we have defined, though not logically impossible. Interviewer writes "3 or 33 days." What do you think? (Note: Question submitted to David Philhour.)

Email Traffic the week of June 18

Message from David Philhour, Chico State
Tues, 19 June 2001

Chris --
Here are the responses to your memo of 6/9
Regarding the cleaning of the data.
Allen will send new datasets.

(Attachment <<Busch.doc>> that responds to Research Team memo)

Message from Allen Lunde, Chico State
Tuesday, 19 June 2001

Chris

I've attached the clean data files as well as the raw data. I will not be able to complete the production report till next week.

Allen

(Attachment: six new diary survey datasets)

Message from Chris Busch, Research Team
Tuesday, 19 June 2001

David,

I've only had a chance to briefly look over this, but it looks great. Thanks for you detailed consideration of these questions.

Chris

Message from Chris Busch, Research Team
Tuesday, 19 June 2001

Allen,

Thanks for the reformulated data and word on the production report.

One thing: the Beaches Plan and Wave 1 Transformation spreadsheets look the same to me. Please let me know if I'm wrong on this.

Have a good journey, wherever you're off to.

Chris

Message from Allen Lunde, Chico State
Tuesday, 19 June 2001

Chris,

I'm going to let David answer this, because I'm not sure.

Allen

Message from David Philhour, Chico State
Tuesday, 19 June 2001

Chris --

Beaches Plan should be a multi-colored mapping of variables from Wave2 thru Wave6 of all variables in the "use" section of the interviews.

Wave 1 Transformation matrix1 consists of 2 worksheets: Machine1 lists more than 1500 "use" variables and their mapping into the variables used for other waves. Variables that contain no data are highlighted in YELLOW. The second sheet called Paper 1 shows the mapping of the paper records for subjects having more than 5 trips in one month. Not that this matrix is transposed and trips are listed in columns and variables in rows. Even

though the variable names are not included the identifiers should leave no doubt as to their position in the final matrix.

Hope this helps.

david

**Message from Chris Busch, Research Team
Wednesday, 20 June 2001**

David,

Thanks for the explanation of the Wave 1 transformation.

What I meant to ask is, regarding the set of files Allen sent to me yesterday, have either the Beaches Plan or Wave 1 Transformation spreadsheets changed since they were sent to me the first time? They both look the same to me as the original version I have. Have you seen the questions I sent to Allen on this? I attach an abbreviated memorandum that excerpts questions relevant to data transformation. I also attach a spreadsheet I created where I try to show how beach trips (focusing on where the beach is listed) can be reorganized in a columnar fashion ("Sketch of Transformation"). Do you have a final copy of the diary survey (which I apparently don't?)

I'm hopeful we can move forward on this track-- developing an explanation of the data transformation process that will be coherent to an outside observer-- as well as on the data checking track.

Thanks,

Chris Busch

**Message from Chris Busch, Research Team
Wednesday, 20 June 2001**

David,

First, thanks to you and Allen for all your work responding to our previous concerns.

The attached memo requests that you address a few more of these data checking questions. (The questions are not nearly so numerous, as you will see. Note also that I've checked these questions against the latest datasets that Allen sent to ensure that they are still relevant.)

Still wondering where we stand on fleshing out our explanation of the data transformation process?

Please advise on your ability to respond to data checking and transformation questions so that I can keep the funders informed.

Thanks very much!

~Chris Busch

(Attachment: Memo to Chico June 21)

**Message from David Philhour, Chico State
Friday, 22 June 2001**

Chris --

Questions regarding codebook and questions in the modules (q14 and beyond) should be addressed to Allen. I worked only on the Use section of the data.

I looked at the transformation plan you sent in your last post and had difficulty matching it to the Beaches Pan.xls that I produced. For example, Why are there 10 rows indicating singleday/singlebeach trips? The designations in Col 3 (before the variable names in the parentheses) do not correspond to any numbers I have.

Allen was going to send you and new copy of the interview schedule that uses the variable names that appear in the Beaches Plan.xls. Did you get this document?

Many of your questions should be answered by this new schedule.

Please make the changes to your copies of the data as indicated in the attached reply to your memo of 6/21

david

**Message from Chris Busch, Research Team
Friday, 22 June 2001**

David,

In my "sketch of transformation of beach survey data" spreadsheet I wasn't trying to match your Beaches Plan spreadsheet. To match your Beaches Plan spreadsheet, I would have to understand it. But with only the survey text and 1-page explanation you wrote in hand, I haven't been able to completely understand the Beaches Plan spreadsheet.

You write: "Fore example, Why are there 10 rows indicating singleday/singlebeach trips? The designations in Col 3 (before the variable names in the parentheses) do not correspond to any numbers I have."

There are 10 rows for single day, single beach trips because a respondent could provide detailed information on up to 5 trips for each month. I attach the survey questions that show precisely the manner in which the respondent can do this. The attachment shows the 4a-4l loop that I refer to-- this is where the respondent gives detailed information on the trip. In column 3 (or C, if you prefer), I list the questions that take the respondent to this 4a-4l loop. The 10 questions that do so are: 4, 8a,8b,8c,8d, 9,10a,10b,10c, and 10d.

Do you see why there are 10 rows for single day, single beach trips now?

I may not have the absolutely final survey text, though the diary survey was already being implemented on the date that is listed at the top of the survey text that I do have. So I don't believe it is possible that the final diary survey text differs very much from the one I have.

You write: "Allen was going to send you and new copy of the interview schedule that uses the variable names that appear in the Beaches Plan.xls. Did you get this document?"

No, I don't have this. Since Allen is out of town, if you have this, please send it ASAP. Do you have a copy of the survey text? Part of the problem may be that we are coming at this from different angles.

Thank you.

Chris Busch

p.s. The "sketch of transformation of beach survey data" was developed with a number of different possible purposes in mind. I put together the spreadsheet in part to disentangle in my mind the location in the survey text of the questions on the different types of trips. I had also hoped that you could look over what I produced so that you could confirm that my understanding is correct. In the "sketch" spreadsheet I sought to show in a simple way how the various data could be reorganized so that each single beach visited falls in a single column-- this is column Q4B in the transformed datasets. I had also thought the spreadsheet could assist in the larger effort of providing documentation on the transformation of raw to final data format.

**Message from Chris Busch, Research Team
Friday, 22 June 2001**

David,

You write: "Questions regarding codebook and questions in the modules (q14 and beyond) should be addressed to Allen. I worked only on the Use section of the data."

So, I have extracted the relevant questions from the memo I sent yesterday. Please address these.

Diary 2

CSID 15815, RECID 1033. Q6B given as 6, but looks like it should be 2. Agree?

Diary 3

CSID 18387, RECID 855-856. These show same start and end date but WT =2. I suspect WT should be changed to =1? Agree?

Diary 5

Regarding this case-- CSID 19986 (PREID 1050).

My initial comment.

End date listed is in month before start date.

Chico response

*****date changed to 8/9-8/17.

My comment

There's still a problem here. The change means that the trip detailed in this row is in fact contained in the trip detailed in the two rows above. Suggest we change WT and dates so that this is a beach visited on the larger trip outlined in the two rows above. Agree?

Diary 6

CSID 19867, RECID 777: Q6 (return month) is listed as 12, outside the two-month acceptable range that we have defined, though not logically impossible. Interviewer writes "3 or 33 days." What do you think?

End questions

Do you per chance know when Allen will be back in town?

thanks

Chris

**Message from Chris Busch, Research Team
Friday, 22 June 2001**

Allen,

Based on David Philhour's message of today, I attach a revised memo with questions on datasets. I have noted in the revised memo which questions I have sent separately to David. Perhaps of greatest concern is the large number of variables missing in Diary 3.

Do you have a final version of the diary survey that you can send me? (I'll ask the research team as well, though Michael is out of the country. Also, you may know better which version is actually "final" since you did the programming.)

David writes, "Allen was going to send you and new copy of the interview schedule that uses the variable names that appear in the Beaches Plan.xls. Did you get this document?" Please send ASAP.

Please advise as to your schedule. I'm wondering when you are back in town?

Thanks,

Chris Busch

(Attachment: Memo to Chico 22 June)

MEMORANDUM

TO: Beach Project Funders
FROM: The Research Team
RE: Looking for and scrutinizing outliers
DATE: 29 May 2001

This memo contains notes on the search for outlier values (both within and outside the range of acceptable values) and the scrutiny of these.

Diary 1

1. Multi day trip length. Positive values are for multi day trips.

length	Freq.	Percent	Cum.
0	1646	98.33	98.33
1	14	0.84	99.16
2	7	0.42	99.58
3	3	0.18	99.76

5		2	0.12	99.88
6		1	0.06	99.94
17		1	0.06	100.00

Total		1674	100.00	

Problem with CSID 19986, RECID 1655. For return date, shows 30 in month column and nothing in day. Should be 1/30.

Note17 day trip (CSID 13952, RECID 846) looks reasonable. Winter holiday with a group.

Cases where start and end dates show a one day trip (i.e. they are the same day) and the WT variable shows a multi day trip (WT = 2).

- 10744, 186 (format is: CSID, RECID)
- 10744, 187
- 12382, 465
- 12382, 466
- 13679, 769
- 18860, 1442
- 18891, 1447
- 19244, 1522
- 19244, 1523

2. Parking cost in dollars.

Q4I		Freq.	Percent	Cum.
0		448	64.83	64.83
1		49	7.09	71.92
2		27	3.91	75.83
3		23	3.33	79.16
4		16	2.32	81.48
5		42	6.08	87.55
6		6	0.87	88.42
7		20	2.89	91.32
8		6	0.87	92.19
10		1	0.14	92.33
16		1	0.14	92.47
20		1	0.14	92.62
40		4	0.58	93.20
75		39	5.64	98.84
88888		8	1.16	100.00

Total		691	100.00	

Possible Problems

- The \$75 answers fall outside the range of acceptable answers. These occur for CSID 15997 (RECID 1152-1168) and CSID 19943 (RECID 1620-1641).

- The four instances of \$40 parking fees look questionable to me. These are all reported by the same respondent for one day trips to Crystal Cove. In three cases the respondent reports going alone and once with one other person. So a large group of people does not explain the figure. CSID 10297, RECID 80-83.

3. Number of people accompanying respondent to the beach.

Q4J	Freq.	Percent	Cum.
0	431	34.40	34.40
1	356	28.41	62.81
2	128	10.22	73.02
3	86	6.86	79.89
4	46	3.67	83.56
5	30	2.39	85.95
6	33	2.63	88.59
7	11	0.88	89.47
8	4	0.32	89.78
9	49	3.91	93.70
10	5	0.40	94.09
11	2	0.16	94.25
12	2	0.16	94.41
13	19	1.52	95.93
14	1	0.08	96.01
15	6	0.48	96.49
18	1	0.08	96.57
20	1	0.08	96.65
40	1	0.08	96.73
99999	41	3.27	100.00
Total	1253	100.00	

Notes

- The trip with 40 people was a single beach, multi day trip to beach code 302, “an unknown SD Co. beach.” CSID 19875 (RECID 1614). Similar response in Diary 2.
- Spikes in 13s and 9s come from contributions of one respondent each (respectively, CSID 10149, RECIDs 13-29, and CSID 12629, RECIDs 506-552). Can’t tell what kind of trips because of WT problem.

Diary 2

1. Multi day trip length. Positive values are for multi day trips.

length	Freq.	Percent	Cum.
0	1468	97.35	97.35
1	28	1.86	99.20
2	5	0.33	99.54
3	1	0.07	99.60
4	1	0.07	99.67
7	1	0.07	99.73

8		1	0.07	99.80
32		2	0.13	99.93
34		1	0.07	100.00

Total		1508	100.00	

As described below, my best guess is that each of these trips over thirty days is erroneous. These three trips are discussed first.

- CSID 16098, RECID 1090. End date indicates trip length of one month and one day, and this end date falls outside of diary period. Looks like end date month should be changed to March (3) from April (4). This would make the trip an overnighter.
- CSID 17035, RECID 1179. End date indicates trip length of one month and one day. Looks like end date month should be changed to February from March. This would make the trip an overnighter.
- CSID 17161, RECID 1198. Hard to make sense of this trip. Shows a trip of more than one month. Changing month of return make the trip 4 days, more plausible, but this contradicts WT value of 1, which indicates a one day trip.

Another problem. CSID 15502, RECID 1004, shows an 8 month long trip. When one makes the logical correction (change date of return month to March), this causes a contraction in that start and end dates show a one day trip and the WT variable shows a multi day trip.

Other cases where start and end dates show a one day trip and the WT variable shows a multi day trip.

- CSID 12197 (RECID 454).
- CSID 12888 (RECID 621).
- CSID 12371 (RECID 479).
- CSID 12888 (RECID 622).
- CSID 14297 (RECID 843).
- CSID 14312 (RECID 856).
- CSID 14638 (RECID 893).
- CSID 15539 (RECID 1012).
- CSID 15550 (RECID 1013).

2. Parking cost in dollars.

Q4I		Freq.	Percent	Cum.
0		386	57.36	57.36
1		25	3.71	61.07
2		56	8.32	69.39
3		68	10.10	79.49
4		13	1.93	81.43
5		49	7.28	88.71
6		36	5.35	94.06
7		11	1.63	95.69

8		9	1.34	97.03
18		1	0.15	97.18
32		1	0.15	97.33
40		1	0.15	97.47
88888		17	2.53	100.00

Total		673	100.00	

Possible Problem

- The 32 value looks like an error. This is for CSID 12734, RECID 552, a 4 person, one day trip to Oceanside city beach. RECID 551 is for the same respondent and for a trip with a similar profile, but the cost of parking was only \$2 for RECID 551. Probably the same for RECID 552.

Other Notes

- The 40 value is for a one person, one beach, one day trip (CSID 13077, RECID 635). No obvious reason for such a cost. No obvious signs of data error.
- The \$18 value is for an overnight trip.

3. Number of people accompanying respondent to the beach.

Q4J		Freq.	Percent	Cum.
0		246	23.91	23.91
1		353	34.31	58.21
2		137	13.31	71.53
3		97	9.43	80.95
4		60	5.83	86.78
5		30	2.92	89.70
6		39	3.79	93.49
7		27	2.62	96.11
8		5	0.49	96.60
9		2	0.19	96.79
10		3	0.29	97.08
11		3	0.29	97.38
13		4	0.39	97.76
14		1	0.10	97.86
15		1	0.10	97.96
20		1	0.10	98.06
30		2	0.19	98.25
40		3	0.29	98.54
88888		11	1.07	99.61
99999		4	0.39	100.00

Total		1029	100.00	

Possible Problems

- CSID 19082, RECID 1289-1292, shows four 1 day, 1 beach trips all occurring on the same day (indicated by WT = 1 for all four rows). This seems to imply that these four trips should be recoded as a 1 day, multi beach trip.

