california health interview survey

Report Five

CHIS 2005 Methodology Series

Weighting and Variance Estimation

CALIFORNIA HEALTH INTERVIEW SURVEY

CHIS 2005 METHODOLOGY SERIES

REPORT 5

WEIGHTING AND VARIANCE ESTIMATION

April 2007

This report was prepared for the California Health Interview Survey by Ismael Flores Cervantes, J. Michael Brick, Laura Alvarez-Rojas and Michael E. Jones of Westat.

www.chis.ucla.edu

This report describes the weighting and variance estimation methods used in CHIS 2005. This report presents the steps used to create the analytical weights for analyzing the data from the adult, child, and adolescent interviews.

Suggested citation:

California Health Interview Survey. *CHIS 2005 Methodology Series: Report 5 – Weighting and Variance Estimation*. Los Angeles, CA: UCLA Center for Health Policy Research, 2007.

The California Health Interview Survey is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Funding for CHIS 2005 came from multiple sources: the California Department of Health Services, the California Endowment, the National Cancer Institute, the First 5 California, the Robert Wood Johnson Foundation, the California Department of Mental Health, the California Office of the Patient Advocate, Kaiser Permanente, the San Diego County Human and Health Services Agency, the Marin County Department of Human and Health Services, the First 5 Marin Children and Families Commission, the Center for Public Policy Research at the University of California, Davis, the U. S. Centers for Disease Control and Prevention, the Solano County Health and Social Services Department, and the Humboldt County Department of Health and Human Services.

PREFACE

Weighting and Variance Estimation is the fifth in a series of methodological reports describing the 2005 California Health Interview Survey (CHIS 2005). The other reports are listed below.

CHIS is a collaborative project of the University of California, Los Angeles (UCLA) Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Westat was responsible for the data collection and the preparation of five methodological reports from the 2005 survey. The survey examines public health and health care access issues in California. The CHIS telephone survey is the largest state health survey ever undertaken in the United States. The plan is to monitor the health of Californians and examine changes over time by conducting periodic surveys in the future.

Methodological Reports

The first five methodological reports for CHIS 2005 are as follows:

- Report 1: Sample Design;
- Report 2: Data Collection Methods;
- Report 3: Data Processing Procedures;
- Report 4: Response Rates; and
- Report 5: Weighting and Variance Estimation.

The reports are interrelated and contain many references to each other. For ease of presentation, the references are simply labeled by the report numbers given above.

This report describes the weighting and variance estimation methods from CHIS 2005. The purpose of weighting the survey data is to permit analysts to produce estimates of the health characteristics for the entire California population and subgroups including counties, and in some cases, cities. This report presents the steps used to create the analytical weights for analyzing the data from the adult, child, and adolescent interviews.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Figures

1. CHIS 2005 DESIGN AND METHODOLOGY SUMMARY

1.1 Overview

The California Health Interview Survey (CHIS) is a population-based random-digit dial telephone survey of California's population conducted every other year since 2001. CHIS is the largest health survey conducted in any state and one of the largest health surveys in the nation. CHIS is based at the UCLA Center for Health Policy Research and is conducted in collaboration with the California Department of Health Services and the Public Health Institute. CHIS collects extensive information for all age groups on health status, health conditions, health-related behaviors, health insurance coverage, access to health care services, and other health and health related issues.

The sample is designed to meet and optimize two goals:

1. provide estimates for large- and medium-sized counties in the state, and for groups of the smallest counties (based on population size), and
2. provide statewide estimates for California's overall population, its major racial and ethnic groups, as well as several ethnic subgroups.

The CHIS sample is representative of California's non-institutionalized population living in households.

This series of reports describes the methods used in collecting data for CHIS 2005, the third CHIS data collection cycle, which was conducted between July 2005 and April 2006. The previous CHIS cycles (2001 and 2003) are described in similar series, available at http://www.chis.ucla.edu/methods_main.html.

CHIS data and results are used extensively by federal and State agencies, local public health agencies and organizations, advocacy and community organizations, other local agencies, hospitals, community clinics, health plans, foundations, and researchers. The data are widely used for analyses and publications to assess public health and health care needs, to develop and advocate policies to meet those needs, and to plan and budget health care coverage and services.

1.2 Sample Design Objectives

To achieve the sample design objectives stated above, CHIS employed a multi-stage sample design. First, the state was divided into 44 geographic sampling strata, including 41 single-county strata and three multi-county strata comprised of the 17 remaining counties. Second, within each geographic stratum, residential telephone numbers were selected through random-digit dial (RDD) sampling, and within each household, one adult (age 18 and over) respondent was randomly selected. In those households with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one child were randomly selected; the adolescent was interviewed directly, and the adult most knowledgeable about the child's health completed the child interview.

Table 1-1 shows the 44 sampling strata for CHIS 2005. A sufficient number of adult interviews were allocated to each stratum to support the first sample design objective. The geographic stratification of the state was revised from the design used in previous CHIS cycles, increasing the number of individual counties from 33 to 41.

Table 1-1. California county and county group strata used in the CHIS 2005 sample design

1. Los Angeles	16. Sonoma	31. Napa
2. San Diego	17. Stanislaus	32. Kings
3. Orange	18. Santa Barbara	33. Madera
4. Santa Clara	19. Solano	34. Monterey
5. San Bernardino	20. Tulare	35. Humboldt
6. Riverside	21. Santa Cruz	36. Nevada
7. Alameda	22. Marin	37. Mendocino
8. Sacramento	23. San Luis Obispo	38. Sutter
9. Contra Costa	24. Placer	39. Yuba
10. Fresno	25. Merced	40. Lake
11. San Francisco	26. Butte	41. San Benito
12. Ventura	27. Shasta	42. Colusa, Glen, Tehama
13. San Mateo	28. Yolo	43. Plumas, Sierra, Siskiyou, Lassen,
14. Kern	29. El Dorado	Modoc, Trinity, Del Norte
15. San Joaquin	30. Imperial	44. Mariposa, Mono, Tuolumne, Alpine, Amador, Calaveras, Inyo

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The samples in Marin, Humboldt, and Solano Counties were enhanced with additional funding. Additional samples were also selected statewide and in San Diego County to increase the number of child interviews; telephone numbers selected in these two additional samples were

screened to identify households with children ages 0 to 11. All supplemental samples were implemented with and incorporated into the original statewide RDD sample.

The main RDD CHIS sample size is sufficient to accomplish the second objective. To increase the precision of estimates for Koreans and Vietnamese, areas with relatively high concentrations of these groups were sampled at higher rates. These geographically targeted oversamples were supplemented by telephone numbers associated with group-specific surnames drawn from listed telephone directories to further increase the sample size for Koreans and Vietnamese.

1.3 Data Collection

To capture the rich diversity of the California population, interviews were conducted in five languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, and Korean. These languages were chosen based on analysis of 2000 Census data to identify the languages that would cover the largest number of Californians in the CHIS sample that either did not speak English or did not speak English well enough to otherwise participate.

Westat, a private firm that specializes in statistical research and large-scale sample surveys, conducted the CHIS 2005 data collection under contract with the UCLA Center for Health Policy Research. Westat staff interviewed one randomly selected adult in each sampled household, and sampled one adolescent and one child if present in the household and the sampled adult was the parent or legal guardian. Up to three interviews could have been completed in each sampled household. In households with children where the sampled adult was not the screener respondent, children and adolescents could be sampled as part of the screening interview, and the extended child (and adolescent) interviews could be completed before the adult interview. This "child-first" procedure was new for CHIS 2005 and substantially increased the yield of child interviews. While numerous subsequent attempts were made to complete the adult interview, there were completed child and/or adolescent interviews in households for which an adult interview was not completed. Table 1-2 shows the number of completed adult, child, and adolescent interviews in CHIS 2005 by the type of sample (RDD or supplemental sample).

Table 1-2. Number of completed CHIS 2005 interviews by type of sample and instrument

Type of sample	Adult	Child	Adolescent
Total RDD + supplemental cases	43,020	11,358	4,029
RDD			
Base plus county supplements	41,074	9,605	3,739
Statewide child supplement	525	511	84
San Diego child supplement	1,143	1,160	186
Supplemental samples:			
Korean	199	60	14
Vietnamese	79	22	6

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Interviews in all languages were administered using Westat's computer-assisted telephone interviewing (CATI) system. The average adult interview took 35 minutes to complete. The average child and adolescent interviews took 15 minutes and 20 minutes, respectively. For "child-first" interviews, additional household information asked as part of the child interview averaged almost 8 minutes. Interviews in non-English languages generally took longer to complete. Just over 10 percent of the adult interviews were completed in a language other than English, as were 18 percent of all child (parent proxy) interviews and 7 percent of all adolescent interviews.

Table 1-3 shows the major topic areas for each of the three survey instruments (adult, child, and adolescent).

1.4 Response Rates

The overall response rate for CHIS 2005 is a composite of the screener completion rate (i.e., success in introducing the survey to a household and randomly selecting an adult to be interviewed) and the extended interview completion rate (i.e., success in getting one or more selected persons to complete the extended interview). To maximize the response rate, especially at the screener stage, an advance letter in five languages was mailed to all sampled telephone numbers for which an address could be obtained from reverse directory services. An advance letter was mailed for approximately 67 percent of the sampled telephone numbers. In CHIS 2005, for the first time a $2 bill was included with the advance letter to promote cooperation. CHIS 2005 also included methodological experiments to test the effects on response of the incentive and different advance letter treatments.

Table 1-3. CHIS 2005 survey topic areas by instrument

Health status	**Adult**	**Teen**	**Child**
General health status, height and weight	✓	✓	✓
Days missed from school due to health problems	✓	✓	✓
Health conditions	**Adult**	**Teen**	**Child**
Asthma	✓	✓	✓
Diabetes	✓	✓	
Heart disease, high blood pressure, epilepsy	✓		
Physical disability/need for special equipment	✓		
Parental concerns with child development, attention deficit disorder (ADD)			✓
Mental health	**Adult**	**Teen**	**Child**
Mental health status	✓		✓
Perceived need, use of mental health services	✓		✓
Emotional functioning		✓	
Health behaviors	**Adult**	**Teen**	**Child**
Dietary intake	✓	✓	✓
Physical activity and exercise	✓	✓	✓
Walking for transportation and leisure	✓		
Sedentary time		✓	✓
Body image		✓	
Flu Shot	✓		
Alcohol and tobacco use	✓	✓	
Drug use		✓	
Sexual behavior, STD testing, birth control practices	✓	✓	
Women's health	**Adult**	**Teen**	**Child**
Pap test screening, mammography screening, hormone replacement therapy	✓		
Emergency contraception		✓	
Pregnancy status	✓	✓	
Cancer history and prevention	**Adult**	**Teen**	**Child**
Cancer history of respondent and family history	✓		
Colon cancer screening, prostrate cancer (PSA) test	✓		
Dental health	**Adult**	**Teen**	**Child**
Last dental visit			✓
Dental insurance coverage		✓	✓
Injury	**Adult**	**Teen**	**Child**
Serious injuries (frequency, cause)		✓	✓
Injury prevention behaviors (bike helmets, seatbelts)		✓	

Table 1-3. CHIS 2005 survey topic areas by instrument (Continued)

Food insecurity/hunger	**Adult**	**Teen**	**Child**
Availability of food in household over past 12 months	✓		
Food environment	**Adult**	**Teen**	**Child**
Quality of food stores in area, where does teen/child eat lunch and breakfast		✓	✓
School has vending machines		✓	
Access to and use of health care	**Adult**	**Teen**	**Child**
Usual source of care, visits to medical doctor	✓	✓	✓
Emergency room visits	✓	✓	✓
Delays in getting care (prescriptions, tests, treatment)	✓	✓	✓
Racial/ethnic discrimination in health care, MD discussed diet and exercise	✓		
Communication problems with doctor	✓	✓	✓
Ability and parental knowledge of teen contacting a doctor		✓	
Health insurance	**Adult**	**Teen**	**Child**
Current insurance coverage, spouse's coverage, who pays for coverage	✓	✓	✓
Health plan enrollment, characteristics of plan	✓	✓	✓
Whether employer offers coverage, respondent/spouse eligibility	✓		
Coverage over past 12 months	✓	✓	✓
Reasons for lack of insurance	✓	✓	✓
Public program eligibility	**Adult**	**Teen**	**Child**
Household poverty level	✓		
Program participation (TANF, CalWorks, Public Housing, Food Stamps, SSI, SSDI, WIC)	✓	✓	✓
Assets, alimony/child support/social security/pension	✓		
Eligible for Medi-Cal and healthy families	✓	✓	✓
Reason for Medi-Cal nonparticipation among potential eligibles	✓	✓	✓
Neighborhood and housing	**Adult**	**Teen**	**Child**
Neighborhood safety	✓	✓	
Home ownership, number of rooms, amount of mortgage/rent	✓		
Parental involvement/adult supervision	**Adult**	**Teen**	**Child**
Parental presence after school, parental knowledge of teen's activities		✓	
Child's activities with family			✓

Table 1-3. CHIS 2005 survey topic areas by instrument (Continued)

Child care and school attendance	Adult	Teen	Child
Current child care arrangements			✓
Paid child care	✓		
First 5 California: receipt of parent kit and attitudes towards preschool			✓
Preschool/school attendance, public/private school		✓	✓
Employment	**Adult**	**Teen**	**Child**
Employment status, spouse's employment status	✓		
Work in last week, industry and occupation	✓		
Hours worked at all jobs	✓	✓	
Income	**Adult**	**Teen**	**Child**
Respondent's and spouse's earnings last month before taxes	✓		
Household income (annual before taxes)	✓		
Number of persons supported by household income	✓		
Respondent characteristics	**Adult**	**Teen**	**Child**
Age, gender, height, weight, education	✓	✓	✓
Race and ethnicity	✓	✓	✓
Marital status	✓		
Sexual orientation	✓	✓	
Citizenship, immigration status, country of birth, length of time in U.S., languages spoken at home, English language proficiency	✓	✓	✓

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The CHIS 2005 screener completion rate was 49.8 percent and was higher for households that were sent the advance letter. The extended interview completion rate varied across the adult, child and adolescent interviews. Multiplying the screener and extended rates gives an overall response rate for each type of interview. At the household level, the percentage of households completing one or more of the extended interviews (adult, child, and/or adolescent) is a useful summary of the overall success of the study. For CHIS 2005, the household response rate was 29.5 percent (the product of the screener response rate and the completion rate at the household level of 59.3 percent). The 2005 survey is the first time that a household response rate has been reported because in earlier cycles the adult interview had to be completed before the child or the adolescent interview (i.e., the household rate equaled the adult rate). The adult extended completion rate for 2005 was 54.0 percent, resulting in an overall adult response rate of 26.9 percent for adults. All of the household and person level response rates vary by sampling stratum. For more information about the CHIS 2005 response rates, please see *CHIS 2005 Methodology Series: Report 4 – Response Rates*.

The CHIS response rate is comparable to response rates of other scientific telephone surveys in California, such as the 2005 California Behavioral Risk Factor Surveillance System (BRFSS) Survey with an overall response rate of 29.2 percent. California as a whole and the state's urban areas in particular, are among the most difficult parts of the nation in which to conduct telephone interviews. Survey response rates tend to be lower in California than nationally, and over the past decade response rates have been declining both nationally and in California.

One way to judge the representativeness of a population survey is to "benchmark" its results against those of other reliable data sources. The CHIS 2001 and 2003 samples yielded unweighted and weighted population distributions and rates that are comparable to those obtained from other sources. The demographic characteristics of the CHIS 2001 sample (such as race, ethnicity, and income) were very similar to those obtained from 2000 Census data. CHIS 2001 respondents also had health characteristics and behaviors that were very similar to those found in other extensively used surveys, such as the California BRFSS. The UCLA Center for Health Policy Research is conducting an extensive benchmarking project for CHIS data.

Adults who completed at least approximately 80 percent of the questionnaire (i.e., through Section J on Health Care Utilization and Access), after all follow-up attempts were exhausted to complete the full questionnaire, were counted as "complete." At least some items in the employment and income series or public program eligibility and food insecurity series are missing from those cases that did not complete the entire interview.

Proxy interviews were allowed for frail and ill persons over the age of 65 who were unable to complete the extended adult interview in order to avoid biases for health estimates of elderly persons that might otherwise result. Eligible selected persons were recontacted and offered a proxy option. For 139 elderly adults, a proxy interview was completed by either a spouse/partner or adult child. A reduced questionnaire, with questions identified as appropriate for a proxy respondent, was administered. (Note: questions not administered in proxy interviews are given a value of "-2" in the data files.)

1.5 Weighting the Sample

To produce population estimates from the CHIS data, weights are applied to the sample data to compensate for the probability of selection and a variety of other factors, some

directly resulting from the design and administration of the survey. The sample is weighted to represent the non-institutionalized population for each sampling stratum and statewide. The weighting procedure used for CHIS 2005 accomplishes the following objectives:

- Compensate for differential probabilities of selection for households and persons;
- Reduce biases occurring because nonrespondents may have different characteristics than respondents;
- Adjust, to the extent possible, for undercoverage in the sampling frames and in the conduct of the survey; and
- Reduce the variance of the estimates by using auxiliary information.

As part of the weighting process, a household weight was created for all households that completed the screener interview. This household weight is the product of the "base weight" (the inverse of the probability of selection of the telephone number) and a variety of adjustment factors. The household weight is used to compute a person-level weight, which includes adjustments for the within-household sampling of persons and nonresponse. The final step is to adjust the person-level weight using a raking method so that the CHIS estimates are consistent with population control totals. Raking is an iterative procedure that forces the CHIS weights to sum to known totals from an independent data source. The procedure requires iteration to make sure all the control totals, or raking dimensions, are simultaneously satisfied within a specified tolerance.

Population control totals of the number of persons by age, race, and sex at the stratum level for CHIS 2005 were created primarily from the California Department of Finance's 2004 Population Estimates and 2005 Provisional Population Estimates. The raking procedure used 11 raking dimensions, which are combinations of demographic variables (age, sex, race, and ethnicity), geographic variables (county, Service Planning Area in Los Angeles County, and Health Region in San Diego County), household composition (presence of children and adolescents in the household), and socio-economic variables (home ownership and education). The socio-economic variables are included to reduce biases associated with excluding households without landline telephones from the sample frame. One limitation of using Department of Finance data is that it includes about 2.4 percent of the population of California who live in "group quarters" (i.e., persons living with nine or more unrelated persons). These persons were excluded from the CHIS sample and as a result, the number of persons living in group quarters was estimated and removed from the Department of Finance control totals prior to raking.

1.6 Imputation Methods

Missing values in the CHIS data files were replaced through imputation for nearly every variable. This was a massive task designed to enhance the analytic utility of the files. Westat imputed missing values for a handful of variables used in the weighting process and UCLA staff imputed values for nearly all other variables.

Two different imputation procedures were used by Westat to fill in missing responses for items essential for weighting the data. The first imputation technique was a completely random selection from the observed distribution of respondents. This method was used only for a few variables when the percentage of the items missing was very small. The second technique was hot deck imputation without replacement. The hot deck approach is probably the most commonly used method for assigning values for missing responses. With a hot deck, a value reported by a respondent for a particular item is assigned or donated to a "similar" person who did not respond to that item. The characteristics defining "similar" vary for different variables. To carry out hot deck imputation, the respondents to an item form a pool of donors, while the nonrespondents are a group of recipients. A recipient is matched to the subset pool of donors based on household and individual characteristics. A value for the recipient is then randomly imputed from one of the donors in the pool. Once a donor is used, it is removed from the pool of donors for that variable. Hot deck imputation was used to impute the same items in CHIS 2003 and CHIS 2005 (i.e., race, ethnicity, home ownership, and education).

The UCLA Center for Health Policy Research imputed missing values for nearly every variable in the data files other than those handled by Westat and some sensitive variables in which nonresponse had its own meaning. Overall, item nonresponse rates in CHIS 2005 were modest, with most variables missing valid responses for less than 2% of the sample. However, there were a few exceptions where item nonresponse rate was greater than 20%, such as household income.

The imputation process conducted by UCLA started with data editing, sometimes referred to as logical or relational imputation: for any missing value, a valid replacement value was sought based on known values of other variables of the same respondent or other sample(s) from the same household. For the remaining missing values, hierarchical sequential hot-deck imputation with donor replacement was used. This method replaces a missing value for one respondent using a valid response from another respondent with similar characteristics as defined by a set of control

variables. The control variables were ranked in order from the most to the least important. This procedure allowed control variables to be dropped if certain conditions (such as the minimum number of donors) were not met. The control variables were dropped sequentially, starting from the variable ranked least important. Once a responding case was used as a donor, it was dropped from the donor pool preventing using one donor multiple times.

Control variables always included the following: gender, age group, race/ethnicity, poverty level (based on household income), educational attainment, and region. Other control variables were also used depending on the nature of the imputed variables. Among the control variables, gender, age, race/ethnicity and regions were imputed by Westat. Household income and educational attainment were imputed first in order to impute other variables. Household income, for example, was imputed using the hot-deck method within ranges from a set of auxiliary variables such as income range and/or poverty level.

The imputation order of the other variables followed the questionnaire. After all imputation was done, logical checks and edits were performed once again to ensure consistency between the imputed and nonimputed values on a case-by-case basis.

1.7 Methodology Report Series

A series of five methodology reports is available with more detail about the methods used in CHIS 2005:

- Report 1 – Sample Design;
- Report 2 – Data Collection Methods;
- Report 3 – Data Processing Procedures;
- Report 4 – Response Rates; and
- Report 5 – Weighting and Variance Estimation.

For further information on CHIS data and the methods used in the survey, visit the California Health Interview Survey Web site at www.CHIS.ucla.edu or contact CHIS at CHIS@ucla.edu.

2. WEIGHTING ADJUSTMENTS CHIS 2005 SAMPLE WEIGHTS

This chapter introduces the concept of weighting and provides some background on the weights developed for analyzing CHIS 2005 survey data. Weighting is a process that attempts to make the estimates from the survey respondents representative of the total population that was sampled by accounting for the chances of selecting units into the sample and making adjustments for imperfections in the sample. The process begins with a base weight that is then adjusted to account for additional stages of sampling and nonresponse. The base weight is the inverse of the selection probability of the sampled unit. During the weighting, additional information from external sources is used to benchmark the weights and achieve consistency between totals from the survey data and the external sources.

Although this chapter deals with the weight adjustments for the CHIS 2005 sample, the chapter begins with the general reasons why fully adjusted weights should be used. It also describes the details, advantages, and disadvantages of weighting. In CHIS 2005, the random digit dial (RDD) sample (including supplemental samples screened for the presence of children) and surname list samples are combined and weighted together in a single file.

2.1 Weighting Approach

Weights are applied to CHIS 2005 sample data to estimate aggregate statistics at the state and county levels. In particular, sample weighting was carried out to accomplish the following objectives:

- Compensate for differential probabilities of selection and sampling rates for households and persons;
- Reduce biases occurring because nonrespondents may have different characteristics from respondents;
- Adjust, to the extent possible, for undercoverage in the sampling frames and in the conduct of the survey; and
- Reduce the variance of the estimates by using auxiliary information.

One set of weights was created for each extended interview (adult, child, and adolescent) for the combined RDD and Korean and Vietnamese surname list samples. Appendix A

shows the CHIS 2005 frame and sample sizes in addition to base weights by sampling stratum for the RDD and Korean and Vietnamese surname lists. These weights can be used to produce estimates at the state and stratum level. The final weights are the result of a series of sequential adjustments made to the base weights. The process for creating the weights for the combined RDD-list sample is described in Chapters 3, 4, 5, and 6.

As part of the weighting process for the RDD samples, a household weight is created for each household that completed the screener interview. This household weight is the base weight (the inverse of the probability of selection of the sample telephone number) adjusted for:

- Subsampling screener refusals for conversion attempt;
- Unknown residential status;
- Supplemental list sample eligibility adjustment;
- Screener interview nonresponse;
- Multiple telephone numbers;
- Supplemental sample households with no children; and
- Household poststratification.

The details of these adjustments are described in Chapter 3.

The poststratified household weight is adjusted to create a person weight for each type of extended interview. The final person weight is based on the poststratified household weight and incorporates the within-household probability of selection of the sampled person, a nonresponse adjustment, and a raking adjustment to control totals. Each adjustment corresponds to a multiplicative weighting factor applied to the weight. For the adult weights, the following factors are included:

- Probability of selection of the adult;
- Extended adult interview nonresponse adjustment; and
- Raking adjustment to person-level control totals.

The child and adolescent weights are somewhat more complex because of the method used to sample children (see *CHIS 2005 Methodology Series: Report 1 - Sample Design*). For these weights, the adjustment factors include:

- Use of child-first procedure for households in which adolescents and children are sampled at the end of the screener interview and may be interviewed before the adult interview;
- Section G adult extended interview nonresponse adjustment for households in which adolescents and children are sampled at the end of Section G of the adult interview;
- Probability of selection of the child or adolescent;
- Extended child and adolescent interview nonresponse adjustment; and
- Raking adjustment to person-level control totals.

The expressions for the weighting factors and adjustments for the person weights are given in Chapters 4, 5, and 6. The derivation of the population control totals is described in Chapter 7. The imputation process and the imputed variables to support the weighting process are described in Chapter 8. Chapter 9 discusses the methods for variance estimation for CHIS 2005.

Appendix B contains tables that show the effect of each step of the weighting process at the household and person levels. Throughout this report, we refer to specific tables and rows in Appendix B that indicate how the weights were adjusted.