- Contradiction in CSID 15539 (RECID 1012). WT value (=2) shows multi day, but trip has same start and end date, thus indicating single day trip. This row shows 30 people on the respondent's trip.

Other Notes

- 40#1, CSID 19047(RECID 205). Looks like a group/family trip to Huntington City Beach.
- 40#2, CSID 19082(RECID 1293). Some aspects of the trip are similar to others reported trips for this respondent with much fewer people, but no stronger indications of data error.
- 40#3, CSID 19875 (RECID 1614). This is the same respondent who reported a similar trip in Diary 1. Probably reason to accept validity.
- 2nd 30, CSID 18891 (RECID 1271) looks okay.
- 20 value looks okay (CSID 18389, RECID 1248). Respondent's other trip involves 10 other people.

Diary 3

1. Multi day trip length. Positive values are for multi day trips.

Can't look at trip length because of data problem. There are no values for Q6B, Q6DY = no end dates = no multi day trips?

2. Parking cost in dollars.
3. Number of people accompanying respondent to the beach.

Diary 4

1. Multi day trip length. Positive values are for multi day trips.

length	Freq.	Percent	Cum.
0	1339	96.05	96.05
1	18	1.29	97.35
2	16	1.15	98.49
3	6	0.43	98.92
4	4	0.29	99.21
5	2	0.14	99.35
6	1	0.07	99.43
7	3	0.22	99.64
10	1	0.07	99.71
27	1	0.07	99.78
30	1	0.07	99.86
35	2	0.14	100.00
Total	1394	100.00	

Possible Problems

- CSID 14898 (RECID 757). Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.
- CSID 15522 (RECID 836). Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.
- CSID 12350 (RECID373). End date listed is one day before start date.

Other Notes.

- 30 day trip. CSID 16198 (RECID 931). Beaches visited are Mission Beach and Oceanside City Beach.
- 27 day trip. CSID 13988. Hermosa Beach.
- 35 day trip (CSID 12039, RECID 342) starting July 4th. 4 people to Pismo Beach. Perhaps the eight in the return month column should be a seven? Is a trip ever too long to be reasonable?

2. Parking cost in dollars.

Q4I	Freq.	Percent	Cum.
0	378	48.96	48.96
1	15	1.94	50.91
2	37	4.79	55.70
3	26	3.37	59.07
4	35	4.53	63.60
5	73	9.46	73.06
6	129	16.71	89.77
7	48	6.22	95.98
8	3	0.39	96.37
10	6	0.78	97.15
12	1	0.13	97.28
28	1	0.13	97.41
40	1	0.13	97.54
88888	19	2.46	100.00
Total	772	100.00	

- \$40. CSID 12945 (RECID 471). 1 day trip to Venice beach with three people. Maybe they got a parking ticket?
- \$28. CSID 13679 (RECID 598) . 1 day, 1 person trip to Manhattan Beach.

3. Number of people accompanying respondent to the beach.

Q4J	Freq.	Percent	Cum.
0	153	14.19	14.19
1	297	27.55	41.74
2	178	16.51	58.26
3	117	10.85	69.11
4	104	9.65	78.76
5	56	5.19	83.95
6	31	2.88	86.83

7	17	1.58	88.40
8	27	2.50	90.91
9	9	0.83	91.74
10	11	1.02	92.76
12	3	0.28	93.04
13	4	0.37	93.41
14	5	0.46	93.88
15	4	0.37	94.25
17	3	0.28	94.53
19	1	0.09	94.62
20	8	0.74	95.36
24	1	0.09	95.45
25	2	0.19	95.64
26	1	0.09	95.73
35	1	0.09	95.83
40	7	0.65	96.47
88888	38	3.53	100.00

Total	1078	100.00	

Notes

- 4 of the 40s attributable to one respondent. CSID 1186 (RECID 216-219). These look acceptable. Regular biweekly outings.
- 40 #5, CSID 12606(RECID 410). 1 day trip to Newport. Looks reasonable. Person took other trips with relatively large groups.
- 40 #6. CSID 14752 (715). 1 day trip to Dockweiler. No reason to doubt.
- 40#7. CSID 18389 (1051). Again looks okay.
- 35. CSID 19986 (1312). 1 day trip to Bolsa Chica. Can't reject.

Diary 5

1. Multi day trip length. Positive values are for multi day trips.

length	Freq.	Percent	Cum.
0	1534	94.17	94.17
1	42	2.58	96.75
2	18	1.10	97.85
3	13	0.80	98.65
4	7	0.43	99.08
5	3	0.18	99.26
6	2	0.12	99.39
7	4	0.25	99.63
8	1	0.06	99.69
15	2	0.12	99.82
16	3	0.18	100.00

Total	1629	100.00	

Possible Problems

- CSID 31755 (RECID 1355). Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.
- CSID 16258 (RECID 806). Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.
- CSID 10515 (RECID 70). Row shows same start and end date (implicitly defines as single day trip). WT variable shows value of 2, thus implying a multi-day trip.
- CSID 30638 (RECID 1116). End date listed is one day before start date.
- CSID 19986 (PREID 1050). End date listed is in month before start date.
- CSID 30748, RECID 1152, shows one day trip (WT=1) but data on hours at beach reads 24, which implies an overnight trip.

Other notes.

- 16 day trips all for CSID 14978 (rows 697-699—only 1 RECID # for these three rows). Cannot disqualify.
- 15 days trips also look okay. For CSID 19986 (PREID 1049-1050).

2. Parking cost in dollars.

Q4I	Freq.	Percent	Cum.
0	509	58.51	58.51
1	41	4.71	63.22
2	21	2.41	65.63
3	51	5.86	71.49
4	18	2.07	73.56
5	75	8.62	82.18
6	61	7.01	89.20
7	31	3.56	92.76
8	1	0.11	92.87
10	4	0.46	93.33
15	1	0.11	93.45
20	1	0.11	93.56
36	1	0.11	93.68
88888	55	6.32	100.00
Total	870	100.00	

Notes

- Not sure, but \$36 may be an error. This CSID 19572 (RECID 953). RECID 954 is a trip by same respondent to same beach (Newport) also with one companion, but paid nothing for parking for this second trip. Car is the type of transportation for both trips.
- \$20 is for a trip (CSID 30748, RECID 1152) to Del Mar City Beach. Large group (6 other people), may explain this. Also, WT indicates 1 day trip, but otherwise data Q4G show 24 hours at beach, so this may be parking for two cars, for two days.

3. Number of people accompanying respondent to the beach.

Q4J	Freq.	Percent	Cum.
0	301	25.95	25.95
1	294	25.34	51.29
2	196	16.90	68.19
3	107	9.22	77.41
4	68	5.86	83.28
5	84	7.24	90.52
6	43	3.71	94.22
7	17	1.47	95.69
8	8	0.69	96.38
9	6	0.52	96.90
10	9	0.78	97.67
11	4	0.34	98.02
12	1	0.09	98.10
13	2	0.17	98.28
14	2	0.17	98.45
15	3	0.26	98.71
16	1	0.09	98.79
18	1	0.09	98.88
19	1	0.09	98.97
20	3	0.26	99.22
24	2	0.17	99.40
25	2	0.17	99.57
34	1	0.09	99.66
35	2	0.17	99.83
40	2	0.17	100.00
Total	1160	100.00	

Note. Inspection of values 24-40 show they are not inconsistent with other data entries.

Diary 6

- Multi day trip length. Positive values are for multi day trips.

length	Freq.	Percent	Cum.
0	1242	97.18	97.18
1	14	1.10	98.28
2	9	0.70	98.98
3	5	0.39	99.37
4	1	0.08	99.45
6	1	0.08	99.53
9	3	0.23	99.77
11	1	0.08	99.84
12	1	0.08	99.92
88859	1	0.08	100.00
Total	1278	100.00	

Note. Inspection of longer trips suggests these check out.

Apparent problems

- CSID 18407 (RECID 650). Return month, Q6 in this case, shows May (5), which is beyond diary time period and a likely error.
- CSID 14884 – Can’t tell how long the trip is because of missing month, but return date is January 1, and interviewer includes comment in BC variable, “2 mo.! Don’t think so.” So, this trip should probably be removed.
- CSID 10805 (RECID 88-89). Shows two 1 day, 1 beach trips that occur on the same day. This seems to imply WT should be recoded so that these are both part of the same 1 day, multi beach trip.
- CSID 19027 (PREID 742). All dates missing. Interviewer note in BC column indicates part of same trip as PREID 741.

Cases where start date is missing in addition to month

- 35475, 1197 (format is CSID, PREID)
- 32223, 1152
- 32223, 1153
- 32223, 1155
- 32223, 1156
- 31507, 1059
- 31507, 1060

2. Parking cost in dollars.

Q4I	Freq.	Percent	Cum.
0	331	67.55	67.55
1	42	8.57	76.12
2	27	5.51	81.63
3	11	2.24	83.88
4	8	1.63	85.51
5	27	5.51	91.02
6	25	5.10	96.12
7	12	2.45	98.57
8	1	0.20	98.78
10	2	0.41	99.18
13	1	0.20	99.39
88888	3	0.61	100.00
Total	490	100.00	

Nothing mysterious here.

3. Number of people accompanying respondent to the beach.

Q4J	Freq.	Percent	Cum.
0	230	35.17	35.17
1	189	28.90	64.07
2	58	8.87	72.94
3	64	9.79	82.72
4	40	6.12	88.84
5	23	3.52	92.35

6		9	1.38	93.73
7		4	0.61	94.34
8		1	0.15	94.50
9		2	0.31	94.80
11		1	0.15	94.95
13		1	0.15	95.11
15		1	0.15	95.26
19		1	0.15	95.41
25		1	0.15	95.57
35		1	0.15	95.72
40		2	0.31	96.02
88888		26	3.98	100.00

Total		654	100.00	

Note. Inspection of values 19-40 does not give reason to reject these. They are generally supported by other data, especially corresponding large number of children on the trip (number of people under 18 (Q4K)).

MEMORANDUM

TO: Beach Project Funders
FROM: The Research Team
RE: Recent Activities To Check Data
DATE: 29 May 2001

CONTENTS

1. INTRODUCTION
 2. REPORTS ON PRELIMINARY DATA REVIEW AND ANALYSIS
 3. EMAIL CORRESPONDENCE WITH CHICO REGARDING DATA
- APPENDIX 1 - Supplement to Report #1
APPENDIX 2 – Supplement to Report # 4
APPENDIX 3 – Supplement to Report # 5

1. INTRODUCTION

This memorandum reports on the checking of the survey data received from Chico Survey Research Center over the past few months. The checking was performed by Chris Busch with guidance from Michael Hanemann, David Layton, and Linwood Pendleton.

The memorandum consists of two sections:

1. A compilation of reports to the other researchers by Chris Busch summarizing the results of his work to analyze the survey data and check for problems and inconsistencies.
2. A reproduction of the email dialogue between the research team and Chico that took place as the research team sent inquiries to Chico regarding issues arising during Chris Busch's analysis of the data.

In addition, there are three appendices containing data that supplement reports #1, #4 and # 5.

We used two techniques to review and check the data. First, we prepared and analyzed summary statistics and descriptive statistics on beach usage and beach activities to see if there were any results that appeared implausible or unreasonable. Secondly, we checked to see that the flow of respondents through the different branches of the survey is logically consistent. The many skip patterns in the survey mean that different members of the panel may face different questions in any given diary survey. For example, the diary survey asks respondents how far they went in the water only if they said "yes" to a question about whether or not they went in the water. Thus, the number of answers to the question about how far the respondent went in the water should equal the number who said "yes, they went the water" in answering the previous question.

Through this process we have found a number of errors, and the issues we have raised led to some important clarifications about how the way the raw responses from the CATI system was transformed into the data spreadsheets that Chico sent us.

Because this transformation of the raw responses to the CATI system into data spreadsheets is a key step in the preparation of the data that we will use in our analysis, we are asking Chico State Survey Research Center to provide a report detailing how they did this.

2. REPORTS ON PRELIMINARY DATA REVIEW AND ANALYSIS

The five reports presented below were first circulated for discussion by the research team. Detailed descriptive and summary statistics that accompanied the reports are included as appendices to this memorandum.

CHRIS BUSCH'S REPORT #1 – 24 FEBRUARY 2001

(Note: only Diary 2, 3, and 4 had been delivered at the time of writing.)

Introductory Comments

I've done some work to assess the quality of the diary datasets delivered thus far. This very preliminary investigation suggests that these datasets are in fairly good shape. Each of the three datasets is in the same format and the codebooks are intelligible.

One item of concern for the research team is that the data is in only one of two forms we requested at the 17 November 2000 meeting. At that time, Allen agreed to deliver the data (1) as a matrix where each row is an individual respondent and (2) as a matrix where each row is a trip to the beach. As delivered, the data is only in the second of these formats.

Here is an overview table that summarizes beach-going behavior by month.

Overview Table

	Feb.	March	April	May	June	July
Percentage of panel that went to the beach	27.6%	36.5%	26.5%	35.0%	35.3%	46.4%
Mean number of trips	2.85	2.89	3.20	2.49	2.80	2.88
Panel size	633	633	505	505	530	530

These results are broadly in line with expectations. Trips to the beach generally increase as spring turns to summer and weather improves. The spike in the percentage of the panel going to the beach in March is somewhat strange, but may be partly attributable to springtime school holidays. The relatively large mean number of trips in April reflects at least in part the fact that there were the fewest number of single trips to the beach reported in that month. Note that the mean number of trips would be lower for February and March if one particularly avid beachgoer who went to the beach everyday during both months had been excluded from the calculation.

Initial Questions

Based on my preliminary review, I have a few quibbles and questions that I list below.

1. Inconsistent coding of yes/no responses. Sometimes “no” is coded <5> and sometimes it is coded <3>. Consistent coding would be useful.
2. Why do we have separate response codes for “Not Sure” and “Don’t Know”?
3. Q3 asks the respondent if s/he went to the beach in March. Possible response are: 1—yes, 3—no, and 5—man. What does “man” mean here?

Descriptive and summary statistics for the three diaries delivered at the time of writing are listed as **Appendix 1** (see page 25 below).

CHRIS BUSCH’S REPORT #2 – 24 MARCH 2001

(This report followed delivery of a new collection of datasets for Diaries 2,3, and 4.)

New Datasets for Diary Surveys 2,3, and 4

The new "final clean datasets" that Allen has provided do not differ from the first collection of datasets, at least with respect to the variables I reported on in my preliminary review of the first collection of datasets. The new codebooks are precisely the same as the first set, right down to the same number of characters contained in the documents.

Thinking About A Reason for the Spike in Beach Usage in March

Upon reflection, I wonder if we should be concerned about the fact that the percentage of the panel visiting the beach at least once in March was second highest over the February to July time period? Only July saw a higher percentage of people going to the beach. More people going to the beach during this relatively cool and rainy month seems counterintuitive.

[For these six months the percentages of the panel that visited the beach at least once are: Feb(27.6%), March (36.5%), April (26.5%), May (35%), June (35.3%), July (46.4%).]

In my initial memo, I suggested that spring break may have something to do with this. But while spring break for many colleges falls in March, for many public schools it falls in April. In Paul Ruud's attendance model he included a spring break variable based on the week before Easter (i.e. in April), and this variable was statistically significant for each of the 6 beaches in both the log-linear and non-linear model formulations.

To look into this further, I collected weather info (using Newport Beach as a proxy for regional weather) to see if there were unusually high temperatures in March last year. This was not the case. In fact, March was cooler than February. A brief synopsis of this weather data was initially included, but is now excluded in favor of the fuller picture of weather data provided in the table that follows.

(Linwood's Comment: You should check to see whether last March was indeed rainy. I don't think it was. Also, March attendance is not significantly different than May and June. I think Spring break is in March for most of LA.)

CHRIS BUSCH'S REPORT #3 – 26 MARCH 2001

In response to Linwood's comment, I provide some greater detail on weather during the course of diary surveys 2-4 in this message. I switched to looking at LAX since it is more central and the month of April was missing for Newport.

The table below provides data on LAX on number of days with at least 0.1 inches of rain, mean daily temperature, mean daily maximum temperature, and mean daily minimum temperature.

Since the temperature measures track each other for the most part, below I'll just list mean daily maximum temperature (in degrees Fahrenheit) and number of days with rain.

Snap Shot of Weather Over Diaries 2-4 (Feb. to July, 2000)

Month	Mean Daily Temperature*	Mean Maximum Temperature*	Mean Minimum Temperature*	Number of Days at Least 0.1 Inches of Rain
Feb.	57.5	64.1	51.4	9
March	57.9	64.5	51.3	3 (3 additional days with trace amounts of rain; no rain after March 9 th)
April	61.0	67.1	54.8	2
May	64.3	70.3	58.2	0
June	68.0	73.8	62.1	0
July	68.6	74.5	62.8	0

*Temperatures in degrees Fahrenheit.

I looked through the day-by-day data for March, and the average figures do not mask a few especially hot days that might have really drawn people to the beach.

(You can view the complete LAX data for February through September at, <http://ftp.ncdc.noaa.gov/pub/upload/581451119751dat.html>)

So, it seems March was not particularly warm, but much less rainy than February. I still find the attendance pattern a bit strange. Perhaps the spike in March attendance can be explained by spring break (as Linwood says he thinks the phenomena occurs mainly during that month) as well as an increase in beach visitation due to a release in pent up demand caused by rains in February and early March? In addition to the three days with significant rain, there were trace amounts of rain early in the month of March as well, but this was all done with by March 9.

(Linwood's comment: I think pent up demand is a good explanation. Also, I think our daily beach users (e.g. runners, etc.) are not sensitive to weather and these beach users have disproportionate weight in the "total beach visit calculations." I'm not particularly worried about this.)

CHRIS BUSCH'S REPORT #4 – 27 MARCH 2001

This report surveys diary data pertaining to beaches visited and activities undertaken at the beach.

1. Most popular beaches

I've tallied beach trips in order to get a picture of the most popular beaches.

Diary 2

1. Newport

2. Santa Monica
3. Seal
4. Huntington City
5. Long Beach

Diary 3

1. Santa Monica
2. Huntington City
3. Newport
4. Manhattan
5. Hermosa

Diary 4

1. Newport
2. Santa Monica
3. Hermosa
4. Huntington State
5. Manhattan (tie)
5. Bolsa Chica (tie)

As Linwood observes, these lists of most popular beaches will reflect the preferences of frequent beach users. For example, Bolsa Chica, listed as one of the favorite beaches for respondents to Diary 4, is not an exceptionally popular beach in terms of total visitation. It was propelled to the top five listing by four respondents who went 17, 16, 9, and 7 times respectively. Future analysis could assess popularity where each respondents' choices have equal weight rather than weighing each beach trip equally. That is, each respondent's favorite beaches could be used to judge overall popularity thereby reducing the influence of repeated trips to the same beach by particular respondents.

2. Beach Activities

An initial question (Q4C) and three follow ups (SUM1, SUM2, and SUM3) were asked about beach activities undertaken for each trip to the beach. I've combined the variables defined by these four questions in order to get an overall view of activities. The separate responses to each question are also tabulated in the sections devoted to each separate diary survey below.

Five most popular activities (most popular listed first)

Diary 2

1. Walking
2. Biking
3. Shopping/dining
4. Jogging
5. People watching

Diary 3

1. Walking
2. Other
3. Biking
4. Swimming
5. Shopping/dinning

Diary 4

1. Walking
2. Sunbathing
3. Swimming
4. Biking
5. Other

Again, these lists of most popular activities will reflect the preferences of frequent beach users. A different approach would be to look at favorite activities by respondent instead of considering activities associated with each beach trip.

Given the approach taken, it makes sense activities like walking and biking, activities that may be part of an exercise routine, are prominent. Of note is the appearance of swimming as the 4th most common activity in Diary 3 and its rise to 3rd most common activity in Diary 4. Also note that sunbathing rocketed to the 2nd favorite activity in Diary 4 during the warm months of June and July. All together, this suggests data on beach activities makes sense.

3. Results On Water-Based Recreation

As would be expected, water-based recreation increases in popularity from Diary 2 to Diary 4 with a particularly notably increase in full water immersion in Diary 4.

Water-Based Recreation Results

	Diary 2	Diary 3	Diary 4
Q4D. Went in Water? # Saying Yes/Total Responses (%Yes)	192/914 (21.0%)	96/454 (21.1%)	179/689 (26.0%)
Q4E. How Deep? # Saying "Got Head Wet"/Total Going in Water (% That Got Head Wet of Those Going in Water)	33/204* (16.2%)	40/153* (26.1%)	83/243* (34.2%)

*It is strange that for each of the diary surveys the number of responses to Q4E exceeds the number of "yes" responses to Q4D. The respondent is only asked Q4E if they answer "yes" to Q4D. There is a logical inconsistency here. More people indicate that they went in the water in Q4E than say "yes" when asked in Q4D if they went in the water at all.

4. Questions

The main item for concern is the apparent logical inconsistency in the responses to Q4D and Q4E as discussed in the preceding section.

One small matter is the fact that Q4E in Diary 4 has five responses coded “5” but this code does not correspond to any given for the question in the codebook. Two potential response codes—“Don’t Know” and “Refused”—are listed for the question, but no such responses appear in the dataset. Perhaps “Don’t Know” and “Refused” responses were subsumed under the code 5 category?

In each case, the codebook lists a potential response of “5 = man” for Q4D which doesn’t make sense. But this response is never actually given, so it seems like nothing to worry about.

5. Detailed Results

Appendix 2 (see page 33 below) gives detailed results for each of the three diaries available at the time of the report. Beach codes and activity codes can be found in each diary codebook and are reproduced at the end of the appendix.

CHRIS BUSCH'S REPORT #5 – 11 APRIL 2001

This report first lists some potential problems that have been identified in the course of the latest work on activities. Next it discusses findings on respondent activities controlling for the frequency of use—a primary activity is designated for each respondent—for each of the six waves.

Some potential problems with the data

- Diary 6 dataset is missing variables: Q2, Q2A,Q3,Q3A. There is a variable called FXQ2 in the dataset that is not listed in the codebook.
- Diary 6 dataset includes a variable, “TRIP”, that is not defined in the codebook. It seems that this assigns the same trip number for each beach visited on a particular beach trip. (Since multiple beaches can be visited on a single trip, this is an issue.) It would be useful to have such a variable for the other diary survey datasets. At present, there isn't one.
- There is a problem with the definition of the GOAL variable (which is the total number of trips for the two-month time period). The GOAL variable reads zero when in a number of cases there have been multiple trips to the beach. GOAL seems to be calculated as the sum of Q2A and Q3A, the number of trips in the first and second months of the survey time period respectively. The problem seems to be that when Q2A and Q3A are missing, GOAL is defined as zero, rather than as missing. I tried to note the cases where this happened for each diary survey in my line-by-line walk through the 6 datasets. I have this information jotted down (that is, other case id numbers where GOAL equals zero and there were some beach trips taken.)
- The variable WT identifies the type of beach trip—single or multi-day, single or multi-beach. It also has a code, “cruise of many beaches,” that seems unnecessary in light of the other four more specific and descriptive choices. This “cruise of many beaches” code does show up in the data, for example 10 times in diary 3 and 91 times in diary 4. Are these single day or multi-day trips?

Findings on activities controlling for frequency of use

Note that the next section describes how the primary activity for each respondent was determined. Full tabulated results for the primary activities of respondents in each diary survey are at the end of the memo. The differences with results when not controlling for frequency of use are not huge. Many of the top activities are the same. The five most popular activities are (top six when one of the top five is other):

Diary 1

1. Walking

2. Biking
3. Other
4. Shopping/Dining
5. Enjoying the view
6. Surfing

Diary 2

1. Walking
2. Biking
3. Picnicking
4. Shopping/Dinning
5. Enjoying the view

Diary 3

1. Walking
2. Swimming
3. Sunbathing
4. Biking
5. Other
6. Boogie Boarding/Body Surfing/Skimboarding

Diary 4

1. Walking
2. Swimming
3. Sunbathing
4. Activities with children
5. Shopping/dining (tie)
5. Enjoying the view (tie)
5. Biking

Diary 5

1. Walking
2. Sunbathing
3. Other
4. Biking
5. Swimming
6. Surfing

Diary 6

1. Walking
2. Shopping/Dining
3. Surfing
4. Other
5. Jogging
6. Biking

Focus on Activities of Interest

Diary #	Surfing Rank	Swimming Rank
1	6	21 (tie)
2	8 (tie)	11 (tie)
3	7 (tie)	2
4	8	2
5	6	5
6	3	7

These results seem reasonable overall. Surfing is a relatively more important activity in the winter when surfers are drawn to good winter waves and others are less likely to go to the beach. Swimming grows in popularity during the warmer months. The stronger showing for swimming during diary 3 (April-May) than during diary 5 (August-September) would not be expected based on temperature alone. The large number of people swimming during diary 3 may reflect the desire of people to get out and enjoy the weather having just emerged from cooler months.

How the primary activity was determined

In many cases, this was easy. For example, a person goes to the beach eight times and lists only walking as their activity every time. That person’s main activity is walking. In some cases there were ties that had to be broken. In general, the first activity listed is given most weight. There is no preference indicated by the formulation of the questions, but it is natural to think that the first response to the question, “what activities?,” is the main activity.

So suppose, respondent X listed their activities for two beach trips as:

walking enjoying the view
walking enjoying the view

When there is a tie such as this, I give greater weight to activities listed first, and so conclude this person’s main beach activity is walking.

If there is a tie with “other,” I chose the activity that is named, since “other” provides no information.

In some cases there is no clear indication of how to break a tie, such as a respondent Y that lists their activities for two beach trips as:

walking enjoying the view
enjoying the view walking

In such situations, I alternate between picking the first and then the second activity listed.

Appendix 3 (see page 57 below) gives detailed tabulated results on activities.

2. EMAIL CORRESPONDENCE WITH CHICO REGARDING DATA

MESSAGE FROM RESEARCH TEAM 12 APRIL 2001

Dear Allen,

Chris Busch has been going through the data hand has identified the following potential problems:

- Diary 6 dataset is missing variables: Q2, Q2A,Q3,Q3A. There is a variable called FXQ2 in the dataset that is not listed in the codebook.
- Diary 6 dataset includes a variable, “TRIP”, that is not defined in the codebook. It seems that this assigns the same trip number for each beach visited on a particular beach trip. (Since multiple beaches can be visited on a single trip, this is an issue.) It would be useful to have such a variable for the other diary survey datasets. At present, there isn’t one.
- There is a problem with the definition of the GOAL variable (which is the total number of trips for the two-month time period). The GOAL variable reads zero when in a number of cases there have been multiple trips to the beach. GOAL seems to be calculated as the sum of Q2A and Q3A, the number of trips in the first and second months of the survey time period respectively. The problem seems to be that when Q2A and Q3A are missing, GOAL is defined as zero, rather than as missing. I tried to note the cases where this happened for each diary survey in my line-by-line walk through the 6 datasets. I have this information jotted down (that is, other case id numbers where GOAL equals zero and there were some beach trips taken.)
- The variable WT identifies the type of beach trip: single or multi-day, single or multi-beach. It also has a code, “cruise of many beaches,” that seems unnecessary in light of the other four more specific and descriptive choices. This “cruise of many beaches” code does show up in the data, for example 10 times in diary 3 and 91 times in diary 4. Are these single day or multi-day trips?

MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER 17 APRIL 2001

(From David Philhour, his responses to questions raised above are in *italics*.)

>Michael --

>

>Allen asked me to look into these problems you were having with the data.

>

> *Diary 6 dataset is missing variables: Q2, Q2A,Q3,Q3A.

>There is a variable called FXQ2 in the dataset that is not listed in the
>codebook.

> ** *The Excel spreadsheet that was sent to you had several
>hidden columns and these variables were among those hidden columns
>(along with some extraneous variables not in codebook E.G. fxq2) . I am
>sending a new spreadsheet with those variables eliminated. Also some fills
>for Q6BR and Q6DYR were discovered that needed to be undertaken.*

>

> *Diary 6 dataset includes a variable, "TRIP", that is not
>defined in the codebook. It seems that this assigns the same trip number
>for each beach visited on a particular beach trip. (Since multiple beaches
>can be visited on a single trip, this is an issue.) It would be useful to
>have such a variable for the other diary survey datasets. At present, there
>isn't one.