2.2 Weight Adjustment Method

In an ideal survey, all the units in the inference population are eligible to be selected into the sample and all those in the sample participate in the survey. In practice, neither of these conditions occurs. Some units are not eligible for the sample (undercoverage) and some of the sampled units do not respond (nonresponse). If undercoverage and nonresponse are not addressed, then estimates from the survey will be biased. In CHIS 2005, the weights of those who respond are adjusted to represent the undercovered persons and nonrespondents. An overview of the approaches used to account for these two sources of missing data begins with adjusting for nonresponse.

Nonresponse results in biases in survey estimates when the characteristics of respondents differ from those of nonrespondents. The size of the bias depends on the magnitude of this difference and the response rate (see Groves, 1989). The purpose of adjusting for nonresponse is to reduce the bias. A weighting class adjustment (see Brick and Kalton, 1996) method is the type

of nonresponse adjustment procedure used in CHIS 2005. In this procedure, nonresponse adjustment weights are computed and applied separately by cell, where a cell is defined using characteristics known for both nonrespondents and respondents. For example, the county associated with each telephone number is known, even if there are some misclassifications in the assignment. Thus, county can be used to define cells, and weighting adjustments can be computed separately for each of these cells. The more similar either response patterns or survey characteristics are within the cells, the larger the bias reduction in the adjustment.

The drawback to nonresponse adjustment is that it increases the variability of the weights and increases the sampling variance of the estimates (Kish, 1992). A nonresponse adjustment is beneficial only when the reduction in bias more than compensates for the increase in variance. When the cells contain sufficient cases and the adjustment factors do not become inordinately large, the effect on variances is often modest. Large adjustment factors usually occur in cells with small numbers of respondents. To avoid this situation, cells with few cases are "collapsed" or combined to form a new cell with a larger number of cases.

The operational objective for nonresponse adjustment in CHIS 2005 was to define adjustment cells for which response rates vary considerably and to avoid cells with either a small number of cases or a large adjustment factor. Since county-level estimates are important, the county was nearly always included in the definition of the cells. Oh and Scheuren (1983) discuss some of the statistical features associated with making these adjustments.

As noted above, nonresponse adjustment classes can be formed only if data are available for both responding and nonresponding units. Since the nonresponse adjustment is done for each stage of data collection, the data available for forming cells are different for each stage. For screening interviews, the nonresponse unit is a household (or more accurately a telephone number), and data must be available for all households. For extended interviews, the nonresponse adjustment is done by type of person (adult, child, or adolescent). At this level, data from the screening interview can be used to define cells.

The approach to adjusting for undercoverage is somewhat different from that for nonresponse because uncovered units or persons were never eligible to be sampled. The undercoverage adjustment procedure uses data from external sources (control totals) in a process called poststratification (Holt and Smith, 1979). The primary objective of poststratification is to dampen potential biases arising from a combination of response errors, sampling frame undercoverage, and nonresponse. A secondary objective is to reduce sampling errors, which is

important because CHIS 2005 sample sizes within counties are fairly modest for some subclasses. In general, the sample is poststratified to as many independent figures as possible, subject to some constraints. In this discussion we use poststratification loosely and intend it to include raking, a form of multidimensional poststratification (see Brackstone and Rao, 1979). In CHIS 2005, the control totals are mainly derived from the 2004 California Department of Finance Population Estimates and 2005 California Department of Finance Provisional Population Estimates (State of California, Department of Finance, 2006, 2006b), the 2004 American Community Survey (U.S. Census Bureau, 2003), and the Census 2000 Summary File 1 for California published by the U.S. Census Bureau (U.S. Census Bureau, 2001). Details of the creation of the control totals at the person level are described in Chapter 7.

3. HOUSEHOLD WEIGHTING

The first step in the weighting process for CHIS 2005 is creating a household weight for each completed screener interview. The household weight is not used for analytical purposes because the only data captured at the household level in the screener interview are for sampling purposes. However, the household weight is a key element for the computation of the person weights (i.e., adult, child, and adolescent).

This chapter is divided into nine sections, each describing the steps involved in creating the household weights. The first section reviews the creation of base weights. Subsequent sections describe the adjustments made to the base weights. These adjustments account for refusal conversion subsampling, unknown eligibility residential status, supplemental list sample eligibility, unknown presence of children in the household, screener nonresponse, households with multiple telephone numbers, and subsampling in child supplemental samples. The final section of the chapter describes how the household weights are poststratified to control totals of the number of households in California.

Knowledge of the sampling methods used in CHIS 2005 is an essential ingredient to understanding the weighting procedures. We assume anyone interested in the weighting procedures is already familiar with the sampling approach – details in *CHIS 2005 Methodology Series: Report 1 - Sample Design*.

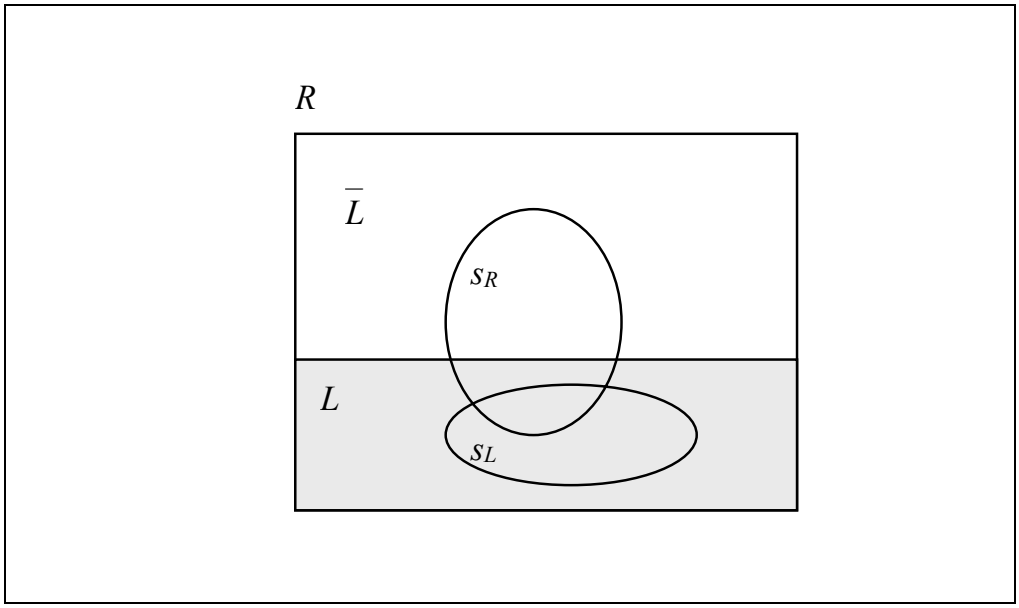
3.1 Base Weights

Each telephone number in the sample is assigned a base weight. Base weights are computed as the inverse of the probability of selection of the telephone number. In CHIS 2005, telephone numbers were drawn from four frames (RDD frame and Korean, Vietnamese, and Korean-Vietnamese [1] surname list frames); therefore, the base weights reflect the multiple probabilities of selection of telephone numbers from these different frames.

Figure 3-1 shows the relationship between the RDD frame and a single surname list frame (i.e., Korean, Vietnamese or Korean-Vietnamese list) for a single sampling stratum. The

[1] A separate frame was created for surnames that could not be readily identified as of Korean or Vietnamese origin. Because this frame contains surnames that are associated with both ethnic groups, this frame is referred to as the Korean-Vietnamese surname list.

figure also shows the relationship of the type of samples drawn from each frame. In order to create the household base weights, we consider all telephone households in California as either being on the supplemental list only (L) or as only being eligible for sampling from the RDD sample ($\overline{L}$) as shown in Figure 3-1. The relationships are discussed in detail below.



* The figure is not drawn to scale. The sizes of the list frame (L) and list and RDD samples (s_L and s_R) are smaller than shown in the figure.

Figure 3-1. Relationship between the RDD frame (R), RDD sample (s_R) sample, List frame (L) and list sample (s_L) for a single stratum

The notation in the figure follows:

R	the RDD frame containing all telephone numbers;
L	the list frame (i.e., surnames);
$\overline{L}$	all telephone numbers not found on the list. We assume that all the numbers in the list are found in R, and $R = L \cup \overline{L}$;
s_R	the simple random sample drawn from the frame R; and
s_L	the simple random sample drawn from the frame L.

Define the following:

N_R	the number of telephone numbers in the frame R;
N_L	the number of telephone numbers in the frame L;
n_R	the sample size (number of telephone numbers) in s_R; and
n_L	the sample size (number of telephone numbers) in s_L.

Notice that the RDD sample s_R can be separated into two parts: s_{RL}, the portion of s_R that is found in the list (L) and $s_{R\bar{L}}$, the portion of s_R that is not found in the list $(\bar{L})$. The sample sizes for each portion are n_{RL} and $n_{R\bar{L}}$, respectively, and $n_R = n_{RL} + n_{R\bar{L}}$.

Consider L and $\bar{L}$ as two separate strata within the frame R. Since s_R is a simple random sample within R, the sample $s_{R\bar{L}}$ can be viewed as a simple random sample of size $n_{R\bar{L}}$ drawn from the $N_{\bar{L}}$ elements from stratum $\bar{L}$. Similarly, the sample s_{RL} can be viewed as a simple random sample of size n_{RL} drawn from the N_L elements from stratum L. In stratum L, there is a second sample s_L (the list sample). Since both samples s_L and s_{RL} are simple random samples, we can view them as a single sample of size $n_{RL} + n_L$ drawn from the N_L elements from stratum L. Notice that s_{RL} and s_L are not necessarily mutually exclusive; i.e., s_{RL} and s_L may contain duplicate telephone numbers. These numbers were removed from s_L during the sample selection.

The base weights can be expressed as follows:

- For sampled records that could only be sampled from the RDD (RDD numbers not found in the list L):

$$HHBSW_{\bar{L}i} = \frac{N_{\bar{L}}}{n_{R\bar{L}}}$$

- For sampled records from the list and sampled records from the RDD that are found in the list L (duplicate telephone numbers were eliminated from the list):

$$HHBSW_{Li} = \frac{N_L}{n_{RL} + n_L}$$

Creating these weights required to classify every telephone number by whether or not it was on the list irrespective of how it was sampled. It is easy to show that the resulting weights are composite weights derived by averaging the RDD and list samples using a composite factor proportional to the sample sizes. Thus, this base weight produces an unbiased estimate in the traditional design-based framework.

The total number of telephone numbers in the RDD frame and list frames (N_R and N_L) are computed separately. The RDD sample was drawn using a list-assisted approach from a stratified frame of 100 banks[2] with at least one listed telephone number in the state of

[2] A bank is defined as 100 consecutive telephone numbers with the same first eight digits including area code.

California. Using this approach, a bank is drawn from the frame and two digits are randomly generated to complete the sampled telephone number. Therefore, the total number of telephone numbers in the RDD frame in stratum *h*, N_{Rh}, is computed as

$$N_{Rh} = 100 \cdot NBANKS_h,$$

where $NBANKS_h$ is the number of 1+ banks in March of 2005 in stratum *h*. A "1+" bank is defined as a 100 bank with at least one working telephone number.

Records on the list frames were assigned to RDD sampling strata by linking telephone exchanges to the counties in the same way as for the RDD sample. The list size by stratum (N_{Lh}) is the number of records in the list assigned to stratum *h*.

As described in *CHIS 2005 Methodology Series: Report 1 - Sample Design,* the RDD sample was drawn from strata defined as counties or groups of counties except for Los Angeles, San Diego, Orange, and Santa Clara. In Los Angeles County, 13 subsampling strata were created by the combination of areas with high/low concentration of Koreans and Vietnamese and eight Special Planning Areas (SPAs). In San Diego County, eight substrata were created by the combination of six health regions in combination with high/low concentration of Koreans and Vietnamese. Two substrata based on the concentration of Koreans and Vietnamese were created for Orange and Santa Clara Counties. The definition of the sampling strata and substrata, in addition to the number of telephone numbers in the frame, the number of sample cases, and base weights by frame type (RDD, Korean, Vietnamese, and Korean-Vietnamese lists), is shown in Appendix A, Table A-1. Table B-1 in Appendix B (rows 1.1 to 1.3) lists the sample counts, sums of base weights, and coefficient of variation by sampling stratum.

3.2 Refusal Subsampling Adjustment

The base weights were adjusted to reflect the differential refusal conversion efforts made during data collection. Subsampling of refusals was one of the techniques used in CHIS 2005 for increasing the efficiency of data collection (Brick et al., 2005). In CHIS 2005, as well as in previous cycles of CHIS, refusal cases comprised the majority of screener nonresponse and substantial effort was required to gain their cooperation. In this procedure a larger sample of telephone numbers than would otherwise be selected is drawn in the first phase. Each number in this first-phase sample is randomly assigned a flag during in second-phase subsample. When

refusals are encountered at the screening stage of data collection, only numbers in the flagged in subsample are eligible for refusal conversion follow-up. The numbers subsampled for refusal follow-up are generally fielded first so that any refusal cases can be worked completely. That is, all of the appropriate scheduling procedures (including holding periods for refusal cases) may be fully implemented and more resources devoted to completing extended interviews. The principles of subsampling for refusal conversion are well established (see Hansen and Hurwitz, 1946, and Elliott, Little, and Lewitzky, 2000). The method is used in other surveys, including the American Community Survey conducted by the U.S. Census Bureau.

The rationale for refusal subsampling depends on two observations: refusal cases comprise the majority of screener nonresponse in CHIS (as in most RDD surveys); and substantial effort is expended to gain cooperation in households in which a member refuses to participate in the study at the screener level. The subsampling of refusals shifts some resources from the less productive, labor-intensive task of refusal conversion to the more productive task of completing extended interviews. The subsampling rates are computed trying to balance the data collection costs and precision of the estimates.

The main disadvantage of this method is that a weighting adjustment is required to account for the subsampling. Those cases that refuse and are subsampled are weighted to represent themselves and the cases that refuse and are not subsampled. This weighting decreases the precision of the survey estimates, but only very slightly. Another disadvantage is that response rates should be weighted. Unweighted rates should not be used to assess response patterns because they do not reflect the subsampling of refusal conversion cases, and unweighted response rates in CHIS cannot be compared with other surveys.

In CHIS 2005, refusal subsampling was only at the screener level; there was no subsampling of extended interview refusals. A flag was randomly assigned to 60 percent of the telephone numbers from the RDD sample during sample selection. In other words, refusal subsampling procedures were used only in cases from the main RDD samples; weights for child supplemental and surname list sample cases were not adjusted for refusal conversion.

In order to adjust the weights for screener interview refusal subsampling, telephone numbers were classified into screener refusal groups using their refusal status (i.e., whether the respondent ever refused) and the value of the refusal conversion flag as shown in Table 3-1:

Table 3-1. Screener refusal groups for RDD sample

Screener refusal group	Respondent ever refused screener interview?	Refusal Subsampling Flag	Description
NRef	No	N/A	Households where respondent did not refuse the screener interview (includes complete and noncomplete interviews)
RefC	Yes	Yes	Households where respondent refused the screener interview and refusal conversion procedures were used
RefNC	Yes	No	Households where respondent refused the screener interview and refusal conversion procedures were not used

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The refusal subsampling adjusted weight, $HHA1W_i$, is:

$$HHA1W_i = HHA1F_i * HHBSW_i$$

where $HHA1F_i$ is the refusal subsampling adjustment factor computed as:

$$HHA1F_c = \begin{cases} \dfrac{\sum_{i \in RefC, RefNC} HHBSW_i \delta_i(c)}{\sum_{i \in RefC} HHBSW_i \delta_i(c)} & If\ i \in RefC \\ 0 & If\ i \in RefNC \\ 1 & If\ i \in NRef \end{cases},$$

where the groups *RefC*, *RefNC*, and *NRef* are defined in Table 3-1, $HHBSW_i$ is the base weight, and $\delta_i(c)$ is 1 if the number is in sampling stratum *c* and is zero otherwise.

Towards the end of the field period, additional telephone numbers sample were released in selected strata to meet the target for the number of completed interviews. Refusal conversion was not used for these additional releases. As a result, the overall refusal conversion subsampling rate for the main RDD sample was 48 percent in CHIS 2005. Table B-1 in Appendix B (rows 2.1 through 2.5) shows the sum of the weights before and after the refusal conversion subsampling adjustment.

3.3 Subsampling Cases with a Mailable Address in CHIS 2005

In CHIS 2003, the efficiency of the sample was improved by stratifying the telephone numbers by mailable status and subsampling the strata at different rates (Brick, Judkins, Montaquila, and Morganstein, 2002). In preparation for CHIS 2005, the effectiveness of oversampling mailable numbers was examined, as had been done in previous years. The optimal subsampling rate for nonmailable numbers proved to be close to 0.90 (whereas in the past we had subsampled 0.75 of the nonmailable numbers). The main reason for the difference in 2005 was that the nonmailable numbers were purged more efficiently than they had been in previous years. Since there was so little gain in efficiency for the additional complexity associated with this procedure, mailable status subsampling was not implemented, and mailable status weighting adjustment was not required in CHIS 2005.

3.4 Unknown Residential Status Adjustment

Telephone numbers with unknown residential status are those that could not be classified by residential status at the end of data collection despite being dialed many times. They are telephone numbers answered only by answering machines or some combination of answering machine and ring no answer (screener disposition code of *NM*) or never answered even by a machine (screener disposition of *NA*, ring no answer). Prior to adjusting the weights to account for telephone numbers with unknown residential status, the number of eligible residential telephone numbers among those numbers with unknown residential status was estimated. This estimate was also used in the computation of the response rates described in *CHIS 2005 Methodology Series: Report 4 - Response Rates*.

In CHIS 2005, the estimated proportion of unknown residential telephone numbers considered residential (p_{res}) was computed separately for the RDD and list samples. The value of p_{res} for the RDD sample was computed using an empirical method using newly available vendor services. This method is very similar to the methods used by Shapiro et al. (1995), but avoids the associated operational problems. A random sample of undetermined numbers was sent to a vendor to determine the connectivity status of the numbers. The vendor queried the telephone numbers through a nationwide TELCO (telephone central office) verified network with data created through various telecommunications partnerships. In contrast, actual calls were made to local telephone companies to achieve the same type of information in Shapiro et al. The result of an automated

query indicated the connectivity status of a telephone number including information such as use and type of service. The results of the query were used to compute the proportion p_{res}.

The value of p_{res} was computed using the results of a subsample of 10,000 telephone numbers with unknown residential status (numbers with a screener result code of *NA* "ring no answer" or *NM* "answering machine") queried through the connectivity network. The proportion p_{res} for the RDD sample was computed with 16 categories defined by urban status[3], mailable address status of the telephone number, and the answering machine status given by the interviewer based on the content of the machine's message (i.e., possible residential answering machine, possible nonresidential answering machine, or unknown).

Table 3-2 shows the values of p_{res} for the RDD sample computed using the vendor method. As expected, the estimated proportion of residential households is much lower for answering machines coded as "nonresidential" compared to those coded as "residential." For example, the estimated proportion of residential households in urban strata with mailable addresses and answering machines coded as possible residential is 92.2 percent, while the estimated proportion of those coded as rural mailable- possible nonresidential is 28.6 percent. The lowest percentages of residential telephone numbers are for the numbers that were not mailable and were never answered ("no machine") or had answering machine messages coded as nonresidential.

Table 3-2. Estimated residential proportion for the CHIS 2005 RDD sample

Urban status	Mail status	Answering machine code	p_{res}
Urban	Mailable	No machine	0.562
Urban	Mailable	Possible residential	0.922
Urban	Mailable	Possible nonresidential	0.644
Urban	Mailable	Unknown	0.745
Rural	Not mailable	No machine	0.217
Rural	Not mailable	Possible residential	0.687
Rural	Not mailable	Possible nonresidential	0.286
Urban	Not mailable	Unknown	0.183
Rural	Not mailable	No machine	0.217
Rural	Not mailable	Possible residential	0.687
Rural	Not mailable	Possible nonresidential	0.286
Urban	Not mailable	Unknown	0.183
Not urban	Mailable	No machine	0.661
Not urban	Mailable	Possible residential	0.924
Not urban	Mailable	Possible nonresidential	0.600
Not urban	Mailable	Unknown	0.759

[3] For this purpose, urbanicity was defined using the sampling strata with a population of over 500,000 persons. Stratum 1 (Los Angeles) through Stratum 15 (San Joaquin) in Table 1-1 are urban, while the remaining strata are rural.

Table 3-2. Estimated residential proportion for the CHIS 2005 RDD sample (continued)

Urban status	Mail status	Answering machine code	p_{res}
Not urban	Not mailable	No machine	0.292
Not urban	Not mailable	Possible residential	0.664
Not urban	Not mailable	Possible nonresidential	0.281
Not urban	Not mailable	Unknown	0.350

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Table 3-3 shows the values of p_{res} for the list samples. Since there were no differences by type of list sample, the values of p_{res} were computed combining the cases from the lists.

Table 3-3. Estimated residential proportion for CHIS 2005 list samples

Answering machine code	p_{res}
No machine	0.799
Answer machine possible residential	0.945
Answer machine possible business	0.383
Answer machine unknown	0.828

The estimated proportion of residential households among the unknown residential telephone numbers p_{res} is then used to adjust the weights for unknown residential status. The residential status adjusted weight, $HHA2W_i$, is

$$HHA2W_i = HHA2F_i * HHA1W_i,$$

where $HHA2F_i$ is the unknown residential status adjustment factor computed as:

$$HHA2F_i = \begin{cases} \dfrac{\sum_{i \in RES} HHA1W_i + \sum_{i \in UNK\_RES} p_{res} * HHA1W_i}{\sum_{i \in RES} HHA1W_i} & If\ i \in RES \\ 0 & If\ i \notin RES \end{cases},$$

where the group *RES* denotes telephone numbers identified as residential and *UNK_RES* denotes telephone numbers with unknown residential status. This adjustment is done separately for the RDD (within sampling stratum) and list samples. After this adjustment, only telephone numbers known to be households have positive weights. Table B-1 in Appendix B (rows 3.1 through 3.5) shows the sum of weights before and after making the adjustment for unknown residential status.

3.5 Supplemental List Sample Eligibility Adjustment

The supplemental surname list samples (i.e., Korean, Vietnamese, Korean-Vietnamese list samples), were screened to identify adults of Korean and Vietnamese descent. The weights for these samples were adjusted to account for households in which the ethnic origin of the adults (i.e., Korean or Vietnamese) could not be determined. This differs from the main RDD sample where every adult in the household was eligible for sampling. Telephone numbers from the surname list samples were eligible only if one or more adults in the household considered themselves of Vietnamese or Korean descent. (Question SC6A1 of the screener interview asked, "Do any of these adults who live in your household consider themselves to be Korean or Vietnamese or of Korean or Vietnamese descent?")

Households with at least one adult from one of these ethnic groups are referred to as "list-eligible" households. If a household from the supplemental sample was found to be list-eligible, then one adult from these ethnic groups was selected for the extended interview. If the household was not list-eligible (i.e., no adults of Vietnamese or Korean descent), then the screener interview was terminated and the case was coded as a list-ineligible.

Screening on list eligibility and retaining only list-eligible households in the surname supplemental samples was a simple method for increasing the number of Korean and Vietnamese extended interviews in CHIS 2005. The information on the ethnic origin of the adults was used to eliminate unnecessary interviews of adults with different ethnic origin, who were eligible in the RDD sample.

Household list eligibility could not be determined for nonresponding households in the surname list samples, and the weights had to be adjusted for unknown list eligibility. The weights of the households with unknown list eligibility were distributed between the list-eligible and ineligible households in the surname samples. The implicit assumption in this adjustment was that the proportion of list-eligible/ineligible households among the households with unknown list eligibility was the same as the observed proportion in the sample with known eligibility. The cases were classified in response groups as indicated in Table 3-4.

Table 3-4. List eligibility response groups

List eligibility response status group		Description
R_E	List-eligible	Household from the surname list sample with at least one list-eligible adult (i.e., adult of Korean or Vietnamese descent).
R_IN	List-ineligible	Household from the surname list sample without any list-eligible adult (i.e., no adults of Korean or Vietnamese descent).
R_UNK	List eligibility unknown	Household from the surname list sample where the list-eligibility of the adults could not be determined.
R_NA	List eligibility not screened	Household from the main RDD or child supplemental samples (not screened for eligible ethnicity).

The list eligibility nonresponse adjusted household weight, $HHA3W_i$, is computed as

$$HHA3W_i = HHA3F_c * HHA2W_i ,$$

where $HHA3F_c$ is the list eligibility nonresponse adjustment factor computed as

$$HHA3F_c = \begin{cases} \dfrac{\sum_{i \in R\_E,\, R\_IN,\quad R\_UNK} HHA2W_i \delta_i(c)}{\sum_{i \in R\_ER, R\_IN} HHA2W_i \delta_i(c)} & i \in R\_E, R\_IN \\ 0 & i \in R\_UNK \\ 1 & i \in R\_NA \end{cases} ,$$

where the groups *R_E*, *R_IN*, *R_UNK*, and *R_NA* are defined in Table 3-4, and $\delta_i(c)$ is 1 if the number is in list eligibility nonresponse adjustment cell *c* and is zero otherwise. The nonresponse adjustment cells corresponded to the list sample type (i.e., Korean, Vietnamese, Korean-Vietnamese samples). The RDD cases (including the child supplemental samples). Table B-1 in Appendix B (rows 4.1 through 4.5) shows the sum of weights before and after the list eligibility nonresponse adjustment.