>

> ** *The variable "TRIP" was part of the original rostering
>program in CASES. Given that we have re-defined each record as being the
>activities at a particular beach within a particular trip, the trip
>indicator no longer gives an accurate count.*

>

> *There is a problem with the definition of the GOAL variable
>(which is the total number of trips for the two-month time period). The
>GOAL variable reads zero when in a number of cases there have been multiple
>trips to the beach. GOAL seems to be calculated as the sum of Q2A and Q3A,
>the number of trips in the first and second months of the survey time period
>respectively. The problem seems to be that when Q2A and Q3A are missing,
>GOAL is defined as zero, rather than as missing. I tried to note the cases
>where this happened for each diary survey in my line-by-line walk through
>the 6 datasets. I have this information jotted down (that is, other case id
>numbers where GOAL equals zero and there were some beach trips taken.)

> *** *I would argue that GOAL is properly defined. If the
>respondent indicated that they went to no beaches, then they were not asked
>Q2 or Q3 and Goal=0 indicates that the respondent went to 0 beaches during
>that two-month period.*

>

> *The variable WT identifies the type of beach tripsingle or
>multi-day, single or multi-beach. It also has a code, "cruise of many
>beaches," that seems unnecessary in light of the other four more specific
>and descriptive choices. This "cruise of many beaches" code does show up in
>the data, for example 10 times in diary 3 and 91 times in diary 4. Are
>these single day or multi-day trips?

> **** the variable WT (Type of trip) has the value 5 = cruise of
>several beaches (via bicycle or rollerskates or boat) ALL ON A SINGLE
>DAY, This clarification is presented in the Production Report currently
>being created by Allen.

MESSAGE FROM RESEARCH TEAM
21 APRIL 2001

Dear Allen,

Following my phone discussion yesterday, variables that we'd like from the CATI survey management program are:

Date of interview (to check telescoping, survey error etc)
Contact info -- name, address, phone number (so we can mail a summary of findings, etc)

In addition, do you info about whether the person had used his calendar when reporting on trips?

Do you have an indication from the interviewer whether the person seemed confused, or data had to be altered manually, or there might have been some other problem with the interview (which could be useful in explaining discrepancies and inconsistencies that might emerge during the statistical analysis)?

Are there any other variables from the CATI program that might be similarly useful?

Many thanks

Michael

MESSAGE FROM RESEARCH TEAM
30 APRIL 2001

Hi Allen,

Thanks to you and David for addressing the earlier questions I had raised about the datasets.

Michael has asked me to follow up with respect to the GOAL variable.

David's response indicates the he believes the variable is properly defined (recall GOAL is the total number of trips for the wave), but I believe there is a problem with the variable. There are a number of cases in each dataset where the GOAL variable shows a zero (0) value but the data otherwise indicate some beach trips have been taken (i.e. beaches visited or activities undertaken at the beach are listed for the respondent).

Here are some specific instances where this seems to be the case:

Diary 1

CSID 15096

CSID 16198

CSID 17327

Diary 2

CSID 11499

CSID 12047

CSID 14304

CSID 15949

CSID 17293

CSID 19761

Diary 3

CSID 15047

Diary 4

CSID 12888

CSID 14685

Diary 5

CSID 15251

CSID 15949

CSID 16314

CSID 30431

Diary 6

CSID 14844

Would you please look into this?

Thank you,

Chris Busch

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
30 APRIL 2001**

Hi Chris,

From my initial look I would say we are talking a computer anomaly. The data make it clear that these folks have gone to the beach. The data indicate that they went in either December or January, but not how many times. It may have to do with changing the number of times they remembered going. Either way, goal was originally a control variable and not a variable that I intended to be analyzed. We could find the anomalies and make them jive with the data if you would like to use the variable for analysis.

Allen

**INTERNAL RESEARCH TEAM MEMO
30 APRIL 2001**

Michael,

Allen's explanation of a "computer anomaly" isn't very comforting. I think this incident should perhaps encourage us to be particularly vigilant in looking over the data. Still, it doesn't seem that we should obsess about this GOAL variable. There are separate questions asking the number of times the respondent went to the beach for each of the two months in the wave. So we can always construct the goal variable if we need it. Rather than pushing for action for Allen on this front, perhaps we wait until a more pressing data problem emerges. Once I've finished the paper I'm working on that's due

tomorrow, I'll take a look at the datasets to look for similar instances where we might be able to find some logical inconsistencies that we need explained. I have one candidate in mind already. As I recall and mentioned once previously, there appears to be a problem with one (or more? at least one, I'll have to check) of the diary surveys. The problem is that the number of people saying they went in the water is exceeded by the number of respondents to the following question asking how far they went in the water, but only the people who answered yes to the preceding question should have been asked the follow up. (These are questions Q4D--Go in the water? and Q4E--How far?)

~Chris

MESSAGE FROM RESEARCH TEAM
3 MAY 2001

Dear Allen,

In reviewing the data Chris has come up with some questions for you.

Many thanks

Michael

FROM: Chris

Why are there inconsistencies in branching structure results?

I think we need an explanation for apparent inconsistencies in branching structure results, by which I mean that the flows of respondents to various question based on their responses to previous ones don't seem to make sense. I see two issues here, one that is quite common and the other less so.

- **Missing responses.** In almost every case I've looked at, when those who answer "yes" to a question are supposed to be asked a follow up question, there are fewer respondents to the follow up question than there were "yes" responses. The only exception to this is when there are too many responses to the follow up question. Some examples are detailed below and other examples are listed in the attached memo.
- **Excess responses.** More alarming is the fact that sometimes there are more responses to follow up questions than there should be in some cases. The one example I've found of this pertains to the question q4d ("did you go into the water when you were there?") and q4e (if yes to q4d, "which of the following best describes how much you went into the water?").

The problem of excess responses.

If a person answers yes to q4d, they are directed to answer q4e. So total responses to q4e (including “don’t know” or “refused”) should equal “yes” responses to q4d. This is not true for any of the diary surveys. Here is the breakdown for each diary survey.

Diary #: “Yes” responses to q4d/Total responses to q4e.

Diary 1: 163/175

Diary 2: 192/205

Diary 3: 96/153

Diary 4: 179/248

Diary 5: 159/249

Diary 6: 49/78

How can this be?

Full tabulated results on these questions can be found in the attached memo.

Some examples of missing responses from Diary 4.

“1” responses to q2 (187) are supposed to go q2a, but there are only 182 responses to q2a.

“1” responses to q3 (246) are supposed to go q3a, but there are only 241 responses to q3a.

Some more examples of missing responses can be found in the attached memo.

How to understand the branching structure regarding reporting details of trips?

In the text of the diary survey (I’m looking at a version dated February 2000), q4, which is not really a question of the respondent but is rather a point where previous answers are evaluated, determines whether the respondent is asked about every trip (if they took five or less trips) or not (if they took more than 5 trips). The diary survey text indicates that a variable TOTAL will be created based on the separate totals for each of the two months in the wave, and that this variable will be used to route the respondent. Those with TOTAL greater than 5 are sent to a series of questions q11a and etc. I don’t see any q11-type questions in the datasets. Have these been re-labeled, perhaps as q4a and etc. (which are the same questions for people who took 5 or less beach trips)? Also, this TOTAL variable has been renamed “GOAL” yes?

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
4 MAY 2001**

>HI Michael,
>
>I am forwarding this to Chris Sweeney and David Philhour for response. I
>will get back to you early next week with clarification.
>
>Allen
>
>Chris and David,
>
>Could you look at these questions and tell me which ones fit into
>something you can answer and which you need me to do further
>investigation?
>
>David,
>
>How are you coming on your section of the production report?
>
>Allen

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
7 MAY 2001**

>Date: Mon, 7 May 2001 11:35:01 -0700
>X-Mailer: Internet Mail Service (5.5.2650.21)
>
>HI Michael,
>
>I agree with Chris. Most of this variation was built into the program
>because of the variation in the ways a respondent went to the beach. They do
>not all get the same questions. The other source of this could be computer
>error if the respondent changed their mind. I will have Chris and David
>Philhour go over this in greater depth and David is going to give me a
>section for the production report on how the data files were reconstructed
>into the matrix you have. That may explain a lot.
>
>Allen
>
>-----Original Message-----
>From: Sweeney, Christian

>Sent: Monday, May 07, 2001 10:13 AM

>To: Lunde, Allen

>Subject: RE: Data issues

>

>Hi Allen,

>

>In looking over the problems I feel that most of the problems with questions
>having more/less answers than the question before it, may have to deal with
>how the data was merged for people who had multi-day trips vs. people who
>didn't. On the multi-day trips some questions were removed to make the
>survey shorter, correct? This may be part of what causes this problem.

>

>The other cause of some questions not having an equal number of answers as
>the follow up questions, could be caused by interviewers jumping back in the
>questionnaire to change an answer based on the R changing their mind. This
>could change the execution path of the program and therefore the first
>question could have an answer, however, the follow up question will not.
>For example if the R said they went to the beach in Feb. and we answer yes
>to q2 and then we get to q2a and they say they didn't go to the beach at all
>during Feb. or Mar., the interviewer could jump back to the question asking
>whether the R went at all, and if they answer NO, q2a will be skipped and the
>program will jump to the module portion of the questionnaire, skipping the
>diary portion. Although this shouldn't account for all of the differences
>between a question and the follow up question it should account for some of
>the variation.

>

>David may have a better idea of what could cause there to be more/less
>responses to a question that should have an equal number of responses, based
>on the merging of the data. We may need to have a breakdown of examples we
>need to check. We did however talk about this when we were going over the
>data (how some questions have answers that should lead to the follow up
>question not having an answer) and the causes of this were explained or
>dismissed as normal variation, correct?

>

>We should have the trace files for each case and can look into specific
>cases to help explain the cause, however, this would be very time consuming.
>Let me know what becomes of this and if we need to do more research.

>

>Chris S.

>

MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
15 MAY 2001

Hi Michael,

Attached is the latest draft of the production report which includes David's section on the matrix reconstruction. Also attached are two excel files that support the model.

Allen

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
16 MAY 2001**

Hi Michael & Chris (Busch)

I have solved the problem with q4d and q4e responses. There was a skip problem in the program. If q4c-sum3 (or their equivalents) indicated that the respondent participated in an activity that naturally would have taken them into the water (i.e. swimming) they were not supposed to be asked q4d or q4e. However in some cases there was a programming error that took the questionnaire to q4e when it should have gone to q4f. Because this problem only occurred in a few instances, we did not discover it in our data checks. It is easily cleaned up with a macro program. I'll get to that today.

As for problems with q2/q2a and q3/q3a my only theory is that problems occurred when the respondent changed their mind and required the interviewer to attempt to leave the normal skip pattern. The actual dates and lines of data should indicate the number of times they went in a given month. The program did not rely on the respondent to get the number of trips right on the first try, but allowed the respondent to change their mind if they realized that they went to the beach more or less than they originally remembered. We could add a variable or two based on the actual trip data that would be more accurate than q2a or q3a, which only indicate what the respondent thought when first asked.

If there are other issues, please let us know. The more specific, the quicker it is for us to resolve. Allen

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
17 MAY 2001**

>Attached are the data sets with q4d/q4e cleaned up. The problem that
>occurred was program based and easy to locate. I wrote a macro that fixed
>the problem in most cases. There were 4 cases in the diary 1 data set and 1
>case in the diary 2 data set that they macro did not correct. I believe this
>was a recoding error. To keep the data clean I have recoded q4d as don't
>know in these 5 cases and made q4e blank.

>
>Allen

MESSAGE FROM RESEARCH TEAM
18 MAY 2001

(Message forwarded to Allen)

Michael,

Allen's response on q4d/q4e seems excellent.

Here are some other thoughts.

Allen writes "As for problems with q2/q2a and q3/q3a my only theory is that problems occurred when the respondent changed their mind and required the interviewer to attempt to leave the normal skip pattern."

I think by this he means the respondent initially said, "yes", trips were taken during month X and then changed their mind when asked how many trips for the month. But, if this is the case, shouldn't the interviewer have gone back to change the initial response to "no"? I don't know if we need this to be fixed, (incorrect "yes" responses changed to "no"), but it seems we should take Allen up on his offer to put together a variable with an accurate count of the number of trips.

Allen asks us to raise specific issues. One other specific issue of concern raised in the memo I wrote is the the apparent loss of respondents between q4h (what transportation taken to the beach) and q4i (for motor vehicle drivers, how much was parking). Here are the numbers on this, where the first number is the # of 1 responses (=took motor vehicle to beach) to q4h and the second number is total responses to q4i.

Diary 1: 745/708

Diary 2: 726/675

Diary 3: 584/553

Diary 4: 835/781

Diary 5: 919/870

Diary 6: 526/490

My reading of the survey structure suggests that these two numbers should always be equal, but this isn't the case.

I will return to the datasets to look for other specific questions that we may want to raise.

Also, I think we need David Philhour to write some text to accompany and explain the excel files on matrix construction (Beaches Plan, Wave Transformation Matrix). These excel documents are not self-explanatory, at least to me. The production report doesn't provide sufficient detail to substitute for such an explanation. Didn't we ask for the raw data too, in case the constructed datasets need to be verified?

~Chris (Busch)

**MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
21 MAY 2001**

>Hi Folks,

>

>The q4h/q4i question is pretty easy to answer. The question about how much
>was spent on parking was not asked in every case where we asked a person how
>they got to the beach. An example of this is q5g, which asks the respondent
>what mode of transportation they used to get to the beach, but does not have
>a follow-up asking about how much was spent on parking. When the data matrix
>was reconstructed, the answer from q5g would have gone under q4h. Since
>there was no follow-up, there would be no answer for q4i and there would be
>more responses to q4h than q4i. Q4n not only represents the data from q4h,
>but every place where we asked about transportation to the beach.

>

>I did discover that a few of the data sets had some anomalous data in q4i.
>There were a few cases where q4h indicated that the respondent did not take
>motorised transportation to the beach, but did have a response to q4i. This
>was probably due to a skip problem. In every case except one where this
>occurred, the respondent answer either 0 or "don't know". I've cleaned this
>data up and it is attached.

>

>I will see to it that Chris has all the information he needs to understand
>the data matrix. Unfortunately David Layton and David Philhour set this up
>with the understanding that David Layton would be supervising the analysis
>of the data. Since that appears to have changed, it's clear that David
>Philhour needs to work with Chris to make sure he understands the matrix.

>

>Allen

**MESSAGE FROM RESEARCH TEAM
22 MAY 2001**

Allen,

Thanks for the clarification.

The survey text I am looking at (dated 2/12/00) has as q6g the question you call q5g, but point understood.

You write,

>Q4n not only represents the data from q4h,
>but every place where we asked about transportation to the beach. (You mean to write Q4h the first time)
>I will see to it that Chris has all the information he needs to understand
>the data matrix.

Indeed, the fact that q6g is missing from the codebook, which I hadn't noticed, isn't sufficient information to know that data from this question has been rolled in with q4h.

Furthermore, we will need explanation not just so that i can understand the datasets, but so that anyone can understand them. Since an important purpose of this study is to provide scientific justification for natural resource damage estimates that may be challenged in an adversarial setting (that is, in court), the datasets will need to be transparent. Any outside observer should be able to make sense of them.

So if a variable represents data other than that attributed to it in the survey text, that will be important. I suppose this ties in to our request for some description of the matrix construction process to accompany the excel spreadsheets David Philhour produced.

~Chris (Busch)

**INTERNAL CHICO STATE SURVEY RESEARCH CENTER EXCHANGE
22 MAY 2001**

> Thanks David,
>
> I did figure out the problem on q4h/q4i and explained it to Chris. The
> next big task will be to create a new variable that can be used as a more
> reliable substitution for q2a and q3a. We should talk about that.
>
> Allen
>
(Message to Allen Lunde from David Philhour, 22 May 2001)
>
> Allen --
> The q4h/q4i issue arises because Parking\$ infor was not collected in

> the series q5 (Multi-Beach) q6 or q7 questions relating to Multi-Day
> trips. Some of these Type 2,3,4 trips (Variable:WT) get Parking\$ picked up from
> Typical Trip data (x101, x102). I will write up a more detailed description of the
> construction of the Transformation matrices. Does Chris Busch have complete copies
> of the Interview schedule that contains the info on beach use variables q4mt through
> x501?
>
> Will try to get the explanation to you later today.
>
> david

MESSAGE FROM CHICO STATE SURVEY RESEARCH CENTER
23 MAY 2001

Chris (Busch),

See if this is helpful.

Allen

(Message to Allen Lunde from David Philhour, 22 May 2001)

> Here is a more detailed explanation of Beach Plan
> <<Explanation of BEACH PLAN.doc>>

APPENDIX 1 – Supplement to Report # 1

DIARY 2 DESCRIPTIVE AND SUMMARY STATISTICS

GOAL. Total trips for two month period (February and March)

GOAL	Freq.	Percent	Cum.
0	366	57.82	57.82
1	88	13.90	71.72
2	59	9.32	81.04
3	33	5.21	86.26
4	22	3.48	89.73
5	17	2.69	92.42
6	10	1.58	94.00
7	4	0.63	94.63
8	6	0.95	95.58
9	4	0.63	96.21
10	3	0.47	96.68
11	2	0.32	97.00
12	1	0.16	97.16
13	1	0.16	97.31
14	1	0.16	97.47
15	1	0.16	97.63
16	2	0.32	97.95
17	1	0.16	98.10
18	2	0.32	98.42
19	1	0.16	98.58
20	2	0.32	98.89
24	1	0.16	99.05
25	3	0.47	99.53
26	1	0.16	99.68
30	1	0.16	99.84
60	1	0.16	100.00
Total	633	100.00	

SUMMARY GOAL. Total trips for two month period (February and March)

Variable	Obs	Mean	Std. Dev.	Min	Max
goal	633	1.761453	4.395585	0	60

Q1AA. GO TO BEACH IN FEB OR MAR?

- <1> Yes
- <3> Not sure
- <5> No
- <88888> Don't know

Q1AA	Freq.	Percent	Cum.
1	279	44.08	44.08
3	2	0.32	44.39
5	351	55.45	99.84

88888		1	0.16	100.00

Total		633	100.00	

q2. Went to beach in Feb?

<1> Yes
 <3> Not sure
 <5> No
 <88888> Don't know

Q2		Freq.	Percent	Cum.

1		175	61.62	61.62
3		108	38.03	99.65
88888		1	0.35	100.00

Total		284	100.00	

q2a. Number of times went to the beach in Feb.

Q2A		Freq.	Percent	Cum.

1		73	43.98	43.98
2		40	24.10	68.07
3		16	9.64	77.71
4		17	10.24	87.95
5		4	2.41	90.36
6		1	0.60	90.96
7		2	1.20	92.17
8		1	0.60	92.77
9		2	1.20	93.98
10		2	1.20	95.18
12		3	1.81	96.99
13		2	1.20	98.19
25		1	0.60	98.80
29		1	0.60	99.40
88888		1	0.60	100.00

Total		166	100.00	

Summary q2a (Number of times went to the beach in Feb.)

Variable		Obs	Mean	Std. Dev.	Min	Max

q2a		165	2.854545	3.663023	1	29

q3. Went to beach in March?

<1> Yes
 <3> Not sure
 <5> No
 <88888> Don't know

Q3		Freq.	Percent	Cum.

1		231	81.34	81.34

3		53	18.66	100.00

Total		284	100.00	

q3a. Number of times went to the beach in Mar.

Q3A		Freq.	Percent	Cum.
1		95	42.60	42.60
2		58	26.01	68.61
3		25	11.21	79.82
4		14	6.28	86.10
5		7	3.14	89.24
6		7	3.14	92.38
7		2	0.90	93.27
8		2	0.90	94.17
9		1	0.45	94.62
10		3	1.35	95.96
11		1	0.45	96.41
12		1	0.45	96.86
13		1	0.45	97.31
14		1	0.45	97.76
16		1	0.45	98.21
17		2	0.90	99.10
25		1	0.45	99.55
31		1	0.45	100.00

Total		223	100.00	

summary q3a (Number of times went to the beach in Mar.)

Variable		Obs	Mean	Std. Dev.	Min	Max
q3a		223	2.887892	3.676194	1	31

DIARY 3 DESCRIPTIVE AND SUMMARY STATISTICS

GOAL = Total trips for April and May

GOAL		Freq.	Percent	Cum.
0		309	61.19	61.19
1		62	12.28	73.47
2		41	8.12	81.58
3		30	5.94	87.52
4		14	2.77	90.30
5		14	2.77	93.07
6		6	1.19	94.26
7		1	0.20	94.46
8		3	0.59	95.05
9		6	1.19	96.24
10		2	0.40	96.63
11		4	0.79	97.43

13		2	0.40	97.82
14		1	0.20	98.02
15		1	0.20	98.22
16		1	0.20	98.42
17		2	0.40	98.81
19		1	0.20	99.01
24		2	0.40	99.41
26		1	0.20	99.60
27		1	0.20	99.80
36		1	0.20	100.00

Total		505	100.00	

sum goal (total trips in April and May)

Variable		Obs	Mean	Std. Dev.	Min	Max
goal		505	1.59604	3.784259	0	36

q1aa. Went to beach in April or May?

<1> Yes
 <3> Not sure
 <5> No
 <88888> Don't know

Q1AA		Freq.	Percent	Cum.
1		220	43.56	43.56
3		2	0.40	43.96
5		281	55.64	99.60
88888		2	0.40	100.00

Total		505	100.00	

q2. Went to beach in April?

<1> Yes
 <3> No
 <88888> Don't know

Q2		Freq.	Percent	Cum.
1		134	59.29	59.29
3		88	38.94	98.23
88888		4	1.77	100.00

Total		226	100.00	

q2a. Number of times went to beach in April

Q2A		Freq.	Percent	Cum.
1		48	37.21	37.21
2		34	26.36	63.57
3		12	9.30	72.87
4		7	5.43	78.29
5		5	3.88	82.17
6		6	4.65	86.82

7		1	0.78	87.60
8		2	1.55	89.15
9		3	2.33	91.47
10		1	0.78	92.25
11		1	0.78	93.02
12		1	0.78	93.80
15		1	0.78	94.57
16		1	0.78	95.35
18		1	0.78	96.12
27		1	0.78	96.90
88888		4	3.10	100.00

Total		129	100.00	

sum q2a (Number of times went to beach in April)

Variable		Obs	Mean	Std. Dev.	Min	Max
q2a		125	3.2	3.793033	1	27

q3. Went to beach in May?

<1> Yes

<3> No

<88888> Don't know

Q3		Freq.	Percent	Cum.
1		177	78.67	78.67
3		47	20.89	99.56
88888		1	0.44	100.00

Total		225	100.00	

q3a. Number of times went to beach in May

Q3A		Freq.	Percent	Cum.
1		68	41.46	41.46
2		47	28.66	70.12
3		21	12.80	82.93
4		10	6.10	89.02
5		4	2.44	91.46
6		1	0.61	92.07
7		3	1.83	93.90
8		4	2.44	96.34
9		1	0.61	96.95
10		1	0.61	97.56
12		1	0.61	98.17
13		1	0.61	98.78
18		1	0.61	99.39
88888		1	0.61	100.00

Total		164	100.00	

sum q3a (Number of times went to beach in May)

Variable		Obs	Mean	Std. Dev.	Min	Max

DIARY 4 DESCRIPTIVE AND SUMMARY STATISTICS

GOAL. Total trips for two month period (June and July)

GOAL	Freq.	Percent	Cum.
0	250	47.17	47.17
1	83	15.66	62.83
2	58	10.94	73.77
3	36	6.79	80.57
4	23	4.34	84.91
5	17	3.21	88.11
6	13	2.45	90.57
7	9	1.70	92.26
8	5	0.94	93.21
9	5	0.94	94.15
10	5	0.94	95.09
11	3	0.57	95.66
12	8	1.51	97.17
13	1	0.19	97.36
14	3	0.57	97.92
17	2	0.38	98.30
18	3	0.57	98.87
22	2	0.38	99.25
27	1	0.19	99.43
28	1	0.19	99.62
30	1	0.19	99.81
35	1	0.19	100.00
Total	530	100.00	

sum goal (Total trips for two month period, June and July)

Variable	Obs	Mean	Std. Dev.	Min	Max
goal	530	2.256604	4.191787	0	35

q1aa. Went to beach in June or July?

- <1> Yes
- <3> Not sure
- <5> No
- <88888> Don't know

Q1AA	Freq.	Percent	Cum.
1	290	54.72	54.72
3	1	0.19	54.91
5	238	44.91	99.81
88888	1	0.19	100.00
Total	530	100.00	

q2. Went to beach in June?

<1> Yes

<3> No

<88888> Don't know

Q2	Freq.	Percent	Cum.
1	187	63.39	63.39
3	107	36.27	99.66
88888	1	0.34	100.00
Total	295	100.00	

q2a. Number of times went to beach in June

Q2A	Freq.	Percent	Cum.
1	70	38.46	38.46
2	43	23.63	62.09
3	29	15.93	78.02
4	9	4.95	82.97
5	9	4.95	87.91
6	4	2.20	90.11
7	5	2.75	92.86
8	6	3.30	96.15
10	2	1.10	97.25
11	1	0.55	97.80
12	1	0.55	98.35
15	1	0.55	98.90
18	1	0.55	99.45
88888	1	0.55	100.00
Total	182	100.00	

sum q2a (Number of times went to beach in June)

Variable	Obs	Mean	Std. Dev.	Min	Max
q2a	181	2.801105	2.623441	1	18

q3. Went to beach in July?

<1> Yes

<3> No

<88888> Don't know

Q3	Freq.	Percent	Cum.
1	246	83.96	83.96
3	45	15.36	99.32
88888	2	0.68	100.00
Total	293	100.00	

tab q3a. (Number of times went to beach in July)

Q3A	Freq.	Percent	Cum.
1	93	38.59	38.59

2		65	26.97	65.56
3		28	11.62	77.18
4		13	5.39	82.57
5		13	5.39	87.97
6		4	1.66	89.63
7		8	3.32	92.95
8		2	0.83	93.78
9		2	0.83	94.61
10		5	2.07	96.68
12		1	0.41	97.10
14		1	0.41	97.51
15		1	0.41	97.93
16		1	0.41	98.34
17		1	0.41	98.76
27		1	0.41	99.17
88888		2	0.83	100.00

Total		241	100.00	

sum q3a (Number of times went to beach in July)

Variable		Obs	Mean	Std. Dev.	Min	Max

q3a		239	2.882845	3.119955	1	27

APPENDIX 2 – supplement to Report # 4

DIARY 2 DESCRIPTIVE STATISTICS

Q4B	Beach Codes (Beaches visited by beach code.)		
Q4B	Freq.	Percent	Cum.
1	4	0.36	0.36
2	6	0.53	0.89
3	3	0.27	1.16
6	1	0.09	1.24
8	1	0.09	1.33
9	4	0.36	1.69
13	16	1.42	3.11
14	1	0.09	3.20
15	14	1.24	4.44
22	34	3.02	7.47
26	10	0.89	8.36
27	1	0.09	8.44
31	26	2.31	10.76
37	4	0.36	11.11
39	1	0.09	11.20
42	1	0.09	11.29
43	9	0.80	12.09
44	5	0.44	12.53
47	2	0.18	12.71
48	1	0.09	12.80
49	1	0.09	12.89
50	1	0.09	12.98
51	2	0.18	13.16
59	43	3.82	16.98
62	77	6.84	23.82
63	48	4.27	28.09
69	35	3.11	31.20
70	3	0.27	31.47
75	1	0.09	31.56
79	59	5.24	36.80
81	10	0.89	37.69
84	1	0.09	37.78
85	1	0.09	37.87
87	38	3.38	41.24
92	12	1.07	42.31
94	108	9.60	51.91
95	3	0.27	52.18
102	1	0.09	52.27
109	1	0.09	52.36
110	8	0.71	53.07
111	4	0.36	53.42
114	55	4.89	58.31
120	4	0.36	58.67
122	9	0.80	59.47
123	8	0.71	60.18