3.6 Screener Nonresponse Adjustment

In this step, the household weight is adjusted to account for households that did not complete the screener interview. The screener nonresponse adjusted household weight, $HHA4W_i$, is:

$$HHA4W_i = HHA4F_c * HHA3W_i,$$

where $HHA4F_c$ is the screener nonresponse adjustment factor computed as

$$HHA4F_c = \begin{cases} \dfrac{\sum_{i \in SC\_R,\ SC\_NR} HHA3W_i \delta_i(c)}{\sum_{i \in SC\_R} HHA3W_i \delta_i(c)} & i \in SC\_R, \\ 0 & i \in SC\_NR \end{cases},$$

where the group *SC_R* is the set of screener respondents, and *SC_NR* is the set of screener nonrespondents, and $\delta_i(c)$ is 1 if the number is in screener nonresponse adjustment cell *c* and is zero otherwise.

The nonresponse adjustment cells were created separately for the main RDD and child supplemental samples within sampling strata using the telephone mailable status (known address, unknown address) and information on the presence of children in the household. These cells have different response rates due to the effect of the pre-notification letter sent to households with a known address and differences associated with the mailable and nonmailable groups. Nonresponse adjustment cells with less that 30 respondents were collapsed with similar cells within strata.

List-ineligible households (i.e., households with no adults of Korean or Vietnamese origin) from the surname list samples (group *R_IN* defined in the previous section) were considered as screener nonrespondents (group *SC_R*) in this adjustment. Although these cases were households with only list-ineligible adults, they still represented households with eligible adults for the RDD sample extended interview who were screened out. In contrast, households without children from the child supplemental samples that were ineligible for the child supplemental samples are respondents (group *SC_R*) and were adjusted at a later step. Table B-1 in Appendix B (rows 5.1 through 5.5) gives the sum of weights before and after the first screener nonresponse adjustment.

3.7 Multiple Telephone Adjustment

At the end of the screener interview, information about the existence of additional telephone numbers and their use in the household was collected. If more than one telephone number is used for residential purposes (not solely for business, fax or computer use, etc.), the household has a greater probability of selection because it could have been selected through any of the additional telephone numbers in the household. In such cases, the household weight is adjusted to reflect the increased probability of selection. The multiple telephone adjusted household weight, $HHA5W_i$, is computed as:

$$HHA5W_i = HHA5F_i * HHA4W_i,$$

where $HHA5F_i$ is the multiple telephone adjustment factor computed as:

$$HHA5F_i = \begin{cases} 0.5 & \text{If household } i \text{ has more than one residential telephone number} \\ 1 & \text{Otherwise} \end{cases}.$$

In this adjustment, we assume that there is at most one additional residential-use telephone number in the household. In a few cases, the same household was reached through two different telephone numbers found in the sample. For these cases, only one interview was conducted and the pair of telephone numbers was identified based on the information from the respondent. One telephone number was coded as duplicate (no interview associated with this number), and the weight of the other telephone number (with the completed interview) was not reduced by the multiple telephone factor because the household was sampled twice. Table B-1 in Appendix B (rows 6.1 and 6.5) identifies the sum of weights before and after the multiple telephone adjustment.

3.8 Supplemental Child Sample Subsampling Adjustment

In CHIS 2005, a state-level supplemental sample was used to increase the number of child interviews in the state. There was also a county-level supplemental sample used to increase the child interviews in San Diego County. To reduce costs in these supplemental samples households were screened to determine the presence of children and only those households with children under age 12 were retained. In contrast, an adult was always selected in the main RDD sample and surname list samples regardless of presence of children in the household. This screening is essentially a subsampling procedure, where the main RDD sample and the surname

samples are a subsample of the entire sample. In the entire RDD sample and surname samples households with children are always selected, but households without children are subsampled. Hence, the weights for the main RDD sample and surname samples are adjusted to account for the subsampling.

To adjust the weights, households were classified into groups using the information on the presence of children and the type of sample as shown in Table 3-5.

Table 3-5. Supplemental child sample subsampling groups

Group	Description
HH_C	Households with children in the main RDD sample, surname samples or child supplemental samples
RDD_NC	Households with no children in the main RDD sample and surname sample
CHLD_NC	Households with no children in the child supplemental samples

The child supplemental sample adjusted household weight, $HHA6W_i$, is:

$$HHA6W_i = HHA6F_c * HH5W_i,$$

where $HHA6F_c$ is the child supplemental sample adjustment factor computed as

$$HHA6F_c = \begin{cases} \dfrac{\sum_{i \in RDD\_NC, CHLD\_NC} HHA5W_i}{\sum_{i \in RDD\_NC} HHA5W_i} & i \in RDD\_NC \\ 0 & i \in CHLD\_NC \\ 1 & i \in HH\_C \end{cases},$$

where *HH_C*, *RDD_NC*, and *CHLD_NC* are defined in Table 3-5 and *c* is the indicator for the subsampling cell. The adjustment cells were defined by sampling stratum, presence of adolescents in the household and telephone mailable status (known address, unknown address). Note that this adjustment was unity for households with children because in this case the main RDD, surname samples, and child supplemental sample extended interviews were always attempted after the screener interview. Table B-1 in Appendix B (rows 7.1 through 7.5) gives the sum of weights before and after the subsampling adjustment.

3.9 Household Poststratification

The final step in weighting the screener interviews was to poststratify the household weights to household control totals as in previous CHIS cycles. The poststratification adjustment in previous cycles used cells created for households with and without a person under 18 years old by sampling stratum. The control totals were derived from the Census 2000 Summary File 1 for California. In CHIS 2005, no data source contained the total number of households by the presence of persons under 18 years old for each county in California in 2005. However, this type of adjustment was considered important, because response rates differ between households with and without a person under 18 years old and poststratifying household level weights using cells for households with and without a person under 18 years old adjusts for this difference in response rates. Consequently, the household weights were poststratified to 2000 control totals. Using older data is not expected to introduce biases because household estimates are not produced with CHIS 2005 data[4]. The household poststratification weight, $HHA7W_i$, is

$$HHA7W_i = HHA7F_k * HHA6W_i,$$

where $HHA7F_k$ is the poststratification factor for cell *k* given by

$$HHA7F_k = \frac{CNT_k}{\sum_i HHA6W_i \delta_i(k)},$$

where $\delta_i(k)$ is 1 if the number is in the poststratification cell *k* and is zero otherwise, and CNT_k is the control total for cell *k* defined by the existence of persons under 18 years old in the household.

The overall poststratification adjustment factors for the state and all sampling strata are listed in Table B-1 in Appendix B (row 8.3). Because the control totals were for 2000 rather than 2005, the magnitude of this adjustment should not be used as an indirect measure of the undercoverage of the estimate of the total number of households.

[4] Since person-level weights are benchmarked to current population totals in the last step of weighting, any person-level estimate is scaled to the appropriate population total in California as of 2005 (see more details in Section 7-3).

4. ADULT WEIGHTING

An adult final weight was created for each adult who completed the adult extended interview.[5] The initial adult weight is the product of the final household weight and the reciprocal of the probability of selecting the adult from all adults in the household. In subsequent steps, the initial adult weight is adjusted for nonresponse and raked to known control totals. To compensate for undercoverage of adults that could not be interviewed because they reside in households without a landline telephone, the raking adjustment included a dimension designed to reduce the undercoverage bias. Details on creating the adult weights follow.

4.1 Adult Initial Weight

As described in *CHIS 2005 Methodology Series: Report 1 - Sample Design*, one adult was sampled with equal probability from all adults in the household using the Rizzo method (see Rizzo et. al., 2004). The initial adult weight is the product of the final household weight and the inverse of the probability of selection of the adult. The expression for the adult initial weight, $ADA0W_j$, is

$$ADA0W_j = ADCNT_i \cdot HHA7W_i,$$

where $ADCNT_i$ is the total number of adults in household *i*, and $HHA7W_i$ is the poststratified household weight from the previous chapter.

4.2 Adult Nonresponse Adjustment

In some households the screener interview was completed but the sampled adult did not complete the extended adult interview. In addition, in few cases it was discovered during the extended interview that the sampled person was under 18 years of age and hence ineligible. To account for both sampled adults who did not complete the extended interview and for ineligible sampled persons, the adult initial weight was adjusted for extended interview nonresponse. Before

[5] Adult extended interviews are considered complete provided the adult completed through Section J on s health care utilization and access, mental health.

adjusting the weights, the extended interviews were classified into response groups as indicated in Table 4-1.

Table 4-1. Extended interview response groups

Response status group		Description
ER	Eligible respondent	Adult who completed the extended interview
IN	Ineligible	Ineligible person
UNK	Unknown eligibility	Sampled adult could not be contacted and eligibility verified for extended interview

The adult nonresponse adjusted weight, $ADA1W_i$, is

$$ADA1W_i = ADA1F_c \cdot ADA0W_i ,$$

where $ADA1F_c$ is the adult nonresponse adjustment factor given by

$$ADA1F_c = \begin{cases} \dfrac{\sum_{i \in ER,IN,UNK} ADA0W_i \cdot \delta_i(c)}{\sum_{i \in ER,IN} ADA0W_i \cdot \delta_i(c)} & i \in ER, IN \\ 0 & i \in UNK \end{cases},$$

where *ER, ENR* and *IN* are defined in Table 4-1, *c* indicates the adult extended interview nonresponse adjustment cell, and $\delta_i(c) = 1$ if the adult belongs to cell *c* and is zero otherwise.

The variables listed in Table 4-2 were considered in creating the nonresponse adjustment cells. All these variables were considered in previous cycles of CHIS except the child-first interview status. The child-first procedure made it possible for a child and/or adolescent to be interviewed prior to an adult. In such households, the response propensity of the adult was expected to differ from that in households where the child-first procedure was not used. For more on the child-first procedure, see *CHIS 2005 Methodology Series: Report 1 - Sample Design*. A nonresponse analysis showed that sex, child-first interview status, and age group were the best candidates for creating cells. The nonresponse cells were thus formed by classifying adults by these three variables. Cells with fewer than 30 respondents or with large adjustment factors were combined with adjacent cells. All the cells were created within sampling stratum. Appendix B, Table B-2 (rows 2.2 and 2.3) shows the sum of weights before and after the nonresponse adjustment. Ineligible persons were dropped following this weighting step.

Table 4-2. Variables considered for the creation of nonresponse adjustment cells for the adult weights

Variable	Levels
Sex of adult respondent	1. Male 2. Female
Child-first interview	1. Child-first procedures in affect 2. Child-first procedures not in affect
Presence of children and/or adolescents in the household at the screener level	1. Yes 2. No
Adult age group	1. 18-30 years old 2. 31-45 years old 3. 46-65 years old 4. 65 years or older
Household mail status	1. With a mailable address 2. Without a mailable address

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

4.3 Adult Trimming Factors

Before benchmarking the adult weights to the known total number of adults in California in 2005, we examined the distribution of the weights to determine if there were very large weights that could have a large effect on either the estimates or the variances of the estimates. When observations with large weights were found, the weights for these cases were reduced in a process called trimming.

As in CHIS 2001 and CHIS 2003, we computed statistics to identify influential weights that were candidates for trimming. These statistics and other variations were later studied in more detail in Liu et al. (2004). The first statistic is a function of spacing of the weights. Let $w_{(1)},\ldots,w_{(n)}$ be the order statistics for the adult weights $w_1,\ldots,w_n$ and define "spacing" z_i as the distance (difference) between a ranked weight $w_{(i)}$ and the next ranked weight $w_{(i-1)}$ (i.e., $z_i = w_{(i)} - w_{(i-1)}$). The statistic $d5\_space_i$ for a ranked $w_{(i)}$ is defined as

$$d5\_space_i = \frac{z_i}{z_{i-1} + z_{i-2} + z_{i-3} + z_{i-4} + z_{i-5}}.$$

The second statistic used computes the distance between a weight and the next largest weight relative to the size of the weight. The statistic is

$$rel\_space_i = \frac{z_i}{w_{(i)}} \times 10 .$$

We also computed a third statistic defined as

$$diff\_dist_i = distance_i - distance_{i-1},$$

where $distance_i$ is the relative distance for the weight $w_{(i)}$ computed as

$$distance_i = \frac{\left| w_{(i)} - median(\mathbf{w}) \right|}{MAD},$$

where $\mathbf{w} = (w_i, ..., w_n)^t$ and the median absolute deviation $MAD = median(|w_i - median(\mathbf{w})|)$.

The three statistics for the largest 20 weights in each stratum were examined separately. When all three statistics were greater than 1 then the case was a primary candidate for trimming. The final decision on trimming involved the inspection of the weight distribution within sampling stratum.

In CHIS 2005, the trimmed weight $TRMW_i$ is

$$TRMW_i = TFACT_i * ADA1W_i,$$

where $TFACT_i$ is the trimming factor for the sampled adult *i* given by

$$TFACT_i = \begin{cases} 1 & if\ the\ weight\ i\ is\ not\ trimmed \\ t_i & otherwise \end{cases}$$

where $0 < t_i < 1$.

Thirty nine adult records were trimmed[6]. The trimming factor, t_i, was determined as the ratio of the largest weight not to be trimmed and the weight being trimmed rounded up to the nearest hundred. This factor ranged from 0.50 to 0.93. Table B-2 (row 3.1) shows the strata with trimmed weights, and rows 3.2 and 3.3 show the sum of weights before and after trimming.

[6] The trimming was done prior to the raking adjustment; however, it was an iterative process. After the trimming and raking, the distribution of the weights was re-examined, and new decisions were made about trimming. This might have changed the decision about which weights should be trimmed or the magnitude of the trimming factor. If the decision was made, the trimmed and raked weights were discarded and new trimming and raking were undertaken. The number of trimmed weights reported here is at the completion of the overall process.

4.4 Adult Raked Weight

The next step in the adult weighting was raking the trimmed weights to population control totals to produce estimates consistent with the 2004 California Department of Finance (DOF) Population Estimates. Included in the raking adjustment is an undercoverage adjustment for adults in households without landline telephone discussed earlier. The specific control totals and the method used to create them are described in Chapter 7.

Raking is a commonly used estimation procedure in which estimates are controlled to marginal population totals. It can be thought of as a multidimensional poststratification procedure because the weights are poststratified to one set (a dimension) of control totals, then these adjusted weights are poststratified to another dimension. The procedure continues until all dimensions are adjusted. The process is then iterated until the control totals for all dimensions are simultaneously satisfied (at least within a specified tolerance). Raking is also described in more detail in Chapter 7.

The adult raked weight, $RAKEDW_i$, can be expressed as

$$RAKEDW_i = TRMW_i \cdot \prod_{k=1}^{K} RAKEDF_{k_l} ,$$

where $RAKEDF_{k_l}$ is the raking factor for dimension k, level l which adult i is in. For example, if the 4th dimension (k =4) is sex with two levels (l=1 for male and l=2 for female), then the raking factor for this dimension is $RAKEDF_{4_1}$ for the adult male. The raking factors are derived so the following relationship holds for every raking dimension k, and level l,

$$CNT_{k_l} = \sum_i \delta(k_l)_i \cdot RAKEDW_i ,$$

where CNT_{k_l} is the control total, and $\delta(k_l)_i = 1$ if the adult i is in level l of dimension k and zero otherwise. Table B-2 (rows 3.3 and 4.2) shows sum of weights before and after the raking adjustment.

4.5 Adult Benchmarked Weight

In CHIS 2005, an additional benchmarking adjustment was made to the raked weights. This adjustment was not done in the previous cycles of CHIS. As explained in Chapter 7, the weights were raked to known control totals where the source of these totals was the 2004 DOF Population Estimates. Since these estimates were for 2004 they were slightly outdated. The DOF provided 2005 Provisional Population Estimates that were more up to date (see Chapter 7 for more detail). However, the Provisional Population Estimates did not include the entire set of dimensions required. The raked weights were inflated by a stratum-level factor that accounts for any change in the 2004 stratum-level population estimates and the 2005 stratum-level provisional population estimates. The benchmarked weight, $BNCHW_i$, is given by

$$BNCHW_i = BFACT_c * RAKEDW_i$$

where $RAKEDW_i$ is the weight raked to 2004 DOF Population Estimate totals, and $BFACT_c$ is the benchmark factor within cell c, where cells are defined by sampling stratum. The benchmark factor, $BFACT_c$, is given by

$$BFACT_c = TOT_c^{2005} \big/ TOT_c^{2004}$$

where TOT_c^{2005} is the 2005 provisional population estimate in sampling stratum c and TOT_c^{2004} is the 2004 population estimate in stratum c. This benchmarking adjustment assumes that the proportion of group quarters population remained constant between 2004 and 2005. The final adult benchmarked weight is shown in Table B-2 (row 5.1). The factors for this adjustment ranged by sampling stratum between 1.00 and 1.04 (see Appendix B, Table B-2, row 5.3).

5. CHILD WEIGHTING

A final child weight was created for all completed child extended interviews. The steps for the child weighting are similar to those for adults described in the previous chapter. One exception is that an additional weighting adjustment is needed to account for sampled adults who do not respond, because children are often selected in the adult interview. A more complete discussion of this adjustment is given in Section 5.1. The format of this chapter follows that for the adult weighting, with the creation of the child initial weights and the adjustments for nonresponse, trimming, and finally raking.

5.1 Household-Level Adjustment

The main difference between the child (and adolescent) weighting procedures and those of the adults is that adults were always sampled in the screener. Children and adolescents not selected in the child-first procedures were sampled in Section G of the adult extended interview. Consequently, weights for those children and adolescents must be further adjusted to account for nonresponse at the adult interview level. This adjustment was also made in the previous cycles to all households with children and adolescents since child-first sampling was not used in those cycles. For CHIS 2005, the weights of the children and adolescents sampled during the screener were not adjusted for adult nonresponse.

Telephone numbers were classified into completion groups (SECGST) by Section G completion status and their child-first interview status as shown in Table 5-1.

Table 5-1. Section G completion groups

Section G completion group (SECGST)	Child-first interview?	Section G completed by adult?	Description
C1st	Yes	N/A	Households where the child-first interview procedures occurred
NC1stGC	No	Yes	Households where the child-first interview procedures did not occur and section G was completed
NC1stGNC	No	No	Households where the child-first interview procedures did not occur and section G was *not* completed

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

To account for adults who did not complete Section G of the adult interview (hence, no child or adolescent could be sampled), the household final weight $HHA7W_i$ was adjusted. We refer to this adjusted weight as the Section G adjusted household weight, $HHA8W_i$, and it is

$$HHA8W_i = HHA8F_c * HHA7W_i,$$

where

$$HHA8F_c = \begin{cases} \dfrac{\sum_{i \in SECGST = NC1stGC, NC1stGNC} HHA7W_i \delta_i(c)}{\sum_{i \in SECGST = NC1stGC,} HHA7W_i \delta_i(c)} & If\ i \in SECGST = NC1stGC \\ 0 & If\ i \in SECGST = NC1stGNC \\ 1 & If\ i \in SECGST = C1st \end{cases},$$

and where the section G completion groups *C1st*, *NC1stGC*, and *NC1stGNC* are defined in Table 5-1, c denotes the Section G nonresponse adjustment cell, and $\delta_i(c) = 1$ if the adult belongs to cell c and is zero otherwise. Following this adjustment, the weights were positive for all households with sampled adults who completed section G and either completed, partially completed, or did not complete the adult interview[7]. Note that this adjustment can be considered as a household adjustment in addition to the poststratified household weight.

The Section G nonresponse adjustment cells were created within sampling strata using a combination of the mailable status (known address/mailed letter, unknown address) and the presence of children and/or adolescents. The information about the presence of children and adolescents is collected during the screener interview.

5.2 Child Initial Weight

The initial child weight is the product of the adjusted household weight and the probability of sampling the child within the household. In CHIS 2005, the selection of the child was done in two steps. In the first step, one adult was randomly selected among all adults in the household. In the second step, one child was randomly selected among all the children associated with the sampled adult (i.e., the sampled adult is the parent or legal guardian of the child). If the

[7] If the adult interview was not completed, no attempt was made to complete a child/teen interview.

sampled adult did not have an associated child, then no child was sampled even if there were children present in the household. See *CHIS 2005 Methodology Series: Report 1 - Sample Design* for information on the within-household person selection process.

Since the child sampling depends on the relationships among children and adults within the household, these relationships were defined before sampling children. The probability of selection reflects the fact that the sampled child could have been selected through the spouse/partner of the sampled adult if both are the parents or legal guardians[8] of the sampled child. Accordingly, the initial child weight, $CHA0W_j$, is

$$CHA0W_j = \frac{1}{CHPROB_j} HHA8W_i$$

where $HHA8W_i$ is the section G adjusted weight, $CHPROB_j$ is the probability of selecting the j^{th} child associated with the i^{th} sampled adult and is relatively complex. If the sampled adult does not have a spouse/partner living in the household or if the spouse/partner of the sampled adult is not the parent or legal guardian of the sample child, then

$$CHPROB_j = \frac{1}{ADLTCNT} \cdot \frac{SACHMOS_j}{\sum_j SACHMOS_j}$$

where $ADLTCNT$ is the number of adults in the household and $SACHMOS_j$ is the measure of size of child *j*. The measure of size for a child is discussed in detail in *CHIS 2005 Methodology Series: Report 1 - Sample Design*, but we note here that within the same household children age 0 to 5 years have a measure of size twice that of children age 6 to 11 years. If the sample adult has a spouse/partner living in the household and the spouse/partner of the sample adult is the parent or legal guardian of the sample child then

$$CHPROB_j = \frac{1}{ADLTCNT}\left(\frac{SACHMOS_j}{\sum_j SACHMOS_j} + \frac{SACHMOS_j}{\sum_k SPCHMOS_k}\right)$$

where $ADLTCNT$ and $SACHMOS_j$ are defined as before and $SPCHMOS_k$ the measure of size child *k* associated with the spouse/partner of the sample adult. The number of sampled children and sum of the initial weights are in Table B-3 (rows 1.1 and 1.2).

[8] If the spouse/partner of the sampled adult is living in the household.

5.3 Other Child Weighting Adjustments

Adjustments were made to the child weights for extended interview nonresponse, trimming influential weights, and raking to control totals. The raking adjustment includes a dimension to account for children living in households without telephones. The following describes each procedure.

The child nonresponse adjustment is the same as the adult nonresponse adjustment described in Section 4.2, except the adjustment cells are defined differently. We initially created child nonresponse adjustment cells using three variables: household mailable status, sex of child, and age group (0-3, 4-7, and 8-11 years old) within sampling stratum. Since a majority of these cells had fewer than 30 respondents, we collapsed cells to increase the number of respondents in each cell. To do this we inspected adjustment factors separately by mailable status, sex, and age group at the state level to determine the variables with the most variable response rates. Using these results, the cells were defined by sampling strata, sex of child, and age group. Any cells still containing fewer than 30 respondents were collapsed over age group. The two strata with smallest sample sizes were collapsed across both sex and age group. Table B-3 (rows 2.1 to 2.3) shows the number of sample records and sum of weights before and after the nonresponse adjustments.

The next step was to identify and trim large child weights. The process used for trimming the adult weights was applied to the child weights. As a result of applying the procedures, we identified and trimmed a total of 39 child weights in CHIS 2005. The trimming factors range from 0.41 to 0.80. Table B-3 (rows 3.1 through 3.3) shows the distribution of trimmed weights by self-reported strata and the sum of the weights before and after applying the trimming factors.

The trimmed child weights were then raked to population control totals to produce estimates consistent with the California Department of Finance 2004 population estimates. See Chapter 7 for the specific controls used. These weights were then benchmarked to 2005 provisional population estimates. The expression for the raking and benchmarking adjustments are the same as those for adult weights described in Sections 4.4 and 4.5. Table B-3 (rows 4.1 through 4.3) shows the counts and sum of weights before and after the raking adjustments, and Table B-3 (row 5.1) shows the sum of the weights following the benchmark adjustment. The factors for this adjustment ranged by sampling stratum between 1.00 and 1.04 (see Appendix B, Table B-3, row 5.3).

6. ADOLESCENT WEIGHTING

In CHIS 2005, adolescents were sampled and responded to the interview for themselves after parental permission to conduct the interview. In this section, we describe the creation of analytic weights for the adolescent interview.

6.1 Initial Adolescent Weights

The procedures for creating the adolescent weights are the same as those for creating the child weights described in Chapter 5. As with the child weighting, the initial weights for the adolescents incorporates the probability of sampling the adult and the probability of sampling an adolescent among the adolescents associated with the sampled adult. The initial weight, $TNA0W_i$, is

$$TNA0W_i = \frac{1}{TNPROB_j} \cdot HHA8W_i$$

where $HHA8W_i$ is defined in Chapter 5, and $TNPROB_i$ is computed the same as $CHPROB_i$. However, the measure of size is unity for all teens regardless of their age. Table B-4 shows the number of adolescents sampled (row 1.1) and the sum of the initial adolescent weights (row 1.2).

6.2 Other Adolescent Weighting Adjustments

The adolescent initial weight was then adjusted for nonresponse the same way the adult and child initial weights were adjusted. Note that nonresponse for the adolescent interview includes failure to obtain permission for the interview, as well as failure to interview the adolescent once permission was obtained. Table B-4 in Appendix B shows the nonresponse adjusted adolescent weight. Initially the adolescent nonresponse adjustment cells were created using household mailable status, sex of the adolescent, and age group (12-14 and 15-17 years old) within sampling strata. We inspected response rates separately by the three variables at the state level. After reviewing these rates, we created cells using sampling stratum, sex and age group. Cells containing fewer than 30 respondents were collapsed across sex first and then across age group.