124	5	0.44	60.62
127	95	8.44	69.07
129	78	6.93	76.00
131	1	0.09	76.09
136	12	1.07	77.16
138	2	0.18	77.33
144	1	0.09	77.42
145	14	1.24	78.67
149	50	4.44	83.11
151	1	0.09	83.20
155	19	1.69	84.89
158	12	1.07	85.96
159	5	0.44	86.40
160	7	0.62	87.02
161	6	0.53	87.56
168	6	0.53	88.09
169	3	0.27	88.36
171	2	0.18	88.53
172	3	0.27	88.80
173	3	0.27	89.07
178	6	0.53	89.60
181	1	0.09	89.69
182	1	0.09	89.78
183	17	1.51	91.29
184	1	0.09	91.38
185	1	0.09	91.47
190	3	0.27	91.73
191	1	0.09	91.82
194	1	0.09	91.91
196	5	0.44	92.36
197	1	0.09	92.44
205	2	0.18	92.62
218	13	1.16	93.78
226	2	0.18	93.96
234	1	0.09	94.04
241	1	0.09	94.13
243	1	0.09	94.22
246	2	0.18	94.40
291	2	0.18	94.58
300	2	0.18	94.76
301	13	1.16	95.91
302	6	0.53	96.44
303	2	0.18	96.62
304	7	0.62	97.24
999	17	1.51	98.76
88888	10	0.89	99.64
99999	4	0.36	100.00
<hr/>			
Total	1125	100.00	

Diary 2 activities (Q4C, SUM1, SUM2, SUM3)

diary2	Freq.	Percent	Cum.
0	49	1.79	1.79
1	174	6.34	8.13
2	13	0.47	8.60
3	2	0.07	8.68
4	6	0.22	8.90
5	11	0.40	9.30
6	17	0.62	9.92
7	3	0.11	10.03
8	144	5.25	15.28
10	6	0.22	15.49
11	69	2.52	18.01
12	115	4.19	22.20
13	79	2.88	25.08
15	2	0.07	25.15
16	164	5.98	31.13
18	23	0.84	31.97
19	67	2.44	34.41
20	34	1.24	35.65
21	30	1.09	36.75
22	53	1.93	38.68
23	8	0.29	38.97
24	473	17.24	56.22
27	112	4.08	60.30
28	46	1.68	61.98
29	6	0.22	62.19
30	38	1.39	63.58
31	9	0.33	63.91
32	3	0.11	64.02
33	3	0.11	64.13
34	3	0.11	64.24
35	19	0.69	64.93
36	14	0.51	65.44
37	3	0.11	65.55
77	930	33.90	99.45
88888	11	0.40	99.85
99999	4	0.15	100.00
Total	2743	100.00	

Q4C First Activity
(Lists first beach activity named by respondent)

Q4C	Freq.	Percent	Cum.
0	37	3.48	3.48
1	154	14.47	17.95
2	3	0.28	18.23
3	1	0.09	18.33
4	5	0.47	18.80
5	6	0.56	19.36
6	13	1.22	20.58
8	92	8.65	29.23
10	4	0.38	29.61

11	39	3.67	33.27
12	20	1.88	35.15
13	53	4.98	40.13
16	64	6.02	46.15
18	3	0.28	46.43
19	29	2.73	49.15
20	29	2.73	51.88
21	14	1.32	53.20
22	16	1.50	54.70
23	4	0.38	55.08
24	363	34.12	89.19
27	27	2.54	91.73
28	19	1.79	93.52
30	31	2.91	96.43
31	5	0.47	96.90
32	3	0.28	97.18
33	3	0.28	97.46
34	3	0.28	97.74
35	6	0.56	98.31
36	3	0.28	98.59
88888	11	1.03	99.62
99999	4	0.38	100.00

Total	1064	100.00	

SUM1 Second Activity
(Lists second beach activity named by respondent.)

SUM1	Freq.	Percent	Cum.
0	10	0.99	0.99
1	15	1.48	2.47
2	10	0.99	3.46
3	1	0.10	3.56
4	1	0.10	3.66
5	4	0.40	4.05
6	3	0.30	4.35
7	2	0.20	4.55
8	30	2.96	7.51
10	1	0.10	7.61
11	26	2.57	10.18
12	42	4.15	14.33
13	2	0.20	14.53
15	2	0.20	14.72
16	73	7.21	21.94
18	15	1.48	23.42
19	19	1.88	25.30
20	4	0.40	25.69
21	8	0.79	26.48
22	27	2.67	29.15
23	3	0.30	29.45
24	64	6.32	35.77
27	71	7.02	42.79
28	21	2.08	44.86
29	4	0.40	45.26
30	1	0.10	45.36
31	1	0.10	45.45

35		8	0.79	46.25
36		9	0.89	47.13
37		3	0.30	47.43
77		532	52.57	100.00

Total		1012	100.00	

SUM2 Third Activity
(Lists third beach activity named by respondent.)

SUM2		Freq.	Percent	Cum.
0		1	0.21	0.21
1		5	1.04	1.25
5		1	0.21	1.46
6		1	0.21	1.67
7		1	0.21	1.88
8		1	0.21	2.09
10		1	0.21	2.30
11		4	0.84	3.13
12		33	6.89	10.02
13		23	4.80	14.82
16		24	5.01	19.83
19		17	3.55	23.38
21		7	1.46	24.84
22		10	2.09	26.93
23		1	0.21	27.14
24		26	5.43	32.57
27		9	1.88	34.45
28		5	1.04	35.49
29		2	0.42	35.91
30		6	1.25	37.16
31		3	0.63	37.79
35		5	1.04	38.83
36		2	0.42	39.25
77		291	60.75	100.00

Total		479	100.00	

SUM3 Fourth Activity
(Lists fourth beach activity named by respondent.)

SUM3		Freq.	Percent	Cum.
0		1	0.53	0.53
8		21	11.17	11.70
12		20	10.64	22.34
13		1	0.53	22.87
16		3	1.60	24.47
18		5	2.66	27.13
19		2	1.06	28.19
20		1	0.53	28.72
21		1	0.53	29.26
24		20	10.64	39.89
27		5	2.66	42.55
28		1	0.53	43.09
77		107	56.91	100.00

Total	188	100.00
-------	-----	--------

Q4D Go in Water?

- <1> Yes
- <3> No
- <5> Man

- <88888> Don't know
- <99999> Refused

Q4D	Freq.	Percent	Cum.
1	192	20.80	20.80
3	722	78.22	99.02
88888	9	0.98	100.00
Total	923	100.00	

Q4E How Deep?

- <1> Up to your ankles
- <2> Up to your waist
- <3> Up to your neck
- <4> You got your head wet

- <88888> Don't Know
- <99999> Refused

Q4E	Freq.	Percent	Cum.
1	121	59.02	59.02
2	24	11.71	70.73
3	26	12.68	83.41
4	33	16.10	99.51
88888	1	0.49	100.00
Total	205	100.00	

DIARY 3 DESCRIPTIVE STATISTICS

Q4B Beach Codes
(Beaches visited by beach code.)

Q4B	Freq.	Percent	Cum.
1	3	0.41	0.41
3	1	0.14	0.55
7	1	0.14	0.69
9	4	0.55	1.24
11	39	5.38	6.62
13	11	1.52	8.14
15	4	0.55	8.69
20	7	0.97	9.66
21	1	0.14	9.79
22	21	2.90	12.69
26	1	0.14	12.83
28	2	0.28	13.10
29	7	0.97	14.07
31	9	1.24	15.31
32	6	0.83	16.14
33	5	0.69	16.83
34	4	0.55	17.38
37	13	1.79	19.17
39	11	1.52	20.69
42	3	0.41	21.10
44	6	0.83	21.93
50	2	0.28	22.21
51	1	0.14	22.34
59	41	5.66	28.00
62	56	7.72	35.72
63	22	3.03	38.76
69	15	2.07	40.83
75	3	0.41	41.24
77	3	0.41	41.66
79	27	3.72	45.38
81	7	0.97	46.34
87	49	6.76	53.10
92	5	0.69	53.79
94	53	7.31	61.10
106	2	0.28	61.38
110	1	0.14	61.52
114	12	1.66	63.17
122	4	0.55	63.72
123	3	0.41	64.14
124	5	0.69	64.83
127	64	8.83	73.66
129	35	4.83	78.48
136	5	0.69	79.17
144	1	0.14	79.31
145	15	2.07	81.38
149	36	4.97	86.34
155	11	1.52	87.86

158	7	0.97	88.83
159	3	0.41	89.24
160	2	0.28	89.52
161	2	0.28	89.79
163	2	0.28	90.07
168	1	0.14	90.21
169	1	0.14	90.34
172	1	0.14	90.48
173	1	0.14	90.62
176	1	0.14	90.76
178	2	0.28	91.03
182	1	0.14	91.17
185	2	0.28	91.45
186	1	0.14	91.59
190	4	0.55	92.14
196	3	0.41	92.55
205	1	0.14	92.69
216	1	0.14	92.83
218	10	1.38	94.21
227	1	0.14	94.34
262	1	0.14	94.48
263	3	0.41	94.90
268	1	0.14	95.03
291	1	0.14	95.17
300	4	0.55	95.72
301	12	1.66	97.38
303	1	0.14	97.52
304	3	0.41	97.93
88888	6	0.83	98.76
99999	9	1.24	100.00

Total	725	100.00	

Diary 3 activities (Q4C, SUM1, SUM2, SUM3)

diary3	Freq.	Percent	Cum.
0	123	6.15	6.15
1	115	5.75	11.89
2	78	3.90	15.79
3	1	0.05	15.84
4	1	0.05	15.89
5	24	1.20	17.09
6	19	0.95	18.04
7	3	0.15	18.19
8	81	4.05	22.24
9	6	0.30	22.54
10	2	0.10	22.64
11	42	2.10	24.74
12	28	1.40	26.14
13	17	0.85	26.99
16	87	4.35	31.33
18	37	1.85	33.18
19	43	2.15	35.33
20	40	2.00	37.33
21	97	4.85	42.18
22	10	0.50	42.68

23	35	1.75	44.43
24	252	12.59	57.02
26	1	0.05	57.07
27	30	1.50	58.57
28	28	1.40	59.97
29	31	1.55	61.52
31	7	0.35	61.87
32	3	0.15	62.02
33	7	0.35	62.37
34	1	0.05	62.42
35	2	0.10	62.52
36	22	1.10	63.62
37	17	0.85	64.47
38	20	1.00	65.47
39	1	0.05	65.52
40	15	0.75	66.27
42	2	0.10	66.37
77	663	33.13	99.50
88888	9	0.45	99.95
99999	1	0.05	100.00

Total	2001	100.00	

Q4C First Activity
(Lists first beach activity named by respondent)

Q4C	Freq.	Percent	Cum.
0	37	5.15	5.15
1	82	11.42	16.57
2	64	8.91	25.49
3	1	0.14	25.63
4	1	0.14	25.77
5	17	2.37	28.13
6	14	1.95	30.08
8	74	10.31	40.39
9	6	0.84	41.23
10	1	0.14	41.36
11	9	1.25	42.62
12	20	2.79	45.40
13	11	1.53	46.94
16	21	2.92	49.86
18	6	0.84	50.70
19	29	4.04	54.74
20	26	3.62	58.36
21	42	5.85	64.21
22	4	0.56	64.76
23	7	0.97	65.74
24	176	24.51	90.25
27	9	1.25	91.50
28	5	0.70	92.20
29	14	1.95	94.15
31	2	0.28	94.43
32	3	0.42	94.85
33	1	0.14	94.99
36	8	1.11	96.10

37	4	0.56	96.66
38	1	0.14	96.80
39	1	0.14	96.94
40	11	1.53	98.47
42	2	0.28	98.75
88888	8	1.11	99.86
99999	1	0.14	100.00

Total	718	100.00	

SUM1 Second Activity

(Lists second beach activity named by respondent.)

SUM1	Freq.	Percent	Cum.
0	55	7.87	7.87
1	32	4.58	12.45
2	11	1.57	14.02
5	7	1.00	15.02
6	2	0.29	15.31
7	1	0.14	15.45
8	7	1.00	16.45
10	1	0.14	16.60
11	22	3.15	19.74
12	5	0.72	20.46
13	4	0.57	21.03
16	54	7.73	28.76
18	18	2.58	31.33
19	13	1.86	33.19
20	14	2.00	35.19
21	48	6.87	42.06
22	3	0.43	42.49
23	22	3.15	45.64
24	49	7.01	52.65
26	1	0.14	52.79
27	11	1.57	54.36
28	20	2.86	57.22
29	16	2.29	59.51
31	3	0.43	59.94
33	5	0.72	60.66
34	1	0.14	60.80
35	2	0.29	61.09
36	12	1.72	62.80
37	11	1.57	64.38
38	9	1.29	65.67
40	2	0.29	65.95
77	238	34.05	100.00

Total	699	100.00	

SUM2 Third Activity

(Lists third beach activity named by respondent.)

SUM2	Freq.	Percent	Cum.
0	29	6.30	6.30
1	1	0.22	6.52

2	1	0.22	6.74
6	3	0.65	7.39
7	2	0.43	7.83
11	7	1.52	9.35
12	2	0.43	9.78
13	1	0.22	10.00
16	10	2.17	12.17
18	3	0.65	12.83
19	1	0.22	13.04
21	6	1.30	14.35
22	3	0.65	15.00
23	4	0.87	15.87
24	25	5.43	21.30
27	8	1.74	23.04
28	1	0.22	23.26
31	2	0.43	23.70
33	1	0.22	23.91
36	2	0.43	24.35
37	2	0.43	24.78
38	9	1.96	26.74
40	1	0.22	26.96
77	335	72.83	99.78
88888	1	0.22	100.00

Total	460	100.00	

SUM3 Fourth Activity
(Lists fourth beach activity named by respondent.)

SUM3	Freq.	Percent	Cum.
0	2	1.61	1.61
2	2	1.61	3.23
11	4	3.23	6.45
12	1	0.81	7.26
13	1	0.81	8.06
16	2	1.61	9.68
18	10	8.06	17.74
21	1	0.81	18.55
23	2	1.61	20.16
24	2	1.61	21.77
27	2	1.61	23.39
28	2	1.61	25.00
29	1	0.81	25.81
38	1	0.81	26.61
40	1	0.81	27.42
77	90	72.58	100.00

Total	124	100.00	

Q4D Go in Water?

<1> Yes
 <3> No
 <5> Man

<88888> Don't know

<99999> Refused

Q4D	Freq.	Percent	Cum.
1	96	21.10	21.10
3	358	78.68	99.78
88888	1	0.22	100.00
Total	455	100.00	

Q4E How Deep?

- <1> Up to your ankles
- <2> Up to your waist
- <3> Up to your neck
- <4> You got your head wet

<88888> Don't Know

<99999> Refused

Q4E	Freq.	Percent	Cum.
1	70	45.75	45.75
2	40	26.14	71.90
3	3	1.96	73.86
4	40	26.14	100.00
Total	153	100.00	

DIARY 4 DESCRIPTIVE STATISTICS

Q4B

Beach Codes

(Beaches visited by beach code)

Q4B	Freq.	Percent	Cum.
1	2	0.17	0.17
5	1	0.09	0.26
6	1	0.09	0.35
7	6	0.52	0.87
9	2	0.17	1.05
12	1	0.09	1.13
13	6	0.52	1.66
15	1	0.09	1.74
22	58	5.05	6.79
24	1	0.09	6.88
26	10	0.87	7.75
29	4	0.35	8.10
31	40	3.48	11.59
32	5	0.44	12.02
33	1	0.09	12.11
35	2	0.17	12.28
37	29	2.53	14.81
39	2	0.17	14.98
42	5	0.44	15.42
43	9	0.78	16.20
44	9	0.78	16.99
51	1	0.09	17.07
57	1	0.09	17.16
59	74	6.45	23.61
62	38	3.31	26.92
63	65	5.66	32.58
69	39	3.40	35.98
73	1	0.09	36.06
75	7	0.61	36.67
77	3	0.26	36.93
79	56	4.88	41.81
81	15	1.31	43.12
83	3	0.26	43.38
85	2	0.17	43.55
87	58	5.05	48.61
92	28	2.44	51.05
94	94	8.19	59.23
96	4	0.35	59.58
97	1	0.09	59.67
102	1	0.09	59.76
103	2	0.17	59.93
110	6	0.52	60.45
114	39	3.40	63.85
120	5	0.44	64.29
122	17	1.48	65.77
123	11	0.96	66.72

124	16	1.39	68.12
126	1	0.09	68.21
127	88	7.67	75.87
129	46	4.01	79.88
136	3	0.26	80.14
137	1	0.09	80.23
144	1	0.09	80.31
145	8	0.70	81.01
147	1	0.09	81.10
149	50	4.36	85.45
155	5	0.44	85.89
158	11	0.96	86.85
159	12	1.05	87.89
160	3	0.26	88.15
161	8	0.70	88.85
168	3	0.26	89.11
169	3	0.26	89.37
172	3	0.26	89.63
176	4	0.35	89.98
177	2	0.17	90.16
185	8	0.70	90.85
186	1	0.09	90.94
190	6	0.52	91.46
194	1	0.09	91.55
196	15	1.31	92.86
197	2	0.17	93.03
205	2	0.17	93.21
218	7	0.61	93.82
240	3	0.26	94.08
246	1	0.09	94.16
263	1	0.09	94.25
269	2	0.17	94.43
271	1	0.09	94.51
287	1	0.09	94.60
291	4	0.35	94.95
298	1	0.09	95.03
300	5	0.44	95.47
301	14	1.22	96.69
302	3	0.26	96.95
303	6	0.52	97.47
304	4	0.35	97.82
88888	4	0.35	98.17
99999	21	1.83	100.00
<hr/>			
Total	1148	100.00	

Diary 4 activities (Q4C, SUM1, SUM2, SUM3)

diary4	Freq.	Percent	Cum.
0	126	4.21	4.21
1	137	4.57	8.78
2	105	3.51	12.29
3	26	0.87	13.16
5	21	0.70	13.86
6	28	0.93	14.79
7	4	0.13	14.92
8	57	1.90	16.83
9	5	0.17	16.99
10	1	0.03	17.03
11	44	1.47	18.50
12	50	1.67	20.17
13	25	0.83	21.00
14	1	0.03	21.04
16	104	3.47	24.51
17	1	0.03	24.54
18	26	0.87	25.41
19	165	5.51	30.92
20	92	3.07	33.99
21	151	5.04	39.03
22	86	2.87	41.90
23	40	1.34	43.24
24	305	10.18	53.42
27	114	3.81	57.23
28	85	2.84	60.07
29	37	1.24	61.30
30	9	0.30	61.60
31	16	0.53	62.14
32	11	0.37	62.50
33	21	0.70	63.21
34	2	0.07	63.27
35	28	0.93	64.21
36	60	2.00	66.21
37	15	0.50	66.71
38	45	1.50	68.21
77	946	31.59	99.80
88888	5	0.17	99.97
99999	1	0.03	100.00
Total	2995	100.00	

Q4C First Activity
(Lists first beach activity named by respondent)

Q4C	Freq.	Percent	Cum.
0	61	5.49	5.49
1	121	10.89	16.38
2	54	4.86	21.24
3	17	1.53	22.77
5	17	1.53	24.30
6	11	0.99	25.29
8	50	4.50	29.79

9	5	0.45	30.24
11	19	1.71	31.95
12	12	1.08	33.03
13	9	0.81	33.84
14	1	0.09	33.93
16	25	2.25	36.18
17	1	0.09	36.27
18	6	0.54	36.81
19	75	6.75	43.56
20	88	7.92	51.49
21	108	9.72	61.21
22	29	2.61	63.82
23	25	2.25	66.07
24	206	18.54	84.61
27	33	2.97	87.58
28	41	3.69	91.27
29	25	2.25	93.52
30	8	0.72	94.24
31	2	0.18	94.42
32	5	0.45	94.87
33	8	0.72	95.59
34	2	0.18	95.77
35	3	0.27	96.04
36	24	2.16	98.20
37	13	1.17	99.37
38	1	0.09	99.46
88888	5	0.45	99.91
99999	1	0.09	100.00

Total	1111	100.00	

SUM1 Second Activity
(Lists second beach activity named by respondent.)

SUM1	Freq.	Percent	Cum.
0	14	1.31	1.31
1	15	1.40	2.71
2	33	3.08	5.79
3	2	0.19	5.98
5	4	0.37	6.36
6	13	1.21	7.57
7	1	0.09	7.66
10	1	0.09	7.76
11	14	1.31	9.07
12	11	1.03	10.09
13	5	0.47	10.56
16	73	6.82	17.38
18	11	1.03	18.41
19	64	5.98	24.39
20	3	0.28	24.67
21	35	3.27	27.94
22	55	5.14	33.08
23	10	0.93	34.02
24	65	6.07	40.09
27	47	4.39	44.49
28	22	2.06	46.54

29	12	1.12	47.66
30	1	0.09	47.76
31	10	0.93	48.69
32	1	0.09	48.79
33	9	0.84	49.63
35	13	1.21	50.84
36	22	2.06	52.90
38	21	1.96	54.86
77	483	45.14	100.00

Total	1070	100.00	

SUM2 Third Activity
(Lists third beach activity named by respondent.)

SUM2	Freq.	Percent	Cum.

0	22	3.77	3.77
2	14	2.40	6.17
3	7	1.20	7.38
6	3	0.51	7.89
7	2	0.34	8.23
8	7	1.20	9.43
11	8	1.37	10.81
12	27	4.63	15.44
13	11	1.89	17.32
16	5	0.86	18.18
19	23	3.95	22.13
20	1	0.17	22.30
21	8	1.37	23.67
22	2	0.34	24.01
23	5	0.86	24.87
24	32	5.49	30.36
27	5	0.86	31.22
28	9	1.54	32.76
31	4	0.69	33.45
32	3	0.51	33.96
33	2	0.34	34.31
35	7	1.20	35.51
36	5	0.86	36.36
37	1	0.17	36.54
38	23	3.95	40.48
77	347	59.52	100.00

Total	583	100.00	

SUM3 Fourth Activity
(Lists fourth beach activity named by respondent.)

SUM3	Freq.	Percent	Cum.

0	29	12.55	12.55
1	1	0.43	12.99
2	4	1.73	14.72
6	1	0.43	15.15
7	1	0.43	15.58
11	3	1.30	16.88

16		1	0.43	17.32
18		9	3.90	21.21
19		3	1.30	22.51
24		2	0.87	23.38
27		29	12.55	35.93
28		13	5.63	41.56
32		2	0.87	42.42
33		2	0.87	43.29
35		5	2.16	45.45
36		9	3.90	49.35
37		1	0.43	49.78
77		116	50.22	100.00

Total		231	100.00	

Q4D Go in Water?

<1> Yes
 <3> No
 <5> Man

<88888> Don't know
 <99999> Refused

Q4D		Freq.	Percent	Cum.

1		179	25.87	25.87
3		510	73.70	99.57
88888		2	0.29	99.86
99999		1	0.14	100.00

Total		692	100.00	

Q4E How Deep?

<1> Up to your ankles
 <2> Up to your waist
 <3> Up to your neck
 <4> You got your head wet

<88888> Don't Know
 <99999> Refused

Q4E		Freq.	Percent	Cum.

1		110	44.35	44.35
2		43	17.34	61.69
3		7	2.82	64.52
4		83	33.47	97.98
5		5	2.02	100.00

Total		248	100.00	

Beach Codes

<1>1000 Steps Beach
 <2>10th Street-Laguna Poc South

<3>17th Street-Newport Beach
<4>204s-Laguna Poc South
<5>6th Street-Seal Beach
<6>Abalone Cove
<7>Alamitos Bay
<8>Alamitos State Beach
<9>Aliso Creek Beach
<10>Anderson Street-Surfside
<11>Anita-Laguna Poc South
<12>Arroyo Sequit
<13>Balboa Beach
<14>Beach Road-Laguna Poc South
<15>Belmont Shores
<16>Big Corona
<17>Big Dume
<18>Big Rock
<19>Blackies-Newport Beach
<20>Bluebird-Laguna Poc South
<21>Boat Canyon-Laguna Poc North
<22>Bolsa Chica Beach
<23>Broad Beach
<24>Brook Street-Laguna Poc South
<25>Bucky Gully
<26>Cabrillo Beach
<27>Calafia-Laguna Poc South
<28>Calafia Beach Park
<29>Capistrano Beach
<30>Castle Rock
<31>Corona Del Mar Beach
<32>Corral State Beach
<33>County Line Beach-Ventura (Yerba Buena)
<34>Crabs
<35>Crescent Bay-Laguna Poc North
<36>Cress Street-Laguna Poc South
<37>Crystal Cove
<38>Dan Blocker Beach
<39>Dana Point Marina-Laguna Poc South
<40>Depot, The-Laguna Poc South
<41>Divers Cove-Laguna Poc North
<42>Dockweiler Beach
<43>Dog Run Beach-Bolsa Chica
<44>Doheny Beach
<45>Dolphin Street-Seal Beach
<46>Drainpipes
<47>El Matedor Beach
<48>El Morrow
<49>El Pescador Beach
<50>El Porto Beach
<51>El Segundo Beach
<52>Emerald Bay- Laguna Poc North
<53>Escondido Beach
<54>Fishermans Cove-Laguna Poc North
<55>Garbage-Laguna Poc North
<56>Goldenwest
<57>Grand Avenue
<58>Groins, The-Newport Beach
<59>Hermosa Beach

<60>Hole in the Fence-Laguna Poc South
<61>Hole, The-Surfside
<62>Huntington City Beach
<63>Huntington State Beach
<64>Irvine Cove-Laguna Poc North
<65>Jack in the Box-Bolsa Chica
<66>Jetty, The-Surfside
<67>La Piedra
<68>Ladders, The-Newport Beach
<69>Laguna Beach
<70>Laguna Pocket Beaches North
<71>Laguna Pocket Beaches South
<72>Lagunita-Laguna Poc South
<73>Las Tunas Beach
<74>Lasuens-Laguna Poc South
<75>Leo Carillo Beach
<76>Linda Lane-Laguna Poc South
<77>Little Corona Del Mar Beach
<78>Little Dume
<79>Long Beach
<80>Lost Winds-Laguna Poc South
<81>Mailibu Lagoon State Beach
<82>Main Beach-The Wedge
<83>Main Beach-Laguna Beach
<84>Malaga Cove
<85>Malibu (Surfrider)
<86>Malibu Road Beach
<87>Manhattan Beach
<88>Marina Del Ray Swimming Beach
<89>Mariposa Point-Laguna Poc South
<90>Monarch Beach-Laguna Poc South
<91>Moss Street-Laguna Poc South
<92>Mothers Beach
<93>Mountain Road-Laguna Poc South
<94>Newport Beach
<95>Newport Pier
<96>Newport River Jetties-Newport Beach
<97>Nicholas Canyon Beach
<98>North Beach-Laguna Poc South
<99>Oak Street - Laguna Poc South
<100>Old Mans - Bolsa Chica
<101>Ole Hanson Beach
<102>Palos Verdes Estates
<103>Paradise Cove
<104>Pearl Street - Laguna Poc South
<105>Picnic Beach - Laguna Poc North
<106>Pier, The - Laguna Poc South
<107>Pier, The - North Side - Seal Beach
<108>Pier, The - South Side - Seal Beach
<109>Poche Beach
<110>Point Dume Beach
<111>Point Fermin Beach
<112>Point, The - Newport Beach
<113>RAT Beach
<114>Redondo Beach
<115>River Jetty
<116>Rivieras - laguna Poc South

<117>RJs - Newport Beach
 <118>Rock Pile - Laguna Poc North
 <119>Rocks, The - Seal Beach
 <120>Royal Palms Beach
 <121>Saint Anns - Laguna Poc South
 <122>Salt Creek Beach
 <123>San Clemente City Beach
 <124>San Clemente State Beach
 <125>San Juan Creek - Laguna Poc South
 <126>Santa Ana River Mouth Beach
 <127>Santa Monica Beach
 <128>Scottsmans
 <129>Seal Beach
 <130>Second Spot - Laguna Poc South
 <131>Secos
 <132>Shaws Cove - Laguna Poc North
 <133>Sleepy Hollow - Laguna Poc north
 <134>Strands - Laguna Poc South
 <135>Street Numbers (52nd, 53rd, Etc) Newport
 <136>Sunset Beach
 <137>Surfrider (Malibu)
 <138>Surfside Beach
 <139>T Street - Laguna Poc South
 <140>Thalia - Laguna Poc South
 <141>The Bluff - Bolsa Chica
 <142>Three Arch Bay - Laguna Poc South
 <143>Tombstones - Surfside
 <144>Topanga Beach
 <145>Torrance Beach
 <146>Trafalger Street - Laguna Poc South
 <147>Treasure Island - Laguna Poc South
 <148>Upper Redondo Beach
 <149>Venice Beach
 <150>Victoria - Laguna Poc South
 <151>Wedge, The
 <152>West Street Beach
 <153>Westward
 <154>White point County Beach
 <155>Will Rogers Beach
 <156>Woods Cove - Laguna Poc South
 <157>Zonker Harris Beach
 <158>Zuma Beach
 <159>Zuma County Beach>
 <160>San Onofre State Beach-North
 <161>San Onofre State Beach-South
 <162>Beacon's Beach (Leucadia State Beach)
 <163>Black's Beach (Torrey Pines City Beach)
 <164>Boneyard Beach
 <165>Boomer Beach
 <166>Border Field State Park
 <167>Cardiff State Beach
 <168>Carlsbad City Beach
 <169>Carlsbad State Beach (Tamarack Beach)
 <170>Children's Pool Beach
 <171>Coast Boulevard Park
 <172>Coronado City Beach
 <173>Coronado Shores Beach