After the nonresponse adjustments, 55 influential weights were identified and trimmed, with factors ranging from 0.24 to 0.86. Table B-4 (rows 3.1 through 3.3) gives the trimmed weights by self-reported stratum and the sum of the weights before and after applying the trimming factors to the adolescent weights

In the last steps, the adolescent weights were raked to California DOF 2004 Population Estimates and then adjusted to produce estimates consistent with the California DOF 2005 Provisional Population. See Chapter 7 for details on the control totals. The expressions for the raking and benchmark adjustments are exactly the same as in the raking of the adult weights and the child weights. Table B-4 (rows 3.3 and 4.2) show the sum of weights before and after raking and (row 5.1) the sum of weights after the benchmark adjustment. The factors for this adjustment ranged by sampling stratum between 1.00 and 1.04 (see Appendix B, Table B-4, row 5.3).

7. RAKING AND CONTROL TOTALS

This chapter describes the raking procedure and the development of control totals for the CHIS 2005. The first section gives a general overview of raking and why this procedure was used in CHIS 2005 and previous cycles of CHIS. The second section describes the 11 dimensions used to rake the weights. Eight of the dimensions are defined by demographic variables and two are defined by socio-economic variables. The 11th dimension was created to reduce the bias associated with households without a landline telephone. The third section describes how the control totals were derived from the California DOF files.

7.1 Raking Procedure

Raking is an adjustment procedure in which estimates are controlled to marginal population totals. The main advantage of raking over poststratification is that raking allows the use of more auxiliary information. A limitation in poststratification is that each unit falls into only one adjustment cell and the number of respondents in a cell could be too small. With raking, the cell size is based on the distribution of each raking dimension. For example, with poststratification, only some cross-classified age/race/sex categories could be used in the adjustments, whereas, with raking, the full cross-classification is not needed, and important geographic level data such as county can be included as dimensions. Raking may be thought of as a multidimensional poststratification procedure because the weights are basically poststratified to one set (a dimension) of control totals, then these adjusted weights are poststratified to another dimension. After all dimensions are adjusted, the process is iterated until the control totals for all the dimensions are simultaneously satisfied within a specified tolerance. Raking was also used in CHIS 2001 and CHIS 2003. Below, we describe the procedure in more detail. Brackstone and Rao (1979); Deville and Särndal (1992); and Kalton and Flores Cervantes (2003) also describe raking.

For simplicity, consider two auxiliary variables (or dimensions) with C and D classes, respectively. If we cross-classify the two variables into $C*D$ cells and the sample counts in some cells are small, then it is likely that the poststratified estimates may be unstable unless the cells in the cross-tabulation are collapsed. With the 11 dimensions used in CHIS 2005, the potential collapsing would be very extensive.

An alternative approach is to rake the weights to the marginal totals of the variables. The raking adjusted estimator is design-unbiased in large samples and is very efficient in reducing the variance of the estimates if the estimates in the cross-tabulation are consistent with a model that ignores the interactions between variables. Collapsing is sometimes required with raking, but it is not as extensive as with poststratification.

The raked weights can be written as $\tilde{w}_{cd,i} = w_{cd}\hat{\alpha}_c\hat{\beta}_d$, where w_{cd} is the pre-raked weight of an observation in cell (c, d) of the cross-tabulation, $\hat{\alpha}_c$ is the effect of the first variable, and $\hat{\beta}_d$ is the effect of the second variable. Note that in this formulation there is no interaction effect; the weights are determined by the marginal distributions of the control variables. As a result, the sample sizes of the marginal distributions are the important determinants of the stability of the weighting procedure, not the cells formed by the cross-classification of the variables. Deficient cells (cells with small sample sizes) are thus defined in terms of the sample sizes of the marginal distributions, not of the cross-classified cells.

7.2 Raking Dimensions

The 11 dimensions used in CHIS 2005 are shown in Table 7-1. The first eight dimensions in Table 7-1 were created by combining demographic variables (age, sex, race, and ethnicity) and different geographic areas (city, county, region or group of counties, and state). The 9th, 10th, and 11th dimensions use additional variables. The 11th dimension was specifically created to adjust the weights for households without a landline telephone. Section 7.3 describes more details on this adjustment and the variables used to create the levels for this dimension. The raking dimensions for CHIS 2005 are similar to those used in previous administrations. There were also 11 dimensions in the 2001 and 2003 cycles of CHIS.

Table 7-1. Definitions of the dimensions used in raking

Dimension	Level	Description	Categories	
1	Stratum (collapsed where necessary)	Age groups (3) x Sex (2)	11	Under 12 years, male
			12	Under 12 years, female
			21	12 to 17 years, male
			22	12 to 17 years, female
			31	18 years or older, male
			32	18 years or older, female

See note at end of table.

Table 7-1. Definitions of the dimensions used in raking (continued)

Dimension	Level	Description		Categories
2	Stratum (collapsed where necessary)	Age groups (9)	1	Under 6 years
			2	6 to 11 years
			3	12 to 17 years
			4	18 to 24 years
			5	25 to 29 years
			6	30 to 39 years
			7	40 to 49 years
			8	50 to 64 years
			9	65 years or older
3	State	Age groups (13) x Sex (2)	11	Under 4 years, male
			12	Under 4 years, female
			21	4 to 7 years, male
			22	4 to 7 years, female
			31	8 to 11 years, male
			32	8 to 11 years, female
			41	12 to 14 years, male
			42	12 to 14 years, female
			51	15 to 17 years, male
			52	15 to 17 years, female
			61	18 to 24 years, male
			62	18 to 24 years, female
			71	25 to 30 years, male
			72	25 to 30 years, female
			81	31 to 37 years, male
			82	31 to 37 years, female
			91	38 to 45 years, male
			92	38 to 45 years, female
			101	46 to 53 years, male
			102	46 to 53 years, female
			111	54 to 64 years, male
			112	54 to 64 years, female
			121	65 to 77 years, male
			122	65 to 77 years, female
			131	78 years or older, male
			132	78 years or older, female
4	SPAs in Los Angeles Co., HRs in San Diego Co., Remainder of CA	SPAs (8), HRs (6), Remainder of CA (1)	1	SPA 1 – Antelope Valley
			2	SPA 2 – San Fernando
			3	SPA 3 – San Gabriel
			4	SPA 4 - Metro
			5	SPA 5 - West
			6	SPA 6 – South
			7	SPA 7 - East
			8	SPA 8 – South Bay
			9	HR 1 – North Coastal
			10	HR 2 – North Central
			11	HR 3 – Central
			12	HR 4 – South
			13	HR 5 - East
			14	HR 6 – North Inland
			15	Remainder of CA

See note at end of table.

Table 7-1. Definitions of the dimensions used in raking (continued)

Dimension	Level	Description	Categories	
5	Region (collapsed where necessary)	Race/ethnicity (7)		
			1	Latino
			2	Non-Latino White
			3	Non-Latino African American
			4	Non-Latino American Indian
			5	Non-Latino Asian
			6	Non-Latino Native Hawaiian
			7	Non-Latino Two or more races
6	State	Gender (2) x Race/ethnicity (7) x Age groups (4) (collapsed where necessary)	111	Male, Latino, under 12 years
			112	Male, Latino, 12 to 17 years
			113	Male, Latino, 18 to 64 years
			114	Male, Latino, 65 years or older
			121	Male, Non-Latino White, under 12 years
			122	Male, Non-Latino White, 12 to 17 years
			123	Male, Non-Latino White, 18 to 64 years
			124	Male, Non-Latino White, 65 years or older
			131	Male, Non-Latino African American, under 12 years
			132	Male, Non-Latino African American, 12 to 17 years
			133	Male, Non-Latino African American, 18 to 64 years
			134	Male, Non-Latino African American, 65 years or older
			141	Male, Non-Latino American Indian, under 12 years
			142	Male, Non-Latino American Indian, 12 to 17 years
			143	Male, Non-Latino American Indian, 18 to 64 years
			144	Male, Non-Latino American Indian, 65 years or older
			151	Male, Non-Latino Asian, under 12 years
			152	Male, Non-Latino Asian, 12 to 17 years
			153	Male, Non-Latino Asian, 18 to 64 years
			154	Male, Non-Latino Asian, 65 years or older
			161	Male, Non-Latino Native Hawaiian, under 12 years
			162	Male, Non-Latino Native Hawaiian, 12 to 17 years
			163	Male, Non-Latino Native Hawaiian, 18 to 64 years
			164	Male, Non-Latino Native Hawaiian, 65 years or older
			171	Male, Non-Latino Two or more races, under 12 years
			172	Male, Non-Latino Two or more races, 12 to 17 years
			173	Male, Non-Latino Two or more races, 18 to 64 years
			174	Male, Non-Latino Two or more races, 65 years or older
			211	Female, Latino, under 12 years
			212	Female, Latino, 12 to 17 years
			213	Female, Latino, 18 to 64 years
			214	Female, Latino, 65 years or older
			221	Female, Non-Latino White, under 12 years
			222	Female, Non-Latino White, 12 to 17 years
			223	Female, Non-Latino White, 18 to 64 years
			224	Female, Non-Latino White, 65 years or older
			231	Female, Non-Latino African American, under 12 years
			232	Female, Non-Latino African American, 12 to 17 years

See note at end of table.

Table 7-1. Definitions of the dimensions used in raking (continued)

Dimension	Level	Description	Categories	
			233	Female, Non-Latino African American, 18 to 64 years
			234	Female, Non-Latino African American, 65 years or older
			241	Female, Non-Latino American Indian, under 12 years
			242	Female, Non-Latino American Indian, 12 to 17 years
			243	Female, Non-Latino American Indian, 18 to 64 years
			244	Female, Non-Latino American Indian, 65 years or older
			251	Female, Non-Latino Asian, under 12 years
			252	Female, Non-Latino Asian, 12 to 17 years
			253	Female, Non-Latino Asian, 18 to 64 years
			254	Female, Non-Latino Asian, 65 years or older
			261	Female, Non-Latino Native Hawaiian, under 12 years
			262	Female, Non-Latino Native Hawaiian, 12 to 17 years
			263	Female, Non-Latino Native Hawaiian, 18 to 64 years
			264	Female, Non-Latino Native Hawaiian, 65 years or older
			271	Female, Non-Latino Two or more races, under 12 years
			272	Female, Non-Latino Two or more races, 12 to 17 years
			273	Female, Non-Latino Two or more races, 18 to 64 years
			274	Female, Non-Latino Two or more races, 65 years or older
7	State	Asian groups (5) x Age groups (4) (collapsed where necessary)	11	Non-Latino Chinese only, under 12 years
			12	Non-Latino Chinese only, 12 to 17 years
			13	Non-Latino Chinese only, 18 to 64 years
			14	Non-Latino Chinese only, 65 years or older
			21	Non-Latino Korean only, under 12 years
			22	Non-Latino Korean only, 12 to 17 years
			23	Non-Latino Korean only, 18 to 64 years
			24	Non-Latino Korean only, 65 years or older
			31	Non-Latino Filipino only, under 12 years
			32	Non-Latino Filipino only, 12 to 17 years
			33	Non-Latino Filipino only, 18 to 64 years
			34	Non-Latino Filipino only, 65 years or older
			41	Non-Latino Vietnamese only, under 12 years
			42	Non-Latino Vietnamese only, 12 to 17 years
			43	Non-Latino Vietnamese only, 18 to 64 years
			44	Non-Latino Vietnamese only, 65 years or older
			51	Other or non-Asian only, under 12 years
			52	Other or non-Asian only, 12 to 17 years
			53	Other or non-Asian only, 18 to 64 years
			54	Other or non-Asian only, 65 years or older

See note at end of table.

Table 7-1. Definitions of the dimensions used in raking (continued)

Dimension	Level	Description		Categories
8	Stratum (collapsed where necessary)	Race/ethnicity (3) x Age groups (4)	11	Latino, under 12 years
			12	Latino, 12 to 17 years
			13	Latino, 18 to 64 years
			14	Latino, 65 years or older
			21	Non-Latino White, under 12 years
			22	Non-Latino White, 12 to 17 years
			23	Non-Latino White, 18 to 64 years
			24	Non-Latino White, 65 years or older
			31	Non-Latino Non-White, under 12 years
			32	Non-Latino Non-White, 12 to 17 years
			33	Non-Latino Non-White, 18 to 64 years
			34	Non-Latino Non-White, 65 years or older
9	State	Education (4)	1	Not applicable (age < 18 years)
			2	Less than High School
			3	High School grad or GED recipient
			4	At least some college
10	State	# Adults in HH (3)	1	0 or 1 adult,
			2	2 adults,
			3	3 or more adults
11	State	Non-telephone		See Table 7-4

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Before raking, dimensions with levels or cells with fewer than 50 respondents were collapsed with "adjacent" cells. In dimensions 1, 2, 5, and 8 the collapsed cells were created within the geographic regions shown in Table 7-2. As Dimension 5 was defined at the region level, some cells were collapsed across regions if the regions did not contain enough respondents. Dimensions 3, 6, and 7 were defined at the state level because there were too few respondents in many of the cells at lower geographic levels. Dimensions 9, 10 and 11 were defined at the state level because the control totals needed to create these cells (education and type of household defined by number of adults in the household) were not available at the county level. When collapsing the cells, we ensured that there was at least one cell or a group of cells within self-reported stratum. In this way, the raked weights summed to the total number of persons in each stratum.

Table 7-2. Regions in California

Region	Sampling strata
Northern & Sierra Counties	Butte, Shasta, Humboldt, Lake, Mendocino, Yuba, Nevada, Sutter, Colusa-Glenn-Tehama, Del Norte-Lassen-Modoc-Plumas-Sierra-Siskiyou-Trinity, Alpine-Amador-Calaveras-Inyo-Mariposa-Mono-Tuolumne
Greater Bay Area	Santa Clara, Alameda, Contra Costa, San Francisco, San Mateo, Sonoma, Solano, Marin, Napa
Sacramento Area	Sacramento, Placer, Yolo, El Dorado
San Joaquin Valley	Fresno, Kern, San Joaquin, Stanislaus, Tulare, Merced, Kings, Madera
Central Coast	Ventura, Santa Barbara, Santa Cruz, San Luis Obispo, Monterey, San Benito
Los Angeles	Los Angeles
Other Southern California	San Diego, Orange, San Bernardino, Riverside, Imperial

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

7.3 Nonlandline Adjustments

Since CHIS 2005 was a traditional RDD sample, households without a landline telephone, including households with only cellular telephones, did not have a chance of being selected. The resulting undercoverage bias is related to the percentage of households without landline telephones and the differences in characteristics of the persons residing in households with landlines and those without. In the 2000 Census approximately 1.5 percent of households in California did not have a telephone. (It is not clear whether cell-only households are nontelephone or telephone households in this estimate.) Blumberg et al. (2006) estimated that the proportion of households nationally with no telephone service (landline or cellular) remained fairly constant at 1.7 percent in 2005. Some studies (e.g., Ford, 1998; Anderson, Nelson, and Wilson, 1998) have shown that the health characteristics of those with and without access to telephones are not different enough to cause significant noncoverage bias. Based on these findings, it is unlikely that most estimates from CHIS 2005 will have substantial bias because nontelephone households are not sampled. However, some estimates that are highly correlated to income may be subject to greater biases due to this form of undercoverage in RDD surveys.

Even though the percentage of nontelephone households has remained stable over the years, noncoverage bias has become a greater cause for concern in RDD surveys. This concern is because the estimated percentage of U.S. households without *landline* telephone service has increased dramatically in the last decade to 8.4 percent (Blumberg et al., 2006). The undercoverage problem is increasing as the rate of households with cellular telephone service only increases (see Tucker, et al., 2004). In addition, initial studies of the characteristics of persons in cell-only households indicate these persons differ from persons in households with landlines with respect to some health-related characteristics such as insurance coverage. These differences between the characteristics of households with and without landline telephones, however, may be offset as the proportion of cell-only households increases (see Blumberg et al., 2006). Prior to 2005, no special weighting adjustment was developed to account for households with only cellular phones. However, beginning with the 2005 cycle the nontelephone adjustment was focused on reducing the potential bias introduced by exclusion of cell-only households from the survey. Like the nontelephone adjustment in CHIS 2001 and CHIS 2003, this adjustment was included as an additional raking dimension at the person level. The adjustment could not be done at the household level because it required data from the adult extended interview.

The variables used to create the nontelephone adjustments were those identified by Blumberg et al. (2006) as good predictors of the prevalence of households with or without landline telephones. The goal was to create cells within which all households had a similar propensity of having a landline telephone. The control totals were derived for the same cells using the 2004 California Department of Finance (DOF) Population Estimates and the 2004 ACS-PUMS. Table 7-3 shows the definition of the 16 cells used for the nontelephone adjustment in CHIS 2005.

Table 7-3. Nontelephone adjustment cell definition for CHIS 2005

Cell	Household tenure	Age in years	Educational attainment	Number of adults in the household
1	Own	0 to 17	NA	0 or 1
2	Rent	0 to 17	NA	0 or 1
3	Own	0 to 17	NA	2 or more
4	Rent	0 to 17	NA	2 or more
5	Own	18 to 30	Up to high school	NA
6	Own	31 to 64	Up to high school	NA
7	Own	65 and older	Up to high school	NA
8	Own	18 to 30	Greater than high school	NA
9	Own	31 to 64	Greater than high school	NA
10	Own	65 and older	Greater than high school	NA
11	Rent	18 to 34	Up to high school	NA
12	Rent	35 and older	Up to high school	0 or 1
13	Rent	35 and older	Up to high school	2 or more
14	Rent	18 to 34	Greater than high school	NA
15	Rent	35 and older	Greater than high school	0 or 1
16	Rent	35 and older	Greater than high school	2 or more

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

7.4 Raking Factors

Table 7-4 shows the overall adjustment factors for the adult, child, and adolescent weights. The overall adjustment factors were computed as the ratio of the control total to the sum of weights before raking. These factors are, in some sense, a measure of the magnitude of the bias correction for estimates of totals. Since the weights were already adjusted for nonresponse, the raking factor could be used as an indirect measure of undercoverage. This interpretation is not straightforward for CHIS because the weights were adjusted to 2000 household totals (see Section 3.9). The adjustment factors also confound several factors such as reporting error and residual nonresponse error. Nevertheless, they may be used as a rough indicator of relative within-household coverage error. A factor greater than unity suggests relative undercoverage, and a factor less than unity relative overcoverage.

Table 7-4. Overall adjustment raking factors for adult, child, and adolescent interviews by sample characteristics

Characteristic	Adult	Child	Adolescent
Total	1.134	1.077	1.058
Sex			
Male	1.175	1.073	1.051
Female	1.096	1.081	1.065
Age group			
Under 5 years		1.114	
6 – 11 years		1.042	
12 – 17 years			1.058
18-24 years	1.429		
25-29 years	1.436		
30-39 years	1.325		
40-49 years	1.097		
50-64 years	0.928		
65 years and over	0.973		
Race/Ethnicity			
Latino	1.386	1.142	1.235
Non-Latino			
White alone	1.001	0.955	0.902
African American alone	1.328	1.500	1.070
American Indian/Alaska Native alone	1.085	1.341	1.163
Asian alone	1.256	1.192	1.326
Native Hawaiian and Other Pacific Islander alone	1.513	2.390	2.895
Two or more races	0.596	0.780	0.553
Non-Latino Asian ethnic groups			
Chinese only	1.062	1.079	0.884
Korean only	0.850	0.756	0.842
Filipino only	1.975	1.936	2.011
Vietnamese only	1.135	1.072	1.511
Educational Attainment			
Not applicable (age < 18 years)		1.077	1.058
Less than High School,	1.412		
High School grad or GED recipient,	1.247		
Some college	1.057		
College degree or above	1.013		

Table 7-4. Overall adjustment raking factors for adult, child, and adolescent interviews by sample characteristics (continued)

Characteristic	Adult	Child	Adolescent
Household Tenure [a]			
Owner	1.090	1.002	0.994
Renter	1.210	1.175	1.186
Number of adults in the household [a]			
One	1.242	1.297	1.136
Two	1.071	1.025	1.020
Three or more	1.293	1.054	1.160
Number of children in the household [a]			
None	1.107		1.040
One	1.204	1.110	1.080
Two or more	1.248	1.064	1.090
Number of adolescents in the household [b]			
None	1.134	1.076	
One	1.118	1.079	1.060
Two or more	1.161	1.085	1.056

[a] Person level estimate by type of household
Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Table 7-4 shows that the adult, child, and adolescent undercoverage is minimal. For adults, the adjustment factor for males is larger than for females, which is common in household surveys and was the case in CHIS 2001 and CHIS 2003. The factors also indicate a potential for undercoverage for younger adults.

One large overall adjustment factor is for persons who self-reported as having less than a high school education (1.677). The factors for the non-Latino Asian, non-Latino African American and non-Latino Native Hawaiian and Other Pacific Islander groups are also all larger than unity suggesting potential undercoverage. Other factors worth noting include those for persons who self-reported as being two or more races. The factors for these race groups are much less than unity, suggesting the CHIS 2005 estimates of persons of two or more races before raking are much higher than the corresponding DOF 2005 totals. It is likely that the way the race question is asked in CHIS (prompting for "any other race?") encourages reporting of multiple races as compared to the Census methodology. Note that the race-ethnicity overall adjustment factors were computed using the DOF estimates that follow the OMB definition where the number of persons who reported two races or more is smaller than if the Census definition were used.

7.5 Sources Used to Produce the Control Totals for CHIS 2005

In all cycles of CHIS considerable thought was given to the choice of data for the primary source of the control totals. In CHIS 2001, Census 2000 data were originally used because those data were recently compiled. However, as the cycles of CHIS move away from 2000, census data do not reflect the current population as well. During the CHIS 2003 cycle, several sources for control totals were examined, including Census 2000 files (U.S. Census Bureau, 2001), the 2002 American Community Survey (ACS) (U.S. Census Bureau, 2003), and the 2003 California DOF Population Projections (State of California, Department of Finance, 2004). The DOF projections were settled on as the primary source for control totals in 2003 with the hope that they could also be used for future cycles of CHIS. The DOF provides Population Projections at the county level by race, ethnicity, gender and single age for each year from 2000 to 2050. The DOF also provides Population Estimates (State of California, Department of Finance, 2006a). The estimates are updated projections based on current birth and death data. The difference between the DOF projections and estimates is that the former are produced prior to the projected year and the latter after the estimated year. Therefore, the distributions of the DOF Population Estimates are more representative of the population. Based on discussions with UCLA, the 2004 California DOF Population Estimates were used as the main source for the demographic control totals, (i.e., raking dimensions defined by gender, race, ethnicity, age, and stratum) for CHIS 2005. As in previous cycles of CHIS, the population totals had to be adjusted to remove the population living in group quarters who are not included in the survey, and household characteristics needed to be estimated. The 2000 Census files were used to compute the proportion of persons living in group quarters. The 2004 ACS files were also used as a source for educational attainment, household tenure, and household composition.

After raking to the 2004 California DOF Population Estimates, the weights were benchmarked further to 2005 California DOF Provisional Population Estimates (State of California, Department of Finance, 2006b). The 2005 provisional estimates are more up to date, but they are only available by county and not by sex, age, race, and ethnicity. The following describes the control total sources in more detail and how they were used to create raking dimension control totals for CHIS 2005.

7.5.1 California Department of Finance Population Estimate

The main source used for creating raking dimensions is the population estimates published by the California DOF. The estimates are an update of the DOF population projections that are provided at the county level by race, ethnicity, gender and single age for each year. The DOF population projections used the 2000 Census counts not adjusted for the Census 2000 undercount as the baseline. A baseline cohort-component method was used to project population estimates based on fertility/mortality rates and life expectancy by different race-ethnic groups and age cohorts. Special populations (prisons, colleges, and military installations) that have very different demographic and behavioral characteristics from the rest of the population were removed from the baseline and projected separately. However, the DOF held most of the special populations only at the year 2000 level. This factor played an important role in the assumptions made when removing the population living in group quarters from the control totals in CHIS 2005.

The population projections and population estimates are updated yearly to account for observed administrative data such as migration and recorded births and deaths. The DOF population estimates are calculated using the Driver License Address Change method. This method estimates three population age groups (under 18, 18 through 64, and 65 and over) relying on distinct data sources. The county proportions are then estimated by averaging the estimates of three separate methods (see State of California, Department of Finance, 2006b for more detail). The DOF population estimate files[9] contain counts for each county in California by year of age (101)×sex(2)×race(6)×ethnicity(2) as defined in Table 7-5. Because the population estimates provide more accurate totals compared to the population projections, the weights were raked to the 2004 estimates and then benchmarked to the 2005 provisional estimates using a simple factor.

[9] The file that included race group breakdowns for those of Latino ethnicity was provided upon request by the California DOF. The public use file only contains race group breakdowns for non-Latinos.

Table 7-5. Definition of counts available in the 2005 California DOF population estimate files*

Variable	Available counts
Age groups (101)	Age 0
	Age 1
	Age 100 or more
Sex (2)	Male
	Female
Race (6)	White alone
	African American alone
	American Indian/Alaska Native alone
	Asian alone
	Native Hawaiian and Other Pacific Islander alone
	Two or more races
Ethnicity (2)	Latino
	Non Latino

* Available at the county level
Source: State of California, Department of Finance.