<174>Crown Point Shores
<175>D Street Viewpoint
<176>Del Mar Bluffs City Park
<177>Del Mar City Beach
<178>Encinitas Beach
<179>Fletcher Cove Park
<180>Gator Beach
<181>Harbor Beach
<182>Hermosa Terrace Park
<183>Imperial Beach
<184>La Jolla Cove
<185>La Jolla Shores Beach
<186>La Jolla Strand Park
<187>La Playa
<188>Las Pulgas (Red) Beach
<189>Marine Street Beach
<190>Mission Beach
<191>Moonlight Beach
<192>Nicholson Point Park
<193>North Delta Beach
<194>Ocean Beach City Beach
<195>Ocean Beach
<196>Oceanside City Beach
<197>Pacific Beach
<198>Palisades Park
<199>Ponto Beach
<200>San Elijo State Beach
<201>Scripps Beach
<202>Seascape Shores
<203>Silver Strand State Beach
<204>Ski Beach
<205>Solana Beach
<206>South Carlsbad State Beach
<207>South Oceanside Beach
<208>Stone Steps Beach
<209>Sunset Cliffs Park
<210>Swami's
<211>Tide Beach Park
<212>Tijuana River Estuary Reserve
<213>Torrey Pines City Beach
<214>Torrey Pines State Beach
<215>Tourmaline Surfing Park
<216>Trestles Beach
<217>Windansea Beach
<218>Channel Island. Harbor Beach
<219>Emma Wood State Beach (Ventura Overhead)
<220>Fairgrounds Beach
<221>Faria Beach
<222>Faria County Park Out In John's)
<223>Hobson County Park
<224>Hollywood Beach
<225>Hollywood-by-the-Sea (Silver Strand
<226>La Conchita Beach
<227>La Jolla Beach
<228>Little Rincon (Mussel Shoals)
<229>Mandalay County Park
<230>Marina Cove Beach (School Yards)

<231>Marina Park
<232>McGrath State Beach
<233>Mussel Shoals Beach (Little Rincon)
<234>North Beach
<235>Oil Piers Beach (now Seacliff)
<236>Ormond Beach
<237>Oxnard Shores
<238>Oxnard State Beach
<239>Peninsula Beach
<240>Point Mugu Beach (missile range)
<241>Port Hueneme Beach Park
<242>Promenade Park
<243>Rincon Beach State Park
<244>Rincon Parkway North
<245>Rincon Parkway South
<246>San Buenaventura State Beach
<247>Santa Clara Rivermouth Beach
<248>School Yards (Marina Cove)
<249>Seacliff (formerly Oil Piers)
<250>Seaside Wilderness Park
<251>Silver Strand Beach (Hollywood-)
<252>Solimar Beach (offshore reef)
<253>Staircase Beach
<254>Super Tubes
<255>Surfer's Point
<256>Sycamore Cove Beach
<257>Thornhill Broome Beach
<258>Arroyo Burro Beach County Park
<259>Biltmore Hotel Beach
<260>Butterfly Beach
<261>Campus (Goleta Point)
<262>Carpinteria City Beach
<263>Carpinteria State Beach
<264>Cat Canyon Beach
<265>Christi Beach
<266>Coal Oil Point Natural Reserve
<267>Devereux Point (Coal Oil)
<268>East Beach
<269>El Capitan State Beach
<270>Gaviota State Park
<271>Goleta Beach County Park
<272>Goleta Point Beach (Campus)
<273>Hammonds Beach (Hammond's Reef Beach)
<274>Hazards Beach
<275>Hendry's Beach
<276>Isla Vista Beach
<277>Jalama Beach County Park
<278>Leadbetter Beach
<279>Lookout County Park
<280>Mesa Lane Beach
<281>Miramar Beach
<282>Molino Canyon
<283>Nipomo Dunes Preserve
<284>Ocean Beach County Park
<285>Palm Park
<286>Point Sal State Beach
<287>Refugio State Beach

<288> Rincon Beach County Park
<289> Rincon Point
<290> Sands Beach
<291> Santa Barbara Harbor
<292> Shoreline Park
<293> Tajiguas Beach
<294> Thousand Steps Beach
<295> UC Santa Barbara Beach
<296> Vandenberg Air Force Base Beach
<297> Vandenberg Air Force Base Fishing
<298> West Beach

<300> Unknown Orange County beach
<301> Unknown LA County beach
<302> Unknown San Diego County beach
<303> Unknown Ventura County beach
<304> Unknown Santa Barbara County beach

<777> Went to more than one beach that day (q4b)
<777> No more beaches that day (q5i, q7f)
<999> Beach not listed

<888> No more (qxi01, qxj01, qxk01, qxl01)
<999> No more (qgl62, qgl63)

<88888> Don't know
<99999> Refused

Activity Codes

<1> Bicycling	<14> Sailing
<2> Body Boarding/Body surfing /skimboarding	<15> Scuba diving
<3> Boating	<16> Shopping/dining
<4> Canoeing	<17> Snorkeling
<5> Fishing (shore or pier)	<18> Sand football/soccer
<6> Frisbee	<19> Sunbathing
<7> Jet boating/Jet skiing personal water craft	<20> Surfing
<8> Jogging	<21> Swimming
<9> Kayaking	<22> Volleyball
<10> Kite Flying	<23> Wading
<11> Picknicing	<24> Walking
<12> People watching	<25> Water skiing
<13> Rollerblading/roller skates	<26> Windsurfing / boardsailing
<29> Splashing in water	<27> Enjoying the view
<31> Reading	<28> Activities with children
<33> Bar-B-Q	<30> Walking the dog
<35> Beachcombing	<32> Hiking
<37> Watched Fireworks	<34> Amusement park/ arcade
<39> Sleeping	<36> Eating/ drinking
<41> Bonfire	<38> Played in the sand
<43> Relaxing	<40> Camping
	<42> Skateboarding

<0> Other
<77> Nothing else
<88888> Don't Know

APPENDIX 3 – supplement to Report # 5

Tabulated results for main activity diary 1 – diary 6

(Note that last page had activity codes.)

DIARY 1

d1_ma	Freq.	Percent	Cum.
0	24	9.27	9.27
1	24	9.27	18.53
2	1	0.39	18.92
3	3	1.16	20.08
4	3	1.16	21.24
5	2	0.77	22.01
6	2	0.77	22.78
8	10	3.86	26.64
9	1	0.39	27.03
10	1	0.39	27.41
11	9	3.47	30.89
12	7	2.70	33.59
13	9	3.47	37.07
16	21	8.11	45.17
19	3	1.16	46.33
20	11	4.25	50.58
21	1	0.39	50.97
22	3	1.16	52.12
23	3	1.16	53.28
24	105	40.54	93.82
27	13	5.02	98.84
28	2	0.77	99.61
30	1	0.39	100.00
Total	259	100.00	

DIARY 2

d2_ma	Freq.	Percent	Cum.
0	12	4.41	4.41
1	22	8.09	12.50
2	2	0.74	13.24
3	1	0.37	13.60
4	3	1.10	14.71
5	3	1.10	15.81
6	1	0.37	16.18
8	6	2.21	18.38
10	1	0.37	18.75
11	17	6.25	25.00
12	5	1.84	26.84
13	7	2.57	29.41
16	14	5.15	34.56
18	2	0.74	35.29
19	10	3.68	38.97

20	9	3.31	42.28
21	6	2.21	44.49
22	3	1.10	45.59
23	3	1.10	46.69
24	109	40.07	86.76
27	13	4.78	91.54
28	9	3.31	94.85
30	3	1.10	95.96
31	1	0.37	96.32
32	2	0.74	97.06
33	2	0.74	97.79
34	1	0.37	98.16
35	3	1.10	99.26
36	1	0.37	99.63
88888	1	0.37	100.00

Total	272	100.00	

DIARY 3

d3_ma	Freq.	Percent	Cum.
0	11	5.24	5.24
1	11	5.24	10.48
2	9	4.29	14.76
3	1	0.48	15.24
4	1	0.48	15.71
5	6	2.86	18.57
6	3	1.43	20.00
8	8	3.81	23.81
9	1	0.48	24.29
11	4	1.90	26.19
12	3	1.43	27.62
13	3	1.43	29.05
16	8	3.81	32.86
18	1	0.48	33.33
19	19	9.05	42.38
20	8	3.81	46.19
21	22	10.48	56.67
22	1	0.48	57.14
23	2	0.95	58.10
24	60	28.57	86.67
27	1	0.48	87.14
28	6	2.86	90.00
29	5	2.38	92.38
31	1	0.48	92.86
32	1	0.48	93.33
36	4	1.90	95.24
37	2	0.95	96.19
38	1	0.48	96.67
40	5	2.38	99.05
42	1	0.48	99.52
88888	1	0.48	100.00

Total	210	100.00	

DIARY 4

d4_ma	Freq.	Percent	Cum.
0	5	1.71	1.71
1	12	4.10	5.80
2	7	2.39	8.19
3	5	1.71	9.90
5	6	2.05	11.95
6	3	1.02	12.97
8	4	1.37	14.33
9	2	0.68	15.02
11	10	3.41	18.43
12	3	1.02	19.45
13	3	1.02	20.48
16	12	4.10	24.57
18	2	0.68	25.26
19	27	9.22	34.47
20	11	3.75	38.23
21	29	9.90	48.12
22	9	3.07	51.19
23	10	3.41	54.61
24	73	24.91	79.52
27	12	4.10	83.62
28	15	5.12	88.74
29	8	2.73	91.47
30	5	1.71	93.17
31	1	0.34	93.52
32	1	0.34	93.86
33	1	0.34	94.20
35	1	0.34	94.54
36	8	2.73	97.27
37	5	1.71	98.98
38	3	1.02	100.00
Total	293	100.00	

DIARY 5

d5_ma	Freq.	Percent	Cum.
0	56	17.55	17.55
1	22	6.90	24.45
2	8	2.51	26.96
5	7	2.19	29.15
8	10	3.13	32.29
9	1	0.31	32.60
11	6	1.88	34.48
12	3	0.94	35.42
13	1	0.31	35.74
14	1	0.31	36.05
16	11	3.45	39.50
18	2	0.63	40.13
19	34	10.66	50.78
20	12	3.76	54.55
21	20	6.27	60.82

22	5	1.57	62.38
23	5	1.57	63.95
24	94	29.47	93.42
27	6	1.88	95.30
28	10	3.13	98.43
29	3	0.94	99.37
30	2	0.63	100.00

Total	319	100.00	

DIARY 6

d6_ma	Freq.	Percent	Cum.
0	11	5.79	5.79
1	9	4.74	10.53
2	2	1.05	11.58
3	1	0.53	12.11
5	5	2.63	14.74
8	10	5.26	20.00
11	2	1.05	21.05
12	4	2.11	23.16
13	6	3.16	26.32
16	27	14.21	40.53
18	1	0.53	41.05
19	2	1.05	42.11
20	15	7.89	50.00
21	8	4.21	54.21
22	2	1.05	55.26
23	1	0.53	55.79
24	63	33.16	88.95
25	1	0.53	89.47
27	7	3.68	93.16
28	3	1.58	94.74
29	1	0.53	95.26
31	2	1.05	96.32
34	2	1.05	97.37
36	2	1.05	98.42
38	1	0.53	98.95
40	2	1.05	100.00

Total	190	100.00	

ⁱ The actual dates in this calendar correspond to April, 2002. The captions and pictures are as they appeared in 2000.

ⁱⁱTo increase the processing speed of the computer, the program was set to accept only up to thirty individual trips. For the vast majority of respondents, this was sufficient.

ⁱⁱⁱ The ten modes represent the five modes originally described multiplied by the two additional modes for those who went to the beach less than five times and those who went to the beach more than five times.

^{iv} Cases in which the phone number was for the exclusive use of a minor are not included in the number of working residential numbers

^v The determination of diminished capacity is at the discretion of the interviewer. Interviewers are required to discuss the interview with a supervisor prior to coding any case as "bad data".

-
- ^{vi} This response rate for the initial recruitment, as well as the response rate for the replenishment survey were calculated using AAPOR formula RR3.
- ^{vii} Because of anomalies in the way that the CASES program tracks elapsed time, outliers are removed before an average time is taken. For the purpose of this study an outlier is defined as a survey that is reported to have taken less than 2 minutes or more than 1 hour.
- ^{viii} For the purpose of this study, the term eligible is defined as a household where it was determined at least one individual had gone to the beach within 12 months of the date of enumeration. Refusals and partials are not included in this line of Table 2 and Table 4. However cases where eligibility was determined but the survey was never conducted are included in this line of Table 2 and Table 4.
- ^{ix} A total of 887 were initially recruited into the panel. Of that number, 26 were not called during the survey portion of the first diary because they were used in the pretest.
- ^x Response rates for the diary surveys are calculated using AAPOR standards for RR1.
- ^{xi} Elapsed times of one minute were included in this calculation for diary surveys.
- ^{xii} The clean up refers to the fact that a small number of cases were never released into the diary sample because the address data collected during the survey was incorrect or incomplete.
- ^{xiii} Differences in the numbers of cases released between diaries reflect differences due to the fact that some respondents were lost or dropped out of the study between diaries. If the respondent could be found again, or if they could be convinced to rejoin the panel, they were added back into the sample.
- ^{xiv} Part of the dip in response rate can be attributed to a late start in calling.
- ^{xv} Additional respondents had been added after the replenishment survey.
- ^{xvi} In this study actually defined as "Other Non-Response".
- ^{xvii} For this study "Respondent Unavailable", "Incorrect Phone Number", and "Never Answered".
- ^{xviii} For the purpose of these calculations, "eligible" is defined as working household numbers.
- ^{xix} Piekarski et al, "Telephony and Telephone Sampling: The Dynamics of Change".
[[http://www.worldopinion.com/news?cmd=item&id=3966.](http://www.worldopinion.com/news?cmd=item&id=3966)]