The main disadvantage of the DOF estimates is the race categorization. The DOF population estimates follows the U.S. Office of Management and Budget (OMB) race definition known as "modified" race with no separate population counts for "other" race. The DOF estimates comply with the OMB 1997 revised standards for collection, tabulation, and presentation of federal data on race and ethnicity (Office of Management and Budget, 1997). The revised OMB standards identify only five main racial categories (White, Black or African American, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander) and combinations of these categories. In CHIS, respondents who could not identify themselves as any of the five OMB race categories could answer with a sixth category, "some other race," consistent with the 2000 Census data collection method. Recoding of "other race" for CHIS 2005 largely followed Census procedures (see *CHIS 2005 Methodology Series: Report 3 - Data Preparation*).

In order to use the DOF estimates, any sampled person who reported other race (alone or in combination with another race) had to be recoded into the OMB categories. This would have required the imputation of an OMB race category for 5,909 persons (10.1 percent of the sample) who self-reported "other race" only in the CHIS 2005 sample. As an alternative, a variable that combined ethnicity with the OMB race that reduced the number of imputations was proposed and approved by UCLA. The recoding includes an additional level that arranges Latinos of any race into one group as shown in Table 7-6. Because most respondents who reported other race only were Latinos, the number of imputed records was reduced significantly to 74 persons (0.13 percent) who reported a non-Latino other race. The advantage of this additional variable is that it

matches the categories of the population projections available in the DOF files. See Section 8.4 for additional details for the creation and imputation of this variable, OMBSRREO.

Table 7-6. Description of the variable using the OMB race definition

OMBSRREO	Definition
1	Latino
2	Non-Latino White alone
3	Non-Latino African American alone
4	Non-Latino Asian alone
5	Non-Latino American Indian/ Alaska Native alone
6	Non-Latino Pacific Islander alone
7	Non-Latino two or more races

The DOF population estimates include the population living in group quarters. Since the target population in CHIS 2005 excludes persons in group quarters, these persons were removed from the DOF population estimates. At the time control totals were being developed, the DOF did not have separate projections for the population living in group quarters. The Census 2000 files were used to estimate the proportion of persons in group quarters, and these proportions were applied to the DOF estimates.

7.5.2 Census 2000 Files

The Census files were used to derive the control totals for the dimension defined by SPAs in Los Angeles and Health Regions in San Diego County (dimension 4 in Table 7-1). The proportions of the total population in those areas were computed from the 2000 Census files. This assumes that the proportion in these areas with respect to the county did not change between 2000 and 2005[10]. The Los Angeles SPAs and San Diego Health Regions were both defined in terms of Census Tracts.

7.5.3 American Community Survey for California

The American Community Survey (ACS) is a nationwide survey that provides current and detailed demographic, social, economic, and housing data. It is a critical element in the Census Bureau's reengineered 2010 Census plan as it will replace the decennial census long form in future

[10] The population in group quarters was removed from these areas and the county before computing the proportions.

censuses. The 2004 ACS produces population and household estimates for a limited number of characteristics at the state level and for over 800 geographical areas excluding the group quarters populations. There is a 2004 ACS public use micro data file (ACS-PUMS) for California that provides household and population estimates at the state level. For selected counties and large communities, custom tables for a subset of estimates can be downloaded from the Census website (http://www.census.gov/acs/www/). The 2004 ACS includes population estimates for 24 CHIS 2005 strata, but not for the SPAs in Los Angeles County, or San Diego County Health Regions, and some small counties.

Although the 2004 ACS was not used as the primary source of population control totals in CHIS 2005, it was used to compute proportions by educational attainment and type of household (tenure and number of adults in the household) at the state level as these variables were not available in the DOF files. These proportions were applied to the 2005 DOF total population counts to derive the control totals for the raking dimensions defined by these characteristics (dimensions 9, 10, and 11 in Table 7-1). The proportions were calculated at the state level allowing use of the ACS data. Applying the 2004 factors assumed that there were no changes in the population proportions between 2004 and 2005[11] for these variables.

7.6 Producing the Control Totals for CHIS 2005

As with previous cycles of CHIS, the derivation of the control totals was a challenging task in 2005. It involved the selection of the sources of control totals, determining the number of dimensions, and computing the control totals. It also had an impact on the set of variables to be imputed. In CHIS 2005, there were 11 raking dimensions. Deriving the control totals for each dimension independently could lead to inconsistencies between totals across the dimensions and this would cause problems in the raking process.

To overcome these difficulties, we used a procedure developed for the CHIS 2003 cycle in which the control totals for most of the dimensions were computed simultaneously. In this approach, a file was created with totals for all the possible combinations of the levels from most of the raking dimensions in the source files. These totals were then adjusted to remove the population living in group quarters. In the final step, the file was summarized by aggregating the totals by raking dimension. Because all totals were produced from the same file, there were no

[11] The population in group quarters was removed from these areas before computing the proportions.

inconsistencies (the sum across dimensions was constant and the relationship between dimensions using the same variables such as age was fixed) among the dimensions. The details of this procedure are described in the following sections.

As the first step when computing control totals, the population living in group quarters was removed from the population counts. This is a straightforward process when counts of persons in group quarters for all variables and geographic levels are available. However, this information was not available in the DOF files. By assuming that the proportion of the population in group quarters did not change between 2000 and 2005, the Census 2000 SF1 file could be used to compute these proportions. This assumption is the same one used by the California DOF for its population projections.

In past cycles of CHIS, two problems occurred when computing the percentage of the population living in group quarters using the Census SF1 file. The first was the limited number of group quarter counts that can be produced from the SF1 file. Counts are available only by stratum (44) × age group 1 (3) × sex (2) × race (7) and by stratum (44) × age group 1 (3) × sex (2) × ethnicity (3) as defined in Table 7-7. Other counts included totals by stratum (44) and single year of age.[12] The file could not be used to produce population counts by single year of age by the cross-tabulation of race and ethnicity. The file with counts by single age was used to compute the population in group quarters for young and older children and adolescents, i.e., and stratum (44) × age group 2 (5) where age group 2 is defined in Table 7-7. As in previous cycles the assumption was made that the distribution of the population in group quarters was uniform among the age groups; for example, if the percentage of persons 65 years old in group quarters is 1.56 percent, then 1.56 percent of persons 68 years old are assumed to be in group quarters.

The second problem was that the group quarter population counts from the SF1 file are defined for the seven race categories shown in Table 7-7 and not the six OMB race groups used in the DOF file (see Table 7-5). To address the problems, we assumed that the distribution of persons in group quarters by ethnicity (Latino or non-Latino) was also the same within race. For example, if 1.42 percent of the African American population is in group quarters, then 1.42 percent of both Latino African Americans and non-Latino African Americans are assumed to be in group quarters.

[12] Census 200 totals for the population in group quarters by age single was not previously used in previous cycles. These counts were used for the first time in CHIS 2005 to more accurate control of the proportion in group quarters of two age groups for children (0 to 5 and 6 to 12 years old) and adolescents. The proportion of the population in group quarters is not uniform across the groups; for example, the proportion of adolescents in group quarters is much larger than the proportion of younger children.

Table 7-7. Definition of levels of variables for group quarters populations in the Census 2000 SF1 file

Characteristics	Available counts
Stratum (44)	Counties or combinations of multiple counties defined in CHIS 2005
Age group1 (3)	Less than 18 years old 18 to 64 years old 65 years old or older
Age group2 (5)	Less than 6 years old 6 to 11 years old 12 to 17 years old 18 to 64 years old 65 years old or older
Sex (2)	Male Female
Race (7)	White alone African American alone American Indian/Alaska Native alone Asian alone Native Hawaiian and Other Pacific Islander alone Other race alone Two or more races
Ethnicity(3)	Latino Non-Latino White alone Other

Source: U.S. Census Bureau, Census 2000.

Using this assumption, the percentage of the population not living in group quarters in 2000 was computed as follows. A file with 2000 population totals, T_{rc}^{2000}, was created by summarizing the 2000 SF1 into 22,176 cells denoted *rc*, where *r* denotes race and *c* is the cross-tabulation of stratum$(44)\times$ ethnicity$(2)\times$ age$(18)\times$ gender(2). The 18 levels of age (see Table 7-8) corresponded to the cross-tabulation of the levels of age available in the DOF data files and in the definition of the raking dimensions. An advantage of summarizing the file by the levels of *c* was the smaller size of the file (i.e., the file contains population totals by the age groups rather than single age). Note that any age group, race, or ethnicity as defined in the raking dimensions could be created by combining the *c* cells.

We defined the cells *rc* as the cross-tabulation of race and the cell *c* as follows:

$$rc = \text{race}_{\overline{OMB}}(7)\times c\,,$$

where the subscript $\overline{OMB}$ refers to the non-OMB race classification that includes a category for "some other race" available in the SF1 file as shown in Table 7-7.

Table 7-8. Age levels corresponding to the cross-tabulation of the DOF data files and the definition of the raking dimensions

Age group(i)	Description
1	0 to 3 years old
2	4 to 5
3	6 to 7
4	8 to 11
5	12 to 14
6	15 to 17
7	18 to 24
8	25
9	26 to 29
10	30
11	31 to 37
12	38 to 39
13	40 to 45
14	46 to 49
15	50 to 53
16	54 to 64
17	65 to 77
18	78 plus

7.6.1 Removing the Population Living in Group Quarters

We now review how the group quarter population was removed from the DOF files. Define $T_{rc}^{2000\,\overline{GQ}}$ as the 2000 population total that excludes the population in group quarters in cell *rc*. The totals $T_{rc}^{2000\,\overline{GQ}}$ were computed by raking the totals T_{rc}^{2000} to three control totals for the population not living in group quarters. Let $D1_{m}^{2000\,\overline{GQ}}$ be the control total for the fist raking dimension computed as

$$D1_{m}^{2000\,\overline{GQ}} = D1_{m}^{2000} - D1_{m}^{2000\,GQ},$$

where $D1_{m}^{2000}$ is the 2000 total population, $D1_{m}^{2000\,GQ}$ is the 2000 population total living in group quarters, and m is the raking cell defined as $m = strata(44) \times race_{\overline{OMB}}(7) \times age\ group\ \ 1(3) \times sex(2)$.

In the same way, let $D2_n^{2000\ \overline{GQ}}$ be the control total for the second raking dimension for cell *n*, where *n* is defined as $n = strata\,(44) \times ethnicity\,(3) \times age\ group\ 1\,(3) \times sex\,(2)$ as in the SF1. Let $D3_p^{2000\ \overline{GQ}}$ be the control total for the third raking dimension for cell *p*, where *p* is defined as $n = strata\,(44) \times age\ group\ 2\,(5)$ as in the SF1.

Note that $D1_m^{2000\ GQ}$, $D2_n^{2000\ GQ}$, and $D3_p^{2000\ GQ}$ are the 2000 population totals living in group quarters available in the SF1 file. By using raking we ensured that all totals, $T_{rc}^{2000\ \overline{GQ}}$, were consistent and they summed to the control totals.

After raking, the proportion of the 2000 population not living in group quarters in cell *rc* was computed as

$$p_{rc}^{2000\ \overline{GQ}} = \frac{T_{rc}^{2000\ \overline{GQ}}}{T_{rc}^{2000}}.$$

Assuming that the proportion of the population not living in group quarters did not change between 2000 and 2005 within cell *rc*, the proportion $p_{rc}^{2000\ \overline{GQ}}$ could be used to compute $T_{rc}^{2005\ \overline{GQ}}$ defined as the 2005 total population not living in group in cell *rc*, as

$$T_{rc}^{2005\ \overline{GQ}} = p_{rc}^{2000\ \overline{GQ}} * T_{rc}^{2005},$$

where T_{rc}^{2005} is the 2005 total population from the 2005 California DOF file in cell *rc*. However, T_{rc}^{2005} could not be computed using the DOF file due to differences in race categorization between the SF1 and the DOF projection. Instead, the 2005 population estimates, $T_{sc}^{2005\ OMB}$, were available in the DOF file for 19,008 cells (labeled *sc*) defined using the OMB race categories. The cells *sc* were defined by the cross-tabulation of $sc = race_{OMB}\,(6) \times c$, where the subscript *OMB* refers to the OMB race groups that exclude the "some other race" category as shown in Table 7-9, and *c* is defined as before.

Table 7-9. OMB race categories available in the California DOF files

$race_{OMB}(s)$	Description
1-W	OMB White alone
2-AA	OMB Black or African American alone
3-AI	OMB American Indian or Alaska Native alone
4-AS	OMB Asian alone
5-PI	OMB Pacific Islander Native Hawaiian alone
6-TM	OMB Two or more races

In order to examine the relationship between the totals T_{sc}^{OMB} and T_{rc}, consider the following summation:

$$T_c^{OMB} = \sum_s T_{sc}^{OMB} = T_{Wc}^{OMB} + T_{AAc}^{OMB} + T_{AIc}^{OMB} + T_{ASc}^{OMB} + T_{PIc}^{OMB} + T_{TMc}^{OMB} .$$

In the same way, the total population in a cell *c* can be represented by non-OMB race groups as

$$T_c = \sum_r T_{rc} = T_{Wc} + T_{AAc} + T_{AIc} + T_{ASc} + T_{PIc} + T_{Oc} + T_{TMc} .$$

The assignment of the OMB was done within cell *c*; in other words, the total population in the cell *c* stays constant. That is

$$T_c = \sum_r T_{ri} = T_c^{OMB} = \sum_s T_{sc}^{OMB} .$$

When assigning an OMB race value, persons who reported "some other race" alone were assigned one of the OMB race categories. Persons who reported two races, one being "other race," kept the OMB race category but dropped "other race." In other words, they were assigned a single OMB race. Persons who reported more than two races, one of these being "other race," were still considered as having multiple races (the "other race" removed).

In order to illustrate the reallocation, consider the Asian group (ignoring the strata, age group, sex, and ethnicity components of the cell),

$$T_{ASc}^{OMB} = T_{ASc} + p_{AS\_O}^{OMB} * T_{Oc} + p_{AS\_TM}^{OMB} * T_{TMc} ,$$

where

T_{ASc}^{OMB} is the total number of Asians (OMB definition);

T_{ASc} is the total number of Asians (non-OMB definition);

T_{TMc} is the total number of persons with two or more races (non-OMB definition);

$p_{AS\_O}^{OMB}$ is the proportion of persons with some other race alone who were coded as Asian alone when assigning the OMB definition; and

$p^{OMB}_{AS\_TM}$ is the proportion of persons with two or more races who are coded as Asian alone when assigning the OMB definition.

In other words, the OMB Asian alone population (T^{OMB}_{ASc}) is composed of the original non-OMB Asian-alone total (T_{ASc}), the portion of the population who reported "some other race" alone that is allocated to OMB Asian ($p^{OMB}_{AS\_O} * T_{Oc}$), and the population who reported non-OMB Asian-alone and "some other race." Figure 7-1 visualizes how the OMB Asian alone population is formed, where T^{OMB}_{AS}, T_{AS}, and T_{TM} are defined above and T_O is defined as the group who reported "other race" only (omitting the subscript *c* for convenience).

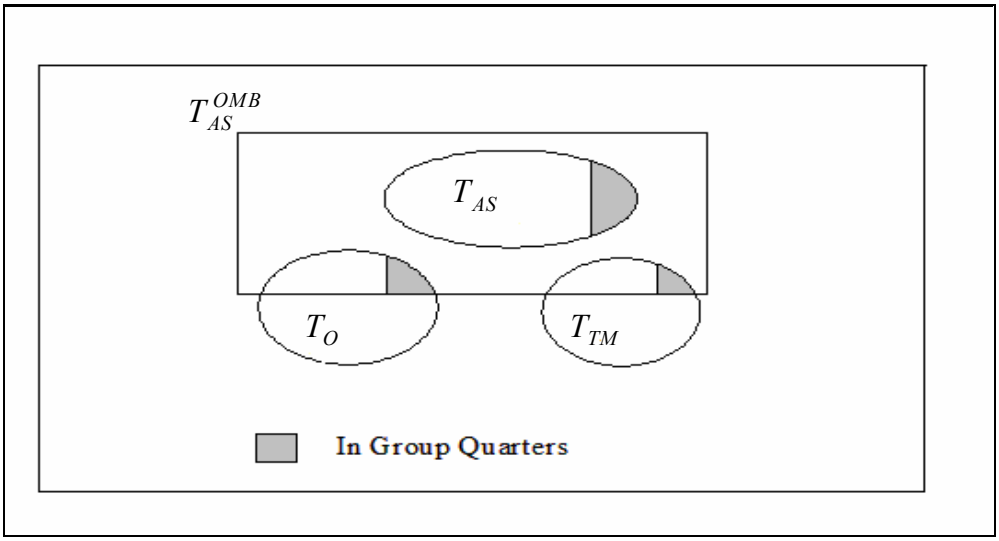


Figure 7-1. Relationship between OMB Asian alone and non-OMB groups

In this example, the proportion of the population in group quarters was known for the non-OMB Asian alone group. In order to compute the proportion of the population not in group quarters for OMB Asian alone we assumed the same proportion holds for the members that were being reclassified into the OMB race group. That is

$$\frac{T^{OMB\,\overline{GQ}}_{AS}}{T^{OMB}_{AS}} \approx \frac{T^{\overline{GQ}}_{AS}}{T_{AS}} = \frac{T^{\overline{GQ}}_{AS} + T_O{}^{\overline{GQ}} + T^{\overline{GQ}}_{TM}}{T_{AS} + T_O + T_{TM}},$$

only for $O \in AS$ and $TM \in AS$, (i.e., OMB race assignment to *AS*).

Generalizing these results to the other groups, the proportion of the population not in group quarters, $p^{\overline{GQ}}_{rc}$, can be computed as

$$p_{rc}^{\overline{GQ}} = \frac{T_{rc}^{\overline{GQ}}}{T_{rc}} \approx \frac{T_{sc}^{OMB\ \overline{GQ}}}{T_{sc}^{OMB}} = p_{sc}^{\overline{GQ}}$$

Under the assumption that the proportion of the population not living in group quarters did not change between 2000 and 2005 the proportion was computed as

$$p_{sc}^{2005\ \overline{GQ}} = p_{sc}^{2000\ \overline{GQ}} = \frac{T_{rc}^{2000\ \overline{GQ}}}{T_{rc}^{2000}}.$$

The proportion $p_{sc}^{2005\ \overline{GQ}}$ was used to compute the 2005 total population not living in group quarters in cell *sc,* $T_{sc}^{2003\ OMB\ \overline{GQ}}$, defined using the OMB race categories, as follows:

$$T_{sc}^{2005\ OMB\ \overline{GQ}} = p_{sc}^{2000\ \overline{GQ}} * T_{sc}^{2005\ OMB} = \frac{T_{rc}^{2000\ \overline{GQ}} * T_{sc}^{2005\ OMB}}{T_{rc}^{2000}}.$$

where T_{rc}^{2000} is computed using the SF1 file, $T_{sc}^{2005\ OMB}$ using the 2005 DOF file and $T_{rc}^{2000\ \overline{GQ}}$ is the 2000 population in cell *rc* not in group quarters, as defined earlier. The 2005 total population not living in group quarters in California is computed as

$$T^{2005\ \overline{GQ}} = \sum_s \sum_c T_{sc}^{2005\ OMB\ \overline{GQ}}.$$

Table 7-10 shows the total population in the 2004 DOF file and the estimated total (and percentage) of the population living in group quarters.

Table 7-10. Population in California

Type	Population	%
In group quarters	863,021	2.4
Not in group quarters	34,796,370	97.6
Total	35,659,391	100.0

7.6.2 Computing the Control Totals

The totals $T_{sc}^{2005\ OMB\ \overline{GQ}}$ were summarized in order to compute the control totals for dimensions 1, 2, 3, 5, 6, and 8. For dimension 7, defined for Asian ethnic groups, the control totals were derived using the same totals but for Asian only. Using the Census 2000 SF1 files, we

computed the percentages for the Asian groups in Table 7-11. The percentages of the Asian groups by ethnicity (Latino, non-Latino) were computed using the 2004 ACS-PUMS file. It was assumed that there were no changes in the distribution of the Asian groups between 2004 and 2005. These percentages were applied to the 2005 DOF projections.

Table 7-11. Census 2000 SF1 Asian groups

Asian Group	Description
1	Chinese alone
2	Korean alone
3	Filipino alone
4	Vietnamese alone
5	Other Asian ethnic group alone

The creation of dimension 4, defined by SPAs in Los Angeles County and Health Regions in San Diego County, used information from the Census 2000 SF1. The Los Angeles County Department of Health (LACDH) produced a listing of Census tracts by SPAs. The 2000 SF1 file was used to compute the percentages of the population in the SPAs by aggregating population counts in the Census tracts. This percentage was applied to the total 2005 DOF population total (excluding group quarters) to produce the controls for dimension 4. A similar procedure was used for San Diego County Health Regions.

For dimensions 9 (adult's education attainment), 10 (number of adults in the household), and 11 (nontelephone adjustment), the percentages of the population were computed using the 2004 ACS-PUMS and then applied to the 2005 DOF population total (excluding group quarters). The underlying assumption was that there were no changes in the distribution of the population between 2004 and 2005.

8. IMPUTATION PROCEDURES

In any household survey, both unit and item nonresponse are virtually unavoidable. We have described how weighting adjustments have been used to compensate for unit nonresponse in CHIS 2005. *CHIS 2005 Methodology Series: Report 4 – Response Rates* discusses unit nonresponse in detail. This chapter focuses on item nonresponse and the imputation for missing responses of the variables used in weighting. The imputed values were needed in the last stages of the weighting process, and only interviews that were considered completed units were subject to imputation. The percentage of missing data and consequent imputation for virtually all of these items is small.

Section 8.1 describes the imputed variables and reviews the different types of imputation techniques used to fill in the missing data. The two imputation techniques employed in CHIS 2005 are random allocation and hot-deck imputation. Sections 8.2 through 8.4 discuss the imputation process for all imputed variables separately. The last section lists the geographic location variables for CHIS 2005. UCLA derived these variables after geocoding the geographic information either collected during the interview (address of respondent, nearest street intersection, self-reported county) or attached to the sample telephone (address for numbers that were mailable or ZIP Code covered by the telephone exchange). Apart from the imputation carried out for variables used in weighting process described in this report, *CHIS 2005 Methodology Series: Report 6 – Data Editing and Imputation* introduces the imputation method used by UCLA.

8.1 Imputed Variables and Methods

Table 8-1 lists the variables imputed in CHIS 2005. The same variables were imputed in CHIS 2003. As noted above, the level of missing data is relatively small. The specific percentages of missing data are given later in the chapter. When the amount of missing data is small and assuming that the data are missing at random (i.e., the missing data have the same distribution as those with complete data within groups defined for imputation), then the bias of the estimates due to the missing data should be relatively small. The imputations may also increase the variance of the estimates, but this effect should be negligible given the low rate of missing data. A flag indicating if the response is imputed accompanies every value.

Table 8-1. Description of imputed variables

Variable name	Description	Extended interview items	Variable type
SRAGE	Self-reported age	AA2, CA3, TA2, KAA2	Demographic
SRSEX	Self-reported sex	AA3, CA1, TA3, KAA3	Demographic
SRTENR	Self-reported household tenure	AK25, KAK25	Socio-economic
SREDUC	Self-reported educational attainment	AH47, KAK47	Socio-economic
SRH	Self-reported Latino	AA4, CH1, TI1	Ethnicity
SRW	Self-reported white	AA5A_6, CH3_6, TI2_6	Race
SRAA	Self-reported African American	AA5A_5, CH3_5, TI2_5	Race
SRAS	Self-reported Asian	AA5A_4, CH3_4, TI2_4	Race
SRAI	Self-reported American Indian/ Alaska Native	AA5A_3, CH3_3, TI2_3	Race
SRPI	Self-reported Native Hawaiian and Other Pacific Islander	AA5A_1, AA5A_2, CH3_1, CH3_2, TI2_1, TI2_2	Race
SRO	Self-reported Other race	AA5A_7, CH3_7, TI2_7	Race
OMBSRREO	OMB self-reported race/ethnicity		Race/ Ethnicity
OMBSRASO	OMB self-reported non-Latino Asian group	AA5E_1- AA5E_18, TI7_1- TI 7_18, CH7_1-CH7_18	Race/ Ethnicity

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

In CHIS 2005 and previous cycles of CHIS, random allocation and hot-deck imputation were used to fill in the missing responses. The first imputation technique is a random selection from the observed distribution. This method is used only when the item missing rate is very small.

The second technique, hot-deck imputation, was used to impute race and ethnicity (including the OMB race-ethnicity variables) as well as household tenure and educational attainment in the 2003 and 2005 cycles. The hot-deck approach is probably the most commonly used method for assigning values for missing responses in large-scale household surveys (Sande, 1983; Ford, 1983). With a hot deck, a value reported by a respondent for a particular item is assigned or donated to a "similar" person who did not respond to that item. In order to carry out hot-deck imputation for CHIS 2005, the respondents to an item form a pool of donors while the

nonrespondents are a group of recipients. A recipient is matched to the subset pool of donors with the same characteristics. The recipient is then assigned a randomly imputed value from one of the donors in the pool. Once a donor is used, it is removed from the donor pool.

8.2 Self-Reported Sex and Age

The percentage of cases where either sex or age was missing in CHIS 2005 is very small across all types of extended interviews (adult, child, and adolescent). Table 8-2 summarizes the number of cases that were imputed for sex and age in CHIS 2005. The sex of only six children and one adolescent were imputed in CHIS 2005; no adults were missing self-reported sex. For the children, the missing data for sex was imputed randomly. A random number was generated for the six missing values. Four children were imputed as male and two female. The sex of the adolescent was imputed using roster information obtained during the screener interview.

Age was imputed in 128 cases in CHIS 2005. A hierarchical process was followed to impute the 114 missing self-reported age values for adults. The process used the values for self-reported age (question AA2 on the adult interview), the self-reported adult age range (question AA2A on the adult interview) asked when the adult refused to provide a specific age, the proxy-reported adult age collected during the child-first interview (question KAA2) if available, and the adult age collected during the screener interview (question ADULTAGE on the screener interview).

Table 8-2. Number and percentage of completed interviews with missing self-reported sex and age

Person type	Number completed	Number missing sex	% missing sex	Number missing age	% missing age
Adult	43,020	0	0.00	114	0.26
Child	11,358	6	0.05	14	0.12
Adolescent	4,029	1	0.02	0	0.00
Total	58,407	7	0.01	128	0.22

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The missing age for adults was imputed as follows. First if an adult had a missing self-reported age, we checked whether the adult age was collected from a proxy adult in the child-first interview. If age was reported, this age was assigned to the sampled adult. If an age was not reported in the child-first interview, the screener age for the sampled adult was checked. If the screener age was within an age range given by the sampled adult, then screener age was used as the

imputed age. If the age range was not reported, then the screener age was used. However, if the screener age was outside the reported age range, then age was randomly selected using the distribution of self-reported age within the reported age range. If no age was collected during the screener but an age range was reported, then age was randomly imputed from the distribution of self-reported age within the reported age range. If no information on age was available from any source, then age was randomly imputed using the distribution of self-reported age of all adult respondents. As an example, assume an adult respondent did not report an age but reported an age range of 40 to 44. Assume also that the proxy reported adult age in the child-first interview was 38 and the age collected in the screener interview when the adult was sampled was 38. This situation could result if the proxy misreported the sampled adult age in both the screening interview and the child-first interview. For this case, the adult age would be imputed using the distribution of the self-reported age of adults age 40 to 44. Assume that the distribution of adult age is such that 41 percent of sampled adults were age 41 or less and 62 percent were age 42 or less. If the random number assigned to the adult had a value of 0.44 then the adult's imputed age would be 42 years old.

8.3 Household Tenure and Educational Attainment

Household tenure and the adult respondent's educational attainment were used to create the raking dimensions 9 and 11. Household tenure had 387 missing responses (0.90 percent). and educational attainment had 129 missing (0.30 percent).

Hot-deck imputation was used to impute missing values for these two variables. The search algorithm CHAID (Kass, 1980) was used to create the hot-deck cells using the variables available for both donors and recipients that are found to be good predictors. A donor was then randomly drawn from the cell and its value for the variable being imputed was assigned to the recipient. Table 8-3 shows the variables considered in CHAID to create the hot-deck cells for educational attainment and household tenure. Table 8-4 shows the distribution of the imputed cases. When calculating the percentages, the denominator for age group is the number of adults in the given age group, and for educational attainment the denominator is all adults.

Table 8-3. Variables used to define hot-deck cells for the imputation of education attainment and household tenure

Variable Name	Description
Educational Attainment	
SRSEX	Self-reported sex
SRRACE_O	Self-reported race
SRH	Self-reported ethnicity
SRAGE	Self-reported age
ADLTFLG	Number of adults in the household
CHLDFLG	Children present in the household
TEENFLG	Adolescents present in the household
POVERTY	Poverty
P_GRAD	Percent college graduates in exchange
P_OWN	Percent home owners in the exchange
P_BLACK	Percent African Americans in the exchange
P_HISP	Percent Latinos in the exchange
CREGION	California Regions
Household Tenure	
ADLTFLG	Number of adults in the household
CHLDFLG	Children present in the household
TEENFLG	Teens present in the household
P_GRAD	Percent college graduates in exchange
P_BLACK	Percent African Americans in the exchange
P_HISP	Percent Latinos in the exchange
P_OWN	Percent home owners in the exchange
POVERTY	Poverty
CREGION	California Regions

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Table 8-4. Counts and percentages of imputed self-reported education attainment and household tenure

	Adult interviews	
	Count	Percentage
Self-reported Education Attainment		
Under 18 years of age	NA	
Less than HS, 18 years of age or older	30	0.07
High School (or equivalent), 18 years of age or older	29	0.07
Some college, 18 years of age or older	34	0.08
BS and above, 18 years of age or older	36	0.08
Total	129	0.30
Self-reported Household Tenure		
Owner	227	0.53
Renter	160	0.37
Total	387	0.09

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

8.4 Self-Reported Race and Ethnicity

As described in Chapter 7, the person weights were raked to control totals from the 2004 California DOF Population Estimates. The California DOF complies with the OMB 1997 revised standards for collection, tabulation, and presentation of federal data on race and ethnicity. The revised OMB standards identify only five main racial categories and combinations of these categories. The main categories are White, Black or African American, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander. Census 2000 allowed a sixth race category ("some other race") for respondents who could not identify with any of the five OMB race categories. Because all public release files of the Census 2000 include six race categories, the Census Bureau released a special file called Modified Race Data Summary file (MRDSF) with 2000 population counts by the five OMB race categories (U.S. Census Bureau, 2002). To produce this file the Census Bureau implemented special procedures to assign and impute an OMB race to those who reported "some other race." The California DOF Estimates used the 2000 MRDSF as the baseline for the time series; as a result, the DOF Estimates include only counts by the five OMB racial categories by county.

Following a procedure similar to the Census 2000, respondents who could not identify themselves as any of the five OMB race categories could answer "some other race" in CHIS. In order to use the DOF estimates as control totals, any sampled person who reported "some other race" (alone or in combination) had to be recoded into one or more of the OMB categories. OMB race was missing for 5,909 persons (10.1 percent) in CHIS 2005. After examining the procedures used by the Census to assign an OMB race, we determined that the assignment of OMB race could not be implemented using the available variables in CHIS 2005 as in Census 2000, because the number of CHIS cases in the geographic area (i.e., stratum) by Latino origin[13] cells is not large enough to guarantee a good assignment. The same situation occurred in CHIS 2003. To reduce the number of records to be imputed, a combined race/ethnic variable (OMBSRREO) that assigned Latinos regardless of race into one group was proposed and approved. The levels of the variable OMBSRREO are given in Table 8-5.

[13] Donors and donees must match on the specific Latino origin (Not Hispanic; Mexican; Puerto Rican, Cuban, Central American and Dominican; South American; Other Spanish).

Table 8-5. OBM race/ethnicity groups (OMBSRREO)

OMBSRREO	Description
1	Latino
2	Non-Latino White
3	Non-Latino African American
4	Non-Latino American Indian Alaskan Native
5	Non-Latino Asian
6	Non-Latino Pacific Islander Native Hawaiian
7	Non-Latino two or more races

By creating a separate group for Latinos, a valid value of OMBSRREO was missing for only 74 persons (0.13 percent) who self-reported as non-Latino and "some other race" alone[14] in CHIS 2005. The reduction in the number of cases is because most of the people who report other race were Latino. Using a variable that combined race-ethnic groups with one level of OMBSRREO for Latino eliminated the need to impute for 5,835 cases who reported Latino "other race" alone.

For continuity with the race and ethnicity variables created in CHIS 2001 and 2003 (see Table 8-1), the same variables were created and imputed in CHIS 2005. We refer to these variables as the "regular" single race and ethnicity variables. The OMB race-ethnicity variable OMBSRREO was created using these regular race and ethnicity variables after imputation. Section 8.4.1 describes the imputation of the regular race and ethnicity variables while Section 8.4.2 describes the creation and imputation of the OMB race variable. Section 8.4.3 discusses the creation and imputation of self-reported Asian ethnic groups.

8.4.1 Imputation of Single Self-Reported Race and Ethnicity

While the procedures used to impute for missing values of sex and age were relatively straightforward, self-reported race and ethnicity presented a greater challenge. Different imputation methods were considered before choosing the final approach. One approach that was considered, but not adopted, was to use the self-reported race and ethnicity of a respondent to impute for any other sampled person with missing values for these items within the household. The reason this approach was not used in any cycle of CHIS is the realization that the method does not account for households with persons of more than one race and ethnicity.

[14] This includes records imputed as non-Latino "other" from the regular CHIS 2005 race imputation.

Instead a hot-deck imputation procedure was developed to deal with the diversity of race and ethnicity within households. Before describing the hot-deck approach, some special features of the race and ethnicity items are worth noting. First, although race is a series of items with subparts, the items we deal with are only those that classify a person as White, African American, Asian, American Indian/Alaska Native, Pacific Islander, or other. Also, these items are treated as either all reported or all missing. In very few cases there were missing values for one of the races but not others, but the data preparation staff was able to replace these missing values using interviewer comments. Finally, some missing values were assigned deterministically based on other items such as country of origin. These deterministic imputations were flagged like all other imputations.

Table 8-6 shows the number and percentage of cases with imputed values by type of extended interview (adult, child, and adolescent). The first columns are those cases where race is imputed, and the next set of columns is for cases where ethnicity is imputed.

Table 8-6. Number and percentage of imputed interviews with missing self-reported race and/or ethnicity

	Imputed race*		Imputed ethnicity	
Type of interview	Count	%	Count	%
Adult	962	2.24	128	0.30
Child	447	3.94	40	0.35
Adolescent	217	5.39	27	0.67
Total	1,626	2.78	195	0.33

* At least one value of race was imputed.
Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The hot-deck imputations were done separately by the completed extended interview structure of the household. In general, the imputation procedure was done at the household level and handled households with the fewest missing values first and then moved to the cases with more missing values. The simplest household structure is where only an adult was interviewed (versus a household with an adult and an adolescent and/or a child). A household with only one adult with missing ethnicity was imputed before a household with only an adult that had both missing race and ethnicity.

The patterns of missing data for race and ethnicity varied by the structure of the household. For the simple case where only an adult was interviewed, the donors were selected from other adult-only households. If the adult was missing both race and ethnicity, both values were imputed from the same donor. If the adult had a reported race but was missing ethnicity, then a donor with the same race (all six race values were placed into a vector and only adults with the

exact same values could be donors) was randomly selected. For an adult with reported ethnicity and missing race, the same procedure was used; only adults in adult-only households with the same value of ethnicity could be donors. Whenever possible, the donors and the recipients were from the same sampling stratum. For cases where the pool formed in this way had too few donors, sampling strata were combined based on geographic and urban status. Once a donor was used, it was removed from the pool for all future hot deck runs.

The same principles were used for more complex household structures. In these cases, some households had missing race and ethnicity for all sampled persons, while in others one or more of the sampled persons might have a reported race and ethnicity. Various combinations, such as a reported ethnicity but not race, were also encountered. Separate hot deck runs were made to accommodate all of these situations. As an illustration, consider households where an adult and a child are interviewed. Assume the adult reported non-Latino ethnicity and Asian race and the child only reported non-Latino ethnicity but no race. The pool of donors for imputing the child's race consists of households where only an adult and a child were interviewed and where the adult reported non-Latino ethnicity and Asian race and the child reported non-Latino ethnicity. The households with other combinations of persons with missing race and/or ethnicity were imputed in an similar way. Table 8-7 shows the counts and percentages of imputed values by self-reported race and ethnicity and type of extended interview (adult, child, and adolescent).

Table 8-7. Counts and percentages of imputed interviews with missing self-reported race and ethnicity by type of extended interview

	Total		Extended interview type					
			Adult		Child		Adolescent	
	Count	%	Count	%	Count	%	Count	%
Self-reported race								
White alone	681	1.17	419	0.97	207	1.82	55	1.37
African American alone	33	0.06	20	0.05	5	0.04	8	0.20
Asian alone	98	0.17	60	0.14	14	0.12	24	0.60
American Indian/ Alaska Native alone	43	0.07	26	0.06	11	0.10	6	0.15
Pacific Islander alone	13	0.02	5	0.01	5	0.04	3	0.07
Other race alone	719	1.23	409	0.95	192	1.69	118	2.93
Two or more races	39	0.07	23	0.05	13	0.11	3	0.07
Total	1,626	2.78	962	2.24	447	3.94	217	5.39
Self Reported Ethnicity								
Latino	39	0.07	18	0.04	14	0.12	7	0.17
Non-Latino	156	0.27	110	0.26	26	0.23	20	0.50
Total	195	0.33	128	0.30	40	0.35	27	0.67
Completed interviews	58,407	100.00	43,020	100.00	11,358	100.00	4,029	100.00

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

8.4.2 Imputation of the OMB Race-Ethnicity Variable

The DOF control totals are defined in terms of OMB race categories for raking dimensions 5, 6, and 8. Persons who reported themselves as Latino "some other race" were assigned an OMB race following procedures similar to those used in Census 2000. Since the OMB assignment is done using the imputed regular single race variables, all sampled persons have nonmissing races values for variables SRW, SRAA, SRAI, SRAS, SRPI, and SRO.

The OMB race-ethnicity variable, OMBSRREO, was assigned as follows:

- If the person self-reported as Latino (SRH=1), the variable OMBSRREO was set to 1. This assignment is independent of the values of the race variables.
- If the person self-reported as non-Latino (SRH=2) and reported OMB race alone or in combination with one or more OMB races (e.g., White alone, White and Black or African American, White and Black or African American and American Indian and Alaska Native) then OMBSRREO was given the value 2, 3, 4, 5, 6 or 7 (see Table 8-5) depending on the values of SRW, SRAA, SRAI, SRAS, and SRPI. In other words, there is no modification of race for non-Latinos who reported a valid OMB race(s).

- If the person self-reported as non-Latino (SRH=2) and reported both an OMB race and "some other race" (SRO=1), then OMBSRREO was assigned using only the specified OMB race(s). For example, non-Latino White and some other race became non-Latino White alone. This scenario is an example of the differences between OMBSRREO and the regular race-ethnicity variables (SRH, SRW, SRAA, SRAI, SRAS, and SRPI). Persons who reported two races, with one of them "some other race" are considered as single race respondents based on the OMB definition.

After the race/ethnicity assignments were made, 74 persons (0.13 percent) remained with missing values of OMBSRREO. These persons self-reported as non-Latino and other race only (SRH=2 and SRO=1). The missing values were imputed using the same procedures used to impute the regular single race variables as described above. In this case, temporary OMB race variables named SRW2, SRAA2, SRAI2, SRAS2, and SRPI2 were created using the values of already imputed SRW, SRAA, SRAI, SRAS, and SRPI. The values of the temporary OMB race variables were set to missing for the cases where the person self-reported as non-Latino and other race only. The missing values were imputed through a series of hot-deck imputations where pools of donors were created by matching the structure of the household and non-missing values of race and ethnicity of the adult, child, or adolescent in the household within geographic areas (i.e., stratum, region, or urban/rural area). For cases where there was no pool of donors based on household structure, missing values were imputed using the value of SRW2, SRAA2, SRAI2, SRAS2, and SRPI2 from another member of the household. Next, the variable OMBSRREO was assigned for the records with SRH=2 and SRO=1 using the imputed of values SRW2, SRAA2, SRAI2, SRAS2, and SRPI2. Table 8-8 shows the counts and percentages of imputed OMBSRREO values by type of extended interview (adult, child, and adolescent).

Table 8-8. Number and percentage of completed interviews with missing OMB race and ethnicity by extended interview type

OMB Race-ethnicity (OMBSRREO)	Total		Extended interview type: Adult		Child		Adolescent	
	Imputed Count	%	Imputed Count	%	Imputed Count	%	Imputed Count	%
1. Latino	NA		NA		NA		NA	
2. Non-Latino White alone	55	0.09	39	0.09	12	0.11	4	0.10
3. Non-Latino African American alone	7	0.01	6	0.01	1	0.01	0	0.00
4. Non-Latino Asian alone	1	0.00	1	0.00	0	0.00	0	0.00
5. Non-Latino American Indian/ Alaska Native alone	9	0.02	9	0.02	0	0.00	0	0.00
6. Non-Latino Native Hawaiian and Other Pacific Islander alone	0	0.00	0	0.00	0	0.00	0	0.00
7. Non-Latino two or more races	2	0.00	1	0.00	1	0.01	0	0.00
Total	74	0.13	56	0.13	14	0.12	4	0.10
Completed interviews	58,407		43,020		11,358		4,029	

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

8.4.3 Self-Reported Asian Ethnic Group

The person weights were raked using a dimension defined for Asian groups (Dimension 7). Since there was only one weight for the combined RDD and supplemental list samples in CHIS 2003 and 2005, we added a variable (OMBSRASO) for a raking dimension that would improve the estimates of the largest Asian ethnic groups in California. The variable OMBSRASO identifies the OMB non-Latino Asian ethnic group and is defined in Table 8-9.

Table 8-9. OMB Non-Latino Asian ethnic groups (OMBSRASO)

OMBSRASO	Description
1	Non-Latino Chinese alone
2	Non-Latino Korean alone
3	Non-Latino Filipino alone
4	Non-Latino Vietnamese alone
5	Other

The process to derive the variable OMBSRASO used the temporary OMB race variable SRAS2 previously created for the imputation of OMBSRREO. For records where SRAS2=1 (self-reported as OMB Asian alone or combined with some other race), five flags indicating the Asian ethnic groups of the respondent were derived using the Asian ethnic group questions in the extended interview (questions AA5E_1 to AA5E_18 for adults, TI7_1 to TI 7_18 for adolescents, and CH7_1 to CH7_18 for children). The name and description of the Asian ethnic group flags are shown in Table 8-10.

Table 8-10. OMB Asian group flags

Variable	Description
SRCH	Self-reported Chinese
SRPH	Self-reported Filipino
SRKR	Self-reported Korean
SRVT	Self-reported Vietnamese
SRASO	Self-reported Other Asian ethnic group

The missing values of the OMB Asian group variables (SRCH, SRPH, SRKR, SRVT, and SRASO) were imputed in the same way as the OMB race variables (or the temporary OMB race variables). A series of hot-deck imputations were run where pools of donors were created by matching the structure of the household and non-missing values of race, ethnicity, and Asian ethnic group of the adult, child, or adolescent in the household within geographic areas (i.e., stratum, region, or urban/rural area). For cases where there was no pool of donors based on household structure, race, ethnicity and Asian ethnic group, missing values were imputed using the values of SRCH, SRPH, SRKR, SRVT, and SRASO from another member of the household. The variable OMBSRASO was then created using the variables SRH, SRAA2, SRAI2, SRAS2, SRPI2, and the variables SRCH, SRPH, SRKR, SRVT, and SRASO after imputation. Table 8-11 shows the counts and percentages of imputed OMBSRASO values by type of extended interview (adult, child, and adolescent).

Table 8-11. Number and percentage of completed interviews with imputed OMB Asian ethnic group by extended interview type

OMB Asian group (OMBSRASO)	Total		Extended interview type: Adult		Child		Adolescent	
	Imputed count	%	Imputed count	%	Imputed count	%	Imputed count	%
1. Non-Latino Chinese	32	0.05	21	0.05	9	0.08	2	0.05
2. Non-Latino Korean	9	0.02	6	0.01	2	0.02	1	0.02
3. Non-Latino Filipino	24	0.04	10	0.02	10	0.09	4	0.10
4. Non-Latino Vietnamese	10	0.02	8	0.02	2	0.02	0	0.00
5. Other	23	0.04	14	0.03	6	0.05	3	0.07
Total	98	0.17	59	0.14	29	0.26	10	0.25
Completed interviews	58,407		43,020		11,358		4,029	

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

8.5 Self-Reported County and Self-Reported Stratum

In CHIS 2005, the geographic location variables such as self-reported county of residence, Los Angeles SPA, San Diego County Health Region, Census tract, and self-reported stratum were assigned after geocoding the geographic information collected during the interview (address of respondent, nearest street intersection, or self-reported county) or attached to the sample telephone number (the mailing address or ZIP Code covered by the telephone exchange). The Table 8-12 shows the variables used in the geocoding process.

Table 8-12. Variables used in geocoding

Variable	Description	Source
AH42	County of residence (self report)	Adult questionnaire
AO1ADDR	Confirmed/corrected street address	Adult questionnaire
AO1CITY	Confirmed/corrected city	Adult questionnaire
AO1ZIP	Confirmed/corrected ZIP Code	Adult questionnaire
AM7	ZIP Code (self report)	Adult questionnaire
AO2ADDR	Street address (self report)	Adult questionnaire
AO2CITY	City (self report)	Adult questionnaire
AM8	Street name of residence (self report)	Adult questionnaire
AM9	Street name of nearest cross street (self report)	Adult questionnaire
M_ADDR	Street address (matched to phone number prior to interview)	Address mailing vendor
M_CITY	City (matched to phone number prior to interview)	Address mailing vendor
M_ZIP	ZIP Code (matched to phone number prior to interview)	Address mailing vendor
S_ZIP	ZIP Code (provided by sample vendor for every phone)	Sample vendor

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

The derived location variables SRSTRATA (self-reported stratum), SRCOUNTY (self-reported county), SR_LASPA (self-reported Los Angeles SPA), SR_HR (self-reported San Diego County Health Region) are household-level variables that were assigned to all adult, child and adolescent records within the same household before creating the raking dimensions. The variable SRSTRATA was used to create the cells for raking dimensions 1, 2, and 8 defined at the stratum or California region level while the variables SRCOUNTY (self-reported county), SR_LASPA (self-reported Los Angeles SPA), and SR_HR (self-reported San Diego County Health Region) were used to create the cells for raking dimension 4 defined for Los Angeles County and San Diego County.

Table 8-13 shows the distribution of adult respondents by self-reported stratum compared with the sampling stratum. Each stratum had migration in and migration out as a result of self-reports not matching the sampling stratum. This table shows that the net effect of cross-stratum migration is small, with the greatest differences for strata with the lowest geographic counts, as indicated by the net agreement ratios (NAR) in the rightmost column of Table 8-13. The NAR is the number of respondents in the sampling stratum divided by the number of respondents in the self-reported stratum. A NAR value less than one indicates more in-migration than out-migration from the stratum, and a value greater than one the reverse. Most values are very close to one, indicating either very little migration or roughly equivalent rates of in- and out-migration.

Table 8-13. Distribution of self-reported strata and sampling strata

Stratum	Sampling stratum	Self-reported stratum	Net agreement ratio
Los Angeles	11,650	11,664	1.00
San Diego	6,036	6,037	1.00
Orange	3,374	3,312	1.02
Santa Clara	1,994	1,993	1.00
San Bernardino	1,877	1,882	1.00
Riverside	1,826	1,845	0.99
Alameda	1,738	1,691	1.03
Sacramento	1,553	1,564	0.99
Contra Costa	1,121	1,189	0.94
Fresno	838	840	1.00
San Francisco	914	897	1.02
Ventura	871	899	0.97
San Mateo	882	924	0.95
Kern	855	850	1.01
San Joaquin	658	658	1.00
Sonoma	605	613	0.99
Stanislaus	642	606	1.06
Santa Barbara	645	639	1.01
Solano	1,626	1,594	1.02
Tulare	668	682	0.98
Santa Cruz	692	669	1.03
Marin	3,930	3,916	1.00
San Luis Obispo	633	635	1.00
Placer	627	627	1.00
Merced	699	732	0.95
Butte	577	589	0.98
Shasta	653	621	1.05
Yolo	663	672	0.99
El Dorado	608	606	1.00
Imperial	607	606	1.00

Table 8-13. Distribution of self-reported strata and sampling strata (continued)

Stratum	Sampling stratum	Self-reported stratum	Net agreement ratio
Napa	610	629	0.97
Kings	697	694	1.00
Madera	638	630	1.01
Monterey	750	774	0.97
Humboldt	1,042	1,038	1.00
Nevada	504	498	1.01
Mendocino	527	519	1.02
Sutter	520	513	1.01
Yuba	540	513	1.05
Lake	475	468	1.01
San Benito	518	519	1.00
Colusa, Glen, Tehama	550	578	0.95
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	492	508	0.97
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	482	474	1.02

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

9. VARIANCE ESTIMATION

This chapter describes the methods and results of computing sampling errors for the CHIS 2005 data. The first section gives an overview of the reason for computing sampling errors and summarizes the precision of estimates for adults, children, and adolescents produced from CHIS 2005. The remainder of the chapter describes the methodology for producing estimates of sampling variability. Section 9.2 is a general review of the two main methods of computing sampling errors or variances of estimates from surveys with complex sample designs like CHIS 2005. Section 9.3 describes a replication method of variance estimation that can be used with the data. Section 9.4 shows how analysts can compute sampling errors for CHIS 2005 estimates using commercially available software.

9.1 Design Effects

To evaluate the precision of sample estimates derived from a survey, sampling errors are computed. Estimates of sampling errors can be used to make inferences about the size of the difference between two population parameters based on the values of corresponding sample estimates, their estimated precision, and the expected probability distribution of such a difference. Suppose an analyst wishes to compare the proportion of employed persons whose employer offers health care benefits in two counties in California. By taking the estimated sampling error of this difference into account, the analyst can make inferences about the size of the difference.

Inferences of this nature require an estimate of the precision or sampling error of the characteristic being investigated. There are a variety of ways of reporting the estimated precision of a survey estimate including:

- A standard error (the standard deviation of the estimate);
- A variance of an estimate (the standard error squared);
- A coefficient of variation (the ratio of the standard error to the estimate); or
- A confidence interval (the estimate plus or minus a multiple of the standard error).

Another way of describing the variability of an estimate from a survey is by using the "design effect." The concept of a design effect was introduced and popularized by Kish (1965) to

account for the additional variability associated with complex sample designs involving stratification and clustering. The design effect is the ratio of the variance of the sample estimate for the survey (with its particular sample design and estimation method) to the variance of a simple random sample of the same sample size.

For a specific sample, the design effect, or *DEFF*, for an estimate from a survey can be estimated as

$$DEFF = \frac{\text{sampling variance of a complex sample}}{\text{sampling variance of a simple random sample}}$$

At the analysis stage, the *DEFF* is useful because many statistical procedures in statistical software assume the data are from a simple random sample when computing sampling errors of estimates. The *DEFF* can, in some circumstances, indicate the appropriateness of this assumption and can be used to adjust the sampling errors of the estimates to produce ones that are closer to the actual sampling errors (Skinner, Holt, and Smith, 1989).

Calculating the design effect for a proportion is straightforward because the variance of an estimated proportion in a simple random sample can be estimated easily. In this case, the estimated *DEFF* for a proportion is

$$DEFF_{PROP} = \frac{v(\hat{p})_{COMPLEX}}{v(\hat{p})_{SRS}}$$

where $\hat{p}$ is the estimated proportion, $v(\hat{p})_{SRS}$ is the variance estimate of the estimated proportion assuming a simple random sample, and $v(\hat{p})_{COMPLEX}$ is the variance of the estimated proportion accounting for the complex sample survey design.

In most surveys, design effects are larger than one. In CHIS 2005, design effects are greater than one mainly because the cases have different estimation weights (Kish, 1992). As will be seen shortly, some design effects from the survey are considerably greater than one for statewide estimates.

Design effects are of primary interest to data users. They reveal that the complex sample design and estimation procedures used resulted in estimates of variances that are greater than what would be obtained from a simple random sample. A simple random sample design was not considered for CHIS 2005, because it would not have achieved the sample sizes for the specific

domains of interest, in particular at the county/stratum level for given resources. The design effects calculated from the CHIS 2005 data indicate that the sample design used in the survey need to be taken into account when analyzing the data.

In CHIS 2005, as in most large-scale surveys, a large number of data items are collected. Each resulting variable has its own design effect. One way to summarize the design effects for the items is to compute *DEFF*s for a number of items and then average them. This average represents the design effects for similar items from the survey, as described in Wolter (1985).

The first panel in Tables 9-1 to 9-3 gives the average, median, minimum, and maximum, *DEFF*s computed for a combination of categorical and continuous variables. The rightmost panel shows the average *DEFT* for the same items. The *DEFT* is the square root of the design effect, and it is similar to the *DEFF* but on the scale of the standard error of the estimate rather than the variance.

Table 9-1 shows the *DEFF*s and *DEFT*s for 38 items selected from the adult interview by stratum. Tables 9-2 and 9-3 present the corresponding *DEFF*s and *DEFT*s for 25 items from the child interviews and for 34 items from the adolescent interviews by stratum, respectively. The *DEFF*s and *DEFT*s for the adult, child, and adolescent are given separately for the categorical and the continuous variables in Appendix C.

The *DEFT* is often considered a more convenient measure than the *DEFF*, because it can be used directly when computing confidence intervals for the estimates. See Verma and O'Muircheartaigh (1980) for a discussion of the use of the *DEFT*. The main reason for presenting the *DEFT*s here is because it dampens some of noise associated with the *DEFF*s. The maximum and minimum values of the *DEFF*s in the tables show that there is considerable variability in these quantities. Taking the *DEFTs* and averaging these values smoothes the variability somewhat. For example, in Table 9-1, the average *DEFF* for Orange County is 1.75, while the range is 3.37. This is one of the larger *DEFF*s in the state. However, the average *DEFT* for Orange County is 1.30, which is more in line with the *DEFT*s of the other counties.

Before reviewing the tables in more detail, it is important to discuss the most important factors that result in design effects larger than one. These factors are

- **Oversampling.** The need for both county and state estimates required oversampling to produce stable estimates for these areas. This oversampling increased the design effect for statewide estimates. Another form of oversampling was the refusal subsampling, but this had a relatively minor effect on the design effects.

- **Within-Household Subsampling.** Only one adult and one child or adolescent was sampled in each household. This subsampling contributed to the differential weights at the person level because persons in households with more persons were subsampled at lower rates. In addition, young children (age 0 to 5 years) were oversampled with a measure of size that was twice that of older children (age 6 to 11 years)

- **Weighting Adjustments.** Differential weights were applied to reduce nonresponse bias and to make the estimates consistent with known population totals. The main reason for including these adjustments was to reduce biases in the estimates, but some of the adjustments may have increased the design effects for some estimates.

Table 9-1 shows the average *DEFT*s for estimates of adult items are between 1.08 to 1.33 in most strata. This implies that for most strata the standard error of the estimates is about 8 to 33 percent greater than the expected standard error of a simple random sample. The average *DEFT* is very consistent by area. The average *DEFT* for the state estimates is 1.37. This is slightly larger than the county-level *DEFT*s as expected because counties were not sampled proportional to their population.

Table 9-1. Average *DEFF* and *DEFT* for estimates from the adult interview

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average
Total	1.93	1.93	3.92	0.12	1.37
Los Angeles	1.52	1.50	2.36	0.92	1.23
San Diego	1.65	1.64	2.35	0.94	1.27
Orange	1.75	1.69	3.75	0.38	1.30
Santa Clara	1.53	1.38	3.60	0.80	1.22
San Bernardino	1.45	1.51	1.96	0.89	1.20
Riverside	1.39	1.31	2.43	0.62	1.16
Alameda	1.54	1.51	3.37	0.73	1.22
Sacramento	1.43	1.33	3.84	0.75	1.18
Contra Costa	1.48	1.42	3.01	0.33	1.20
Fresno	1.47	1.50	2.29	0.61	1.20
San Francisco	1.40	1.37	2.16	0.76	1.17
Ventura	1.41	1.41	2.47	0.85	1.18
San Mateo	1.62	1.55	2.99	0.52	1.25
Kern	1.48	1.47	2.43	0.67	1.21
San Joaquin	1.53	1.48	3.48	0.53	1.21
Sonoma	1.41	1.46	2.48	0.75	1.17
Stanislaus	1.25	1.20	1.96	0.50	1.11
Santa Barbara	1.56	1.52	3.10	0.80	1.23

Table 9-1. Average *DEFF* and *DEFT* for estimates from the adult interview (continued)

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Average	Maximum	Minimum	
Solano	1.39	1.38	2.08	0.75	1.17
Tulare	1.44	1.50	2.31	0.76	1.19
Santa Cruz	1.27	1.18	3.39	0.66	1.11
Marin	1.81	1.88	2.96	0.79	1.33
San Luis Obispo	1.43	1.36	2.68	0.69	1.18
Placer	1.24	1.20	2.26	0.47	1.10
Merced	1.56	1.54	3.35	0.75	1.23
Butte	1.39	1.39	2.08	0.74	1.17
Shasta	1.28	1.32	1.97	0.61	1.12
Yolo	1.40	1.43	2.17	0.83	1.17
EL Dorado	1.22	1.15	3.24	0.74	1.10
Imperial	1.39	1.30	2.65	0.68	1.17
Napa	1.41	1.42	2.26	0.41	1.17
Kings	1.28	1.26	2.16	0.60	1.12
Madera	1.40	1.43	2.25	0.30	1.17
Monterey	1.24	1.18	2.14	0.68	1.10
Humboldt	1.46	1.51	2.35	0.56	1.20
Nevada	1.28	1.25	2.18	0.50	1.12
Mendocino	1.19	1.19	1.72	0.63	1.08
Sutter	1.68	1.76	3.07	0.71	1.28
Yuba	1.42	1.41	2.37	0.82	1.18
Lake	1.41	1.36	2.51	0.71	1.18
San Benito	1.65	1.64	3.83	0.64	1.27
Colusa, Glen, Tehama	1.22	1.23	1.92	0.68	1.10
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.33	1.32	1.99	0.92	1.15
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.26	1.27	1.95	0.68	1.11

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Table 9-2 shows the average *DEFT* for estimates from the child interview in each county. The average *DEFT* at the state level is 1.63. The average *DEFT*s for the counties vary between 1.08 and 1.51; that is, the standard errors of these estimates are between 8 and 50 percent greater than expected from a simple random sample. The average *DEFT*s for estimates from the child interview are larger than those for the adult interview. This result is expected, because the subsampling at the person level for children is typically more variable than it is for adults (the number of children per household is more variable than the number of adults per household). Also, young children (age 0 to 5 years) were sampled at a rate twice that of older children (age 6 to 11 years) contributing further to the higher *DEFT*s. It is worth noting here that the *DEFT*s for the first four counties in Table 9-2 are larger than those of the counties in the rest of the table. This is a result of the differential sampling within these four counties between areas of high and low concentration of minorities. Of these four counties, San Diego County has the smallest (average

DEFT = 1.37). When calculating the *DEFT*s, all samples were used, including the San Diego oversample. This sample increased the number of children in San Diego, which increased the precision associated with child estimates.

Table 9-2. Average *DEFF* and *DEFT* for estimates from the child interview

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average
Total	2.68	2.66	4.66	1.83	1.63
Los Angeles	2.28	2.20	3.89	1.32	1.50
San Diego	1.90	1.85	2.92	0.72	1.37
Orange	2.32	2.15	4.39	1.38	1.51
Santa Clara	2.18	1.95	4.88	0.84	1.45
San Bernardino	1.72	1.66	3.59	0.79	1.30
Riverside	1.64	1.66	2.40	0.66	1.27
Alameda	1.50	1.46	2.87	0.80	1.21
Sacramento	1.50	1.51	2.64	0.97	1.21
Contra Costa	1.43	1.43	2.19	0.56	1.19
Fresno	1.30	1.23	2.05	0.59	1.13
San Francisco	1.40	1.40	2.25	0.56	1.17
Ventura	1.69	1.66	3.53	0.67	1.28
San Mateo	1.88	1.73	5.42	0.50	1.33
Kern	1.63	1.54	2.44	0.52	1.26
San Joaquin	1.62	1.61	4.78	0.51	1.23
Sonoma	1.20	1.25	1.79	0.26	1.08
Stanislaus	1.60	1.51	2.44	0.82	1.25
Santa Barbara	1.40	1.43	2.80	0.44	1.17
Solano	1.59	1.40	3.00	0.51	1.24
Tulare	1.59	1.62	2.62	0.89	1.25
Santa Cruz	1.22	1.26	1.70	0.40	1.09
Marin	1.98	1.66	4.32	0.94	1.39
San Luis Obispo	1.47	1.43	3.38	0.41	1.19
Placer	1.39	1.33	4.21	0.21	1.14
Merced	1.47	1.41	2.36	0.68	1.20
Butte	1.42	1.36	2.28	0.65	1.18
Shasta	1.47	1.47	2.82	0.30	1.18
Yolo	1.34	1.32	2.13	0.45	1.15
EL Dorado	1.66	1.50	4.25	0.81	1.26
Imperial	1.54	1.43	3.19	0.78	1.22
Napa	1.22	1.23	1.78	0.54	1.09
Kings	1.59	1.53	3.00	0.65	1.24
Madera	1.48	1.46	2.41	0.76	1.21
Monterey	1.53	1.48	2.89	0.49	1.22
Humboldt	1.48	1.46	2.97	0.55	1.20
Nevada	1.34	1.17	3.10	0.55	1.13
Mendocino	1.26	1.29	1.92	0.77	1.11
Sutter	1.65	1.61	2.68	0.50	1.26

Table 9-2. Average *DEFF* and *DEFT* for estimates from the child interview (continued)

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average
Yuba	1.26	1.21	2.38	0.15	1.09
Lake	1.44	1.45	3.15	0.40	1.18
San Benito	1.22	1.26	1.95	0.17	1.08
Colusa, Glen, Tehama	1.19	1.19	1.91	0.77	1.09
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.24	1.26	1.87	0.46	1.10
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.58	1.45	3.37	0.00	1.21

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Table 9-3 shows that the average *DEFT* for items from the adolescent interviews are similar to those from the child interviews. Since the sampling for adolescents is similar to that for children, we expect a close correspondence between the two. The average *DEFT*s are slightly lower for adolescents, primarily because there are fewer adolescents per household than children which reduces the variability in the weights. The average *DEFT* for the state estimates is 1.39. For most of the strata, the average *DEFT*s are between 1.02 and 1.35.

Table 9-3. Average *DEFF* and *DEFT* for estimates from the adolescent interview

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average
Total	1.96	1.95	4.27	1.40	1.39
Los Angeles	1.82	1.82	3.51	0.89	1.34
San Diego	1.67	1.65	2.32	0.78	1.29
Orange	1.90	1.75	3.67	0.44	1.35
Santa Clara	1.46	1.43	3.94	0.42	1.19
San Bernardino	1.42	1.41	2.01	0.87	1.18
Riverside	1.39	1.32	2.85	0.74	1.16
Alameda	1.37	1.35	2.05	0.71	1.16
Sacramento	1.70	1.59	3.82	0.27	1.28
Contra Costa	1.16	1.20	1.70	0.37	1.06
Fresno	1.53	1.51	2.56	0.72	1.23
San Francisco	1.24	1.19	4.83	0.45	1.09
Ventura	1.55	1.55	2.56	0.60	1.23
San Mateo	1.50	1.36	4.09	0.44	1.19
Kern	1.20	1.16	2.12	0.14	1.07
San Joaquin	1.30	1.27	2.59	0.60	1.13
Sonoma	1.25	1.22	2.24	0.80	1.11
Stanislaus	1.22	1.17	2.52	0.45	1.09

Table 9-3. Average *DEFF* and *DEFT* for estimates from the adolescent interview (continued)

Stratum	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average
Santa Barbara	1.41	1.40	3.20	0.00	1.16
Solano	1.67	1.56	3.64	0.92	1.28
Tulare	1.27	1.17	2.26	0.57	1.11
Santa Cruz	1.25	1.23	2.84	0.66	1.11
Marin	1.79	1.82	3.09	0.63	1.33
San Luis Obispo	1.09	1.06	1.48	0.85	1.04
Placer	1.51	1.47	2.98	0.71	1.21
Merced	1.45	1.45	2.22	0.47	1.19
Butte	1.13	1.10	1.77	0.49	1.05
Shasta	1.11	1.09	1.79	0.67	1.05
Yolo	1.09	1.14	1.95	0.20	1.02
EL Dorado	1.31	1.18	2.76	0.80	1.14
Imperial	1.72	1.56	2.99	0.83	1.29
Napa	1.09	1.04	1.74	0.51	1.04
Kings	1.27	1.29	1.83	0.43	1.12
Madera	1.47	1.40	3.52	1.06	1.20
Monterey	1.29	1.30	1.87	0.45	1.13
Humboldt	1.29	1.26	2.04	0.81	1.13
Nevada	1.17	1.24	1.70	0.48	1.07
Mendocino	1.47	1.34	3.47	0.62	1.19
Sutter	1.33	1.35	2.30	0.47	1.13
Yuba	1.48	1.48	3.25	0.48	1.20
Lake	1.31	1.27	2.13	0.69	1.14
San Benito	1.27	1.26	2.14	0.00	1.08
Colusa, Glen, Tehama	1.20	1.23	2.17	0.60	1.09
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.15	1.18	1.92	0.67	1.06
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.24	1.35	2.13	0.00	1.08

Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

9.2 Methods for Variance Estimation

Variance estimation procedures have been developed to account for the complex sample design. Using these procedures, factors such stratification, multistage sampling, and the use of differential sampling rates to oversample a targeted subpopulation can be appropriately reflected in estimates of sampling error. The two main methods are replication and linearization or the Taylor series approximation. Wolter (1985) is a useful reference on the theory and applications of these methods. Shao (1996) is a more recent review paper that compares these methods. The rest of this section briefly reviews these methods.

The basic idea behind replication is to draw subsamples from the sample, compute the estimate from each of the subsamples, and estimate the variance of the original sample using the variability of the subsample estimates. Specifically, subsamples of the original "full" sample are selected to calculate subsample estimates of a parameter for which a "full-sample" estimate of interest has been generated. The variability of these subsample estimates about the estimate for the full sample can then be assessed. The subsamples are called replicates, and the estimates from the subsamples are called replicate estimates. Rust and Rao (1996) discuss balanced repeated replication (BRR) and jackknife replication, two general approaches to forming subsamples. They show how the units included in the subsample can be defined using variance strata and units. They also describe how these methods can be implemented using replicate weights.

Replicate weights are created to produce the corresponding replicate estimate. Each replicate weight is computed using the same estimation steps as the full sample weight but using only the subsample of cases comprising each replicate. Once the replicate weights are developed, it is a straightforward matter to compute variance estimates for sample estimates of interest. The variance estimate takes the following form:

$$v\left(\hat{\theta}\right) = c \sum_{k=1}^{G} \left(\hat{\theta}_{(k)} - \hat{\theta}\right)^2 \qquad (1)$$

where

θ is an arbitrary parameter of interest;
$\hat{\theta}$ is the estimate of θ based on the full sample;
$\hat{\theta}_{(k)}$ is the k^{th} estimate of θ based on the observations included in the k^{th} replicate;
G is the total number of replicates formed;
c is a constant that depends on the replication method; and
$v\left(\hat{\theta}\right)$ is the estimated variance of θ.

The other widely used method for variance estimation for complex sample surveys is called linearization and is based on the Taylor series approximation. In this method, the Taylor series linearization of a statistic is formed and then substituted into the formula for calculating the variance of a linear estimate appropriate for the sample design. Linearization relies on the simplicity associated with estimating the variance for a linear statistic even with a complex sample design.

9.3 Design of Replicates

In CHIS 2005, a paired unit jackknife method (JK2[15]), a form of jackknife replication, was selected for computing variances. This section provides details on setting up the replication structure, including the definition of the variance strata and units.

Two major reasons for using replication to estimate variances for CHIS 2005 are operational convenience and the ability to reflect all components of the design and estimation in the estimates of variability. With respect to operational convenience, once replicate weights are constructed, it is very simple to compute estimates of sampling errors. No special care is needed for subgroups of interest, and no knowledge of the sample design is required. If an estimator is needed that was not previously considered, replication methods can be easily used to develop an appropriate estimate of variance. In such a case, variance estimates using a Taylor series approach would require additional work. The variance estimation stratum and unit must also be included in the file for the Taylor series method.

The second reason for using replication is probably more important. Both the nonresponse and raking types of adjustments made in developing the CHIS 2005 analysis weights affect the sampling errors of the estimates produced from the survey. The replicate weights prepared for CHIS reflect all such aspects of weighting. Currently existing software for using the Taylor series method for variance estimation cannot reflect these weighting adjustments. In some Taylor series software poststratification can be taken into account, but only in specific situations.

In the JK2 replication method, adjacent pairs of sampled telephone numbers are treated as having been sampled from the same stratum. Each pair of sampled telephone numbers is treated as an implicit stratum, where each such stratum is defined by the sort order used in the sample selection of telephone numbers. In this method, the constant, c, in equation (1) equals 1. This approach has been used in previous cycles of CHIS and in other RDD studies such as the National Household Education Survey, an ongoing national RDD survey starting in 1991 and with a most recent cycle in 2005 (Hagedorn, et. al., 2006).

The first step in designing the replicate structure is to determine the number of variance estimation strata. In the JK2 method, the number of replicates is equal to the number of

[15] This method is denoted as JK2 in the software program, WesVar, which was used to compute all the sampling errors in this report.

variance estimation strata. The choice of the number of variance estimation strata is based on the desire to obtain an adequate number of degrees of freedom to ensure stable estimates of variance while not having so many as to make the cost of computing variance estimates unnecessarily high. Generally, at least 30 degrees of freedom are needed to obtain relatively stable variance estimates. A number greater than 30 is often targeted because there are other factors that reduce the contribution of a replicate to the total number of degrees of freedom, especially for estimates of subgroups.

For CHIS 2005 and previous cycles of CHIS, we elected to create 80 variance estimation strata, even though many more could have been created. The 80 variance strata were formed as follows. First, the sampled telephone numbers were arranged in the same sort order used in sample selection. Next, adjacent sampled telephone numbers were paired to establish initial variance estimation strata (the first two sampled phone numbers were the first initial stratum, the third and fourth sampled telephone numbers were the second initial stratum, etc). Each telephone number in the pair was randomly assigned to be either the first or second variance unit within the variance stratum. Each pair was sequentially assigned to one of 80 final variance estimation strata (the first pair to variance estimation stratum 1, the second to stratum 2, ..., the 80th stratum pair to stratum 80, the 81st pair to stratum 1, etc.). As a result, each variance stratum had approximately the same number of telephone numbers. The same process was followed for each sampling stratum.

Once the variance strata are created, the replicate weights can be created. The full replicate weights are constructed by first modifying the full sample base weights. The replicate base weight for replicate k for record i is

$$w_i^{(k)} = \begin{cases} 2w_i, & \text{if } i \text{ is in variance stratum } k \text{ and variance unit 1} \\ 0, & \text{if } i \text{ is in variance stratum } k \text{ and variance unit 2} \\ w_i, & \text{if } i \text{ is not in variance stratum } k \end{cases}$$

The same sequence of weighting adjustments used in the full sample weight is then applied to the replicate base weights to create the final replicate weights. Thus, all of the different components of the weighting process are fully reflected in the replicate weights, ranging from household adjustments (nonresponse, adjustment for household noncoverage, and adjustment to control totals) to person adjustments (nonresponse and raking).

9.4 Software for Computing Variances

In the past, most standard statistical software packages assumed a simple random sample when computing estimates of variance. As a result, estimates of variance from these packages had the potential to seriously understate the true variability of the survey estimates. However, in recent years, specialized commercial software has been developed to analyze data from complex surveys (Lepkowski and Bowles, 1996). In this section, we describe the elements needed to compute estimates for CHIS 2005 using some of these programs beginning with the program used to compute the sampling errors in this report, WesVar.

WesVar Version 4.2 (Westat, 2000) is a software package developed and distributed by Westa**t.** WesVar uses replication methods to compute variance estimates. WesVar is an interactive program with a graphical interface that makes it simple to specify the estimates for sampling errors for estimates of interest. The data requests center on sessions called "workbooks." A workbook is a file linked to a specific WesVar data set. In a workbook, the user can request descriptive statistics, as well as analyze and create new statistics. Descriptive statistics of analysis variables can be produced through "table requests" in WesVar. Regression requests support both linear and logistic regression models. Outputs include statistics of interest, such as the sum of weights, means, percentages, along with their corresponding standard errors, design effects, coefficients of variation, and confidence intervals.

To use WesVar with CHIS 2005 data, the only requirements are to identify the full and replicate weights that are on the data file and specify the replication method as JK2. This specification is made when a workbook is opened. All of the standard errors produced will properly account for the sample design and estimation methods because these features are accounted for in the replicate weights.

SUDAAN® (Research Triangle Institute, 2005) is a package developed by Research Triangle Institute to analyze data from complex sample surveys. SUDAAN is available as a standalone package or it can be called using SAS. SUDAAN and WesVar produce the same point estimates. The difference between the two packages is in the method used to compute the variances. While Wesvar uses replication exclusively, SUDAAN can use either a first-order Taylor series expansion approximation (linearization), or replication to compute variances of the estimates. When the Taylor series approximations are used, SUDAAN does not fully take into account complex weighting schemes such as nonresponse adjustments or raking, so the variance estimates will be potentially different than estimates calculated using replication. Medians and

quantiles cannot be computed directly using either the Taylor series method or the replication method, but the same type of approximation is used in WesVar and SUDAAN to compute medians.

For descriptive statistics, SUDAAN offers two procedures: PROC CROSSTAB for categorical variables and PROC DESCRIPT for continuous variables. These procedures can be used to compute statistics of interest, such as sum of weights, means, and percentages along with their corresponding standard errors, design effects, and confidence intervals. Both procedures use the option DESIGN= to specify the type of survey design when calculating variance estimates. If no design type is specified using this option, then a standard "with replacement" design is assumed and linearization is used for variance estimation. Specifying JACKKNIFE assumes the use of replication. In this instance the WEIGHT and NEST statements are also required. SUDAAN also contains procedures for computing other analytic statistics, such as those associated with linear and regression models. Consult the help manuals (also available online) for more detail on the procedures and options available for SUDAAN.

SAS has also introduced procedures to analyze survey data. SAS® Version 9 (SAS Institute, 2005) has two procedures for analyzing survey data: PROC SURVEYMEANS and PROC SURVEYREG. Both use the linearization approach to estimate standard errors. SURVEYMEANS produces estimates of means, proportions, and totals, while SURVEYREG fits linear regression models (logistic regression is not available). No design effects are estimated with either PROC. Estimates of differences or other linear combinations are not available in SURVEYMEANS. These procedures are relatively new in SAS and do not contain as many features as most of the other packages. At the current time, the SAS procedures are the most limited of all the packages discussed here.

Another software package that can be used to analyze survey data is STATA (version 9 is the latest version as of this writing) (STATA Corporation, 2005). STATA is a command driven, fully programmable statistical package used for managing, analyzing, and graphing data. STATA was developed by StataCorp and is available for a variety of platforms, including DOS, Windows, Macintosh, and UNIX. STATA's statistical, graphical, and data management capabilities are fully expandable through programming.

STATA has a family of *svy*- commands to analyze data from sample surveys. The set of analytic methods in STATA is more exhaustive than any other package. The *svy* commands can be used to estimate a variety of quantities such as totals, proportions, means, linear combinations of means, and logistic regression parameters. Two-dimensional tables of totals and proportions, along

with *DEFF*s for proportions can be produced using *svy* **tab.** The command *svy* **mean** can be used to produce the *DEFF*s for proportions by coding the analytical variable with values 0 and 1. To estimate totals using *svy* **total**, a variable with a value of 1 must be created for all records in the file.

Like SUDAAN, STATA can use linearization (**linear** variance type option) or replication (**jack** variance type option) to estimate variances. Again, linearization theory does not account for nonresponse or raking when estimating variances. Besides point estimates (proportions, means, ratios and totals) and their standard errors, STATA can compute confidence intervals, design effects, and misspecification effects. Design and misspecification effects are computed for means and proportions only.

When using linearization theory to estimate variances the software packages referred to above require auxiliary variables that provide information about the sample design. Two variables have been defined and included in the data files (TSVARSTR and TSVRUNIT). TSVARSTR is required for all analyses, but TSVRUNIT is required only when analyses are performed using the combined data. In other words, when separate analyses are done by adults, children or teens the variable TSVRUNIT is not required. The definitions of TSVARSTR and TSVRUNIT are

- **TSVARSTR** (Taylor's series variance stratum). The variable TSVARSTR indicates the variance strata to be used for software that computes estimates of variance using the Taylor series method. The variable TSVARSTR was created by sequentially numbering the sampling strata.

- **TSVRUNIT** (Taylor's series unit). The variable TSVRUNIT indicates the PSU. In this case the PSU is the sampled household. TSVRUNIT was created by sequentially numbering the PSU's within the sampling strata.

The same variables, TSVARSTR and TSVRUNIT, can be used for linearization variance estimation in SUDAAN, SAS, and STATA.

REFERENCES

Anderson, J.E., Nelson, D.E., and Wilson, R.W. (1998). Telephone coverage and measurement of health risk indicators: Data from the national health interview survey. *American Journal of Public Health*, 88, 1392-1395.

Blumberg, S. J., Luke, J. V., Cynamon, M. L. (2006). Telephone coverage and health survey estimates: Evaluating the need for concern about wireless substitution. *American Journal of Public Health*, 96(5), 926-931.

Brackstone, G.J., and Rao, J.N.K. (1979). An investigation of raking ratio estimation. *Sankhya,* 41, 97-114.

Brick, J.M., Judkins, D., Montaquila, J., and Morganstein, D. (2002). Two-phase list-assisted RDD sampling. *Journal of Official Statistics*, 18(2), 203-215.

Brick, J.M., and Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, 5, 215-238.

Brick, J. M., Montaquila, J., Hagedorn, M. C., Roth, S. B., and Chapman, C. (2005). Implications for RDD design from an incentive experiment, *Journal of Official Statistics*, 21(4), 571-589.

Deville, J.C., and Särndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association,* 87, 376-382.

Elliott, M. R., Little, R.J.A., and Lewitzky, S. (2000). Subsampling callbacks to improve survey efficiency. *Journal of the American Statistical Association,* 95(451), 730-738.

Ford, B. M. (1983). *An overview of hot-deck procedures*: *Vol. 2. Incomplete data in sample surveys*. New York: Academic Press.

Ford, E.S. (1998). Characteristics of survey participants with and without a telephone: Findings from the third national health and nutrition examination survey. *Journal of Clinical Epidemiology*, 51, 55-60.

Groves, R.M. (1989). *Survey errors and survey costs*. New York: John Wiley and Sons.

Hagedorn, M., Montaquila, J., Carver, P., O'Donnell, K., and Chapman, C. (2006). National household education surveys program of 2005. *Public-use data file user's manual*: *Vol. I. Study overview and methodology*, (NCES 2006-078). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Hansen, M.H., and Hurwitz, W.N. (1946). The problem of nonresponse in sample surveys. *Journal of the American Statistical Association*, 41, 517–529.

Holt, D., and Smith, T.M.F. (1979). Post Stratification. *Journal of the Royal Statistical Society*. Series A (General), 142(1), 33-46.

Kalton, G., and Flores Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19(2), 81-97.

Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied statistics*, 29, 119-127.

Kish, L. (1992). Weighting for unequal pi. *Journal of Official Statistic,* 8, 183-200.

Kish, L. (1965). *Survey sampling*. New York: John Wiley and Sons.

Lepkowski, J., and Bowles, J. (1996). Sampling error software for personal computers. *The Survey Statistician*, 35, 10-16.

Liu, B., Ferraro, D., Wilson, E., and Brick, J.M. (2004). Trimming extreme weights in household surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association* [CD-ROM]. Alexandria, VA: American Statistical Association.

Office of Management and Budget (1997). *Revisions to the standards for the classification of federal data on race and ethnicity*, 62(210).

Oh, H.L., and Scheuren, F.J. (1983). *Weighting adjustments for unit nonresponse*. Incomplete data in sample surveys: *Vol. 2*. W.G. Madow, I. Olkin, and D.B. Rubin (Eds.). Academic Press, 143-184.

Research Triangle Institute (2005). *SUDAAN user's manual*, *Release 9.0*. Research Triangle Park, NC: Research Triangle Institute.

Rizzo, Lou, Brick J.M., and Park, I. (2004). A minimally intrusive method for sampling persons in random-digit-dial surveys. *Public Opinion Quarterly,* 68, 267-274.

Rust, K.F., and Rao, J.N.K. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5, 282-310.

Sande, I. G. (1983). *Hot-deck imputation procedures*, *incomplete data in sample surveys*: *Vol. 3*. New York: Academic Press.

SAS Institute (2005). *SAS OnlineDoc® version 8,* Cary, NC: SAS Institute, Inc.

Shao, J. (1996). Resampling methods in sample surveys (with discussion). *Statistics*, 27, 203-254.

Shapiro, G., Battaglia, M., Camburn, D., Massey, J., and Tompkins, L. (1995). Calling local telephone company business offices to determine the residential status of a wide class of unresolved telephone numbers in a random-digit-dialing sample. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 975-980.

Skinner, C.J., Holt D., and Smith, T. M. F. (1989). *Analysis of complex surveys*. New York: John Wiley and Sons.

Stata Corporation (2005). *STATA survey data reference manual*, *Release 9*. College Station, TX: Stata Corporation.

State of California, Department of Finance (2004). *Population projections by race/ethnicity for California and it's counties, 2000-2050.* Sacramento, CA.

State of California, Department of Finance (2006). *California county population estimates and components of change by year, July 1, 2000–2005*. Sacramento, CA.

State of California, Department of Finance (2006a). *Estimated race/ethnic population with age and sex detail*, 2000–2004. Sacramento, CA.

State of California, Department of Finance (2006b). *California county population estimates and components of change by year*, July 1, 2000–2005. Sacramento, CA.

Tucker, C., Brick, J.M., Meekins, B., and Morganstein, D. (2004*)*. Household telephone service and usage patterns in the U.S. in 2004. *Proceedings of the Survey Methods Section of the American Statistical Association* [CD-ROM].

U.S. Census Bureau (2001). *2000 Summary File 1* (Prepared by the U.S. Census Bureau). California.

U.S. Census Bureau (2002). *Census 2000 Modified Race Data*. (Prepared by the U.S. Census Bureau).

U.S. Census Bureau (2003). *2002 American Community Survey PUMS file* (Prepared by the U.S. Census Bureau). California.

Verma, V., Scott, C., and O'Muircheartaigh, C. (1980). Sample designs and sampling errors for the world fertility survey. *Journal of the Royal Statistical Society*, A 143, 431-473.

Westat (2000). *WesVar ™ 4.0 user's guide.* Westat, Rockville, MD.

Wolter, K. (1985). *Introduction to variance estimation*. Springer-Verlag.

Appendix A

Table A-1. CHIS 2005 frame sizes[1], sample sizes[2], and base weights by sampling stratum and sampling frame (RDD, Korean and Vietnamese lists)

Sampling stratum	Description	RDD sampling frame			Korean surname list			Vietnamese surname list			Korean/Vietnamese surname list		
		Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight
1.012	Los Angeles, San Fernando SPA – High Density	152,200	8,454	17.98	2,097	30	16.51	242	4	14.24	1,518	29	12.34
1.013	Los Angeles, San Gabriel SPA – High Density	346,000	7,715	45.10	11,907	154	27.76	6,029	81	27.53	5,619	99	22.30
1.014	Los Angeles, Metro SPA – High Density	345,400	7,511	45.93	5,467	71	29.24	384	4	29.54	3,397	57	27.40
1.017	Los Angeles, South SPA – High Density	96,900	5,116	18.98	2,107	25	14.14	379	6	15.79	1,314	28	13.69
1.018	Los Angeles, South Bay SPA – High Density	176,400	6,796	26.05	1,468	20	19.57	636	8	12.72	1,021	20	15.47
1.021	Los Angeles, Antelope Valley SPA – Low Density	220,600	4,572	48.30	262	3	32.75	221	2	20.09	210	3	26.25
1.022	Los Angeles, San Fernando SPA – Low Density	1,604,300	15,965	100.42	4,888	70	45.26	2,332	32	42.40	3,734	76	33.04
1.023	Los Angeles, San Gabriel SPA – Low Density	949,400	12,833	74.07	8,784	119	36.00	5,569	82	34.59	4,474	82	30.44
1.024	Los Angeles, Metro SPA – Low Density	719,400	12,284	58.56	3,497	41	38.86	1,310	24	29.77	2,079	36	25.05
1.025	Los Angeles, West SPA – Low Density	1,036,600	9,478	109.53	3,947	53	39.08	1,217	16	46.81	2,110	39	34.59
1.026	Los Angeles, South SPA – Low Density	693,100	6,371	108.66	1,334	22	38.11	371	5	41.22	1,050	22	45.65

Table A-1. CHIS 2005 frame sizes[1], sample sizes[2], and base weights by sampling stratum and sampling frame (RDD, Korean and Vietnamese lists) (continued)

Sampling stratum	Description	RDD sampling frame			Korean surname list			Vietnamese surname list			Korean/Vietnamese surname list		
		Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight
1.027	Los Angeles, East SPA – Low Density	768,400	9,845	78.14	2,468	34	36.84	768	10	36.57	1,574	23	30.27
1.028	Los Angeles, South Bay SPA – Low Density	1,067,900	13,356	80.03	4,167	52	37.54	2,512	33	32.62	2,605	46	36.69
2.012	San Diego, North Central Health Region – High Density	84,300	5,306	15.92	456	7	12.67	1,149	15	12.49	443	8	10.55
2.013	San Diego, Central Health Region – High Density	53,500	4,998	10.71	92	2	11.50	414	8	8.28	157	3	8.26
2.021	San Diego, North Coastal Health Region – Low Density	384,800	11,527	33.35	820	11	31.54	578	8	27.52	517	8	20.68
2.022	San Diego, North Central Health Region – Low Density	402,000	25,617	15.68	1,392	15	14.81	1,239	20	13.18	934	15	12.45
2.023	San Diego, Central Health Region – Low Density	418,800	13,857	30.20	474	4	31.60	1,338	18	23.89	522	8	21.75
2.024	San Diego, South Health Region – Low Density	235,600	7,864	29.94	357	5	22.31	184	1	30.67	264	3	29.33
2.025	San Diego, East Health Region – Low Density	367,500	10,320	35.59	287	6	23.92	409	6	27.27	241	5	21.91
2.026	San Diego, North Inland Health Region – Low Density	342,000	11,573	29.56	1,213	15	21.28	901	14	20.48	736	16	17.52
3.01	Orange – High Density	837,100	21,176	39.49	6,569	90	26.07	17,272	249	25.90	7,146	131	22.76
3.02	Orange – Low Density	1,878,800	16,891	111.24	6,897	94	42.84	6,233	77	46.51	5,252	101	36.47

Table A-1. CHIS 2005 frame sizes[1], sample sizes[2], and base weights by sampling stratum and sampling frame (RDD, Korean and Vietnamese lists) (continued)

Sampling stratum	Description	RDD sampling frame			Korean surname list			Vietnamese surname list			Korean/Vietnamese surname list		
		Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight
19	Solano	295,300	11,879	24.86	517	8	24.62	456	5	17.54	432	10	13.94
20	Tulare	295,700	4,946	59.86	186	3	23.25	129	2	18.43	117	1	29.25
21	Santa Cruz	262,400	4,862	54.01	415	5	29.64	209	4	20.90	246	3	30.75
22	Marin	328,400	38,003	8.64	592	8	7.05	376	5	8.00	373	9	7.61
23	San Luis Obispo	237,400	4,492	52.84	328	5	27.33	272	4	54.40	224	6	17.23
24	Placer	293,100	4,817	60.82	462	7	33.00	359	3	51.29	338	10	21.13
25	Merced	122,800	3,885	31.592	162	3	18.00	126	1	42.00	124	2	24.80
26	Butte	161,000	3,014	53.307	240	3	48.00	206	3	34.33	154	1	154.00
27	Shasta	151,600	3,387	44.775	141	2	35.25	97	3	16.17	112	2	18.67
28	Yolo	141,900	4,071	34.824	1,077	15	22.91	601	8	23.12	545	10	27.25
29	El Dorado	159,900	4,221	37.903	229	3	25.44	119	2	19.83	134	2	19.14
30	Imperial	95,300	4,690	20.317	117	0	29.25	43	0	21.50	57	2	9.50
31	Napa	116,200	5,293	21.937	125	1	31.25	86	2	21.50	94	3	11.75
32	Kings	71,400	3,747	19.058	72	1	12.00	42	2	10.50	47	0	23.50
33	Madera	88,200	3,960	22.253	72	0	72.00	49	1	16.33	46	0	46.00
34	Monterey	360,800	7,681	46.950	795	11	30.58	412	6	25.75	517	7	34.47

Table A-1. CHIS 2005 frame sizes[1], sample sizes[2], and base weights by sampling stratum and sampling frame (RDD, Korean and Vietnamese lists) (continued)

Sampling stratum	Description	RDD sampling frame			Korean surname list			Vietnamese surname list			Korean/Vietnamese surname list		
		Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight	Frame	Sample	Weight
35	Humboldt	129,000	6,283	20.543	120	2	12.00	77	1	11.00	81	2	16.20
36	Nevada	101,000	2,937	34.389	126	2	25.20	65	1	21.67	84	2	16.80
37	Mendocino	86,700	3,261	26.587	91	2	30.33	72	1	18.00	50	2	8.33
38	Sutter	60,000	3,477	17.274	60	1	12.00	73	1	10.43	31	1	7.75
39	Yuba	51,600	3,090	16.697	101	1	20.20	52	0	17.33	54	1	9.00
40	Lake	60,500	3,107	19.465	73	0	18.25	37	0	37.00	30	0	30.00
41	San Benito	37,100	3,823	9.708	32	0	8.00	25	1	6.25	27	1	6.75
42	Colusa, Glenn, Tehama	78,000	2,976	26.219	47	1	11.75	56	1	18.67	40	0	0.00
43	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	194,800	4,120	47.241	137	1	68.50	115	0	57.50	106	2	35.33
44	Amador, Alpine, Calaveras, Inyo, Mariposa, Mono, Tuolumne	230,300	3,918	58.757	206	3	34.33	160	3	32.00	134	3	26.80
	Total	30,326,700	549,694		137,149	1,832		102,740	1,438		86,313	1,608	

[1] Total number of possible phone numbers in eligible working 100 banks.
[2] Realized number of sampled telephone numbers in strata.
Source: UCLA Center for Health Policy Research, 2005 California Health Interview Survey.

Appendix B

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum

	All strata	Los Angeles	San Diego	Orange	Santa Clara
1. Base weight					
1.1 Sample size	554,572	121,857	91,283	38,809	21,576
1.2 Sum of weight	30,326,700	8,176,600	2,288,500	2,715,900	1,675,200
1.3 Coefficient of variation (CV)	60.01	44.16	33.62	51.44	33.38
2. Refusal conversion subsampling adjustment					
2.1 Sum of weights before adjustment					
a. Household never refused	24,579,323	6,582,723	1,859,816	2,214,287	1,403,287
b. Household refused – selected for ref. conversion	3,113,859	827,386	346,640	239,727	137,902
c. Household refused – not selected for ref. conversion	2,633,518	766,491	82,045	261,886	134,011
2.2 Sum of weights after adjustment					
a. Household never refused	24,579,323	6,582,723	1,859,816	2,214,287	1,403,287
b. Household refused – selected for ref. conversion	5,747,377	1,593,877	428,684	501,613	271,913
c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3 Sum of weights	30,326,700	8,176,600	2,288,500	2,715,900	1,675,200
2.4 Sample size	508,206	109,651	88,157	34,993	19,792
2.5 Coefficient of variation	69.56	53.55	39.99	64.17	44.84
3. Adjusting for unknown residential					
3.1 Sum of weights before adjustment					
a. Residential – respondent	5,410,611	1,364,288	409,990	421,348	244,876
b. Residential – nonrespondent	4,070,857	1,203,592	326,025	362,319	205,885
c. Unknown residential status (NA, NM)	3,089,326	835,513	256,229	272,280	198,382
d. Nonresidential	17,755,906	4,773,207	1,296,256	1,659,952	1,026,057

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	All strata	Los Angeles	San Diego	Orange	Santa Clara
3.2 Sum of weights – allocating unknown residential					
a. Residential – respondent	5,410,611	1,364,288	409,990	421,348	244,876
b. Residential – nonrespondent	4,070,857	1,203,592	326,025	362,319	205,885
c. (NA, NM)	1,372,632	355,263	115,154	115,336	84,610
3.3 Sum of weights before adjustment					
a. Residential – respondent	5,410,611	1,364,288	409,990	421,348	244,876
b. Residential – nonrespondent	5,443,489	1,558,904	441,128	477,816	290,536
c. Estimated residential among unknown	0	0	0	0	0
3.3 Sum of weights	10,854,100	2,923,192	851,119	899,163	535,412
3.4 Sample size	133,610	28,543	25,468	7,763	4,498
3.5 Coefficient of Variation	85.04	67.97	50.01	81.24	69.55
4. Supplemental list-sample eligibility adjustment					
4.1 Sum of weights before adjustment					
a. RDD – household	10,738,114	2,888,941	848,249	883,011	519,935
b. List eligible household	19,416	6,884	517	4,129	3,274
c. List ineligible household	28,879	6,795	774	2,016	3,437
d. Household with unknown list eligibility	67,692	20,572	1,579	10,007	8,765
4.2 Sum of weights before adjustment					
a. RDD household	10,738,114	2,888,941	848,249	883,011	519,935
b. List eligible household	46,614	16,368	1,253	9,944	7,992
c. List ineligible household	69,373	16,303	1,874	4,851	8,211
d. Household with unknown list eligibility	0	0	0	0	0
4.3 Sum of weights	10,854,100	2,921,612	851,377	897,806	536,138
4.4 Sample size	131,851	27,971	25,393	7,485	4,282
4.5 Coefficient of variation (CV)	84.06	66.23	49.89	78.17	64.87

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	All strata	Los Angeles	San Diego	Orange	Santa Clara
5. Screener nonresponse adjustment					
5.1 Sum of weights before adjustment					
a. Respondents	5,404,321	1,365,577	409,541	424,879	245,761
b. Nonrespondents	5,449,780	1,556,035	441,836	472,927	290,377
5.2 Sum of weights after adjustment					
a. Respondents	10,854,100	2,932,180	851,363	907,784	538,462
b. Nonrespondents	0	0	0	0	0
5.3 Sum of weights	10,854,100	2,932,180	851,363	907,784	538,462
5.4 Sample size	84,297	16,902	14,739	4,743	2,638
5.5 Coefficient of variation	77.56	60.75	42.52	70.01	53.74
6. Multiple telephone adjustment					
6.1 Sum of weights before adjustment	10,854,100	2,932,180	851,363	907,784	538,462
6.2 Overall adjustment factor	0.97	0.96	0.96	0.96	0.96
6.3 Sum of weights	10,478,852	2,816,082	820,709	874,640	516,429
6.4 Sample size	84,297	16,902	14,739	4,743	2,638
6.5 Coefficient of variation	79.13	62.41	45.26	71.92	55.09
7. Supplemental child sample subsampling adjustment					
7.1 Sum of weights before adjustment					
a. Households-child supp. w/child-RDD w/o child	9,694,842	2,694,490	433,896	838,840	494,568
b. Households-child supp. w/o children	784,011	121,592	386,813	35,799	21,861
7.2 Sum of weights after adjustment					
a. Households-child supp. w/child-RDD w/o child	10,478,852	2,816,082	820,709	874,640	516,429
b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	All strata	Los Angeles	San Diego	Orange	Santa Clara
7.3 Sum of weights	10,478,852	2,816,082	820,709	874,640	516,429
7.4 Sample size	73,775	16,154	6,979	4,534	2,496
7.5 Coefficient of variation	75.46	63.86	73.92	72.84	55.18
8. Household poststratification					
8.1 Number of completed screeners	73,775	16,154	6,979	4,534	2,496
8.2 Sum of weights before adjustment	10,478,852	2,816,082	820,709	874,640	516,429
8.3 Overall adjustment factor	1.10	1.11	1.21	1.07	1.10
8.4 Sum of weights	11,502,870	3,133,774	994,677	935,287	565,863
8.5 Coefficient of variation	76.14	62.92	74.03	73.01	56.27
9. Section G nonresponse adjustment					
9.1 Sum of weights before adjustment					
a. Household with child 1st procedure	1,003,737	270,245	82,118	87,356	59,331
b. Household w/o child 1st proc.-Sect. G comp.	6,188,825	1,589,367	533,599	474,586	305,893
c. Household w/o child 1st proc.-Sect. G not comp.	4,310,308	1,274,162	378,960	373,345	200,640
9.2 Sum of weights after adjustment					
a. Household with child 1st procedure	1,003,737	270,245	82,118	87,356	59,331
b. Household w/o child 1st proc.-Sect. G comp.	10,499,133	2,863,529	912,559	847,931	506,532
c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3 Sum of weights	11,502,870	3,133,774	994,677	935,287	565,863
9.4 Number HHs w/child 1st procedure	6,998	1,420	1,117	456	276
9.5 Number HHs w/o child 1st proc.–Sect. G comp	41,005	8,420	3,502	2,368	1,360
9.6 Coefficient of variation	80.73	66.29	81.14	75.92	58.52

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
1. Base weight					
1.1 Sample size	14,309	13,731	17,067	12,829	9,502
1.2 Sum of weight	1,255,900	1,253,600	1,493,900	1,129,600	917,900
1.3 Coefficient of variation (CV)	6.90	5.94	10.37	8.44	7.78
2. Refusal conversion subsampling adjustment					
2.1 Sum of weights before adjustment					
a. Household never refused	969,251	951,025	1,258,489	921,014	746,779
b. Household refused – selected for ref. conversion	149,304	152,599	126,099	110,677	94,050
c. Household refused – not selected for ref. conversion	137,346	149,975	109,311	97,909	77,071
2.2 Sum of weights after adjustment					
a. Household never refused	969,251	951,025	1,258,489	921,014	746,779
b. Household refused – selected for ref. conversion	286,649	302,575	235,411	208,586	171,121
c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3 Sum of weights	1,255,900	1,253,600	1,493,900	1,129,600	917,900
2.4 Sample size	12,746	12,095	15,813	11,721	8,706
2.5 Coefficient of variation	31.37	33.02	27.49	28.93	27.08
3. Adjusting for unknown residential					
3.1 Sum of weights before adjustment					
a. Residential – respondent	271,088	281,029	236,970	212,095	164,585
b. Residential – nonrespondent	182,671	197,546	174,556	140,127	114,549
c. Unknown residential status (NA, NM)	117,090	120,314	158,405	112,242	92,465
d. Nonresidential	685,050	654,712	923,969	665,137	546,301

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
3.2	Sum of weights – allocating unknown residential					
	a. Residential – respondent	271,088	281,029	236,970	212,095	164,585
	b. Residential – nonrespondent	182,671	197,546	174,556	140,127	114,549
	c. (NA, NM)	50,845	54,669	68,307	47,836	39,980
3.3	Sum of weights before adjustment					
	a. Residential – respondent	271,088	281,029	236,970	212,095	164,585
	b. Residential – nonrespondent	233,510	252,321	242,662	187,863	154,358
	c. Estimated residential among unknown	0	0	0	0	0
3.3	Sum of weights	504,599	533,349	479,631	399,958	318,943
3.4	Sample size	3,804	3,821	3,697	3,061	2,205
3.5	Coefficient of Variation	46.41	47.27	52.69	49.62	47.23
4.	Supplemental list-sample eligibility adjustment					
4.1	Sum of weights before adjustment					
	a. RDD – household	501,498	530,277	470,872	395,317	316,231
	b. List eligible household	369	477	556	602	325
	c. List ineligible household	946	1,007	2,862	1,340	736
	d. Household with unknown list eligibility	1,786	1,589	5,342	2,699	1,651
4.2	Sum of weights before adjustment					
	a. RDD household	501,498	530,277	470,872	395,317	316,231
	b. List eligible household	874	1,145	1,349	1,463	774
	c. List ineligible household	2,272	2,428	6,871	3,254	1,775
	d. Household with unknown list eligibility	0	0	0	0	0
4.3	Sum of weights	504,644	533,850	479,092	400,033	318,780
4.4	Sample size	3,763	3,788	3,576	3,001	2,170
4.5	Coefficient of variation (CV)	45.21	46.22	49.46	47.56	45.57

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
5. Screener nonresponse adjustment					
5.1 Sum of weights before adjustment					
a. Respondents	270,647	280,413	234,722	211,341	164,007
b. Nonrespondents	233,997	253,437	244,370	188,692	154,772
5.2 Sum of weights after adjustment					
a. Respondents	503,952	532,767	473,846	398,532	318,166
b. Nonrespondents	0	0	0	0	0
5.3 Sum of weights	503,952	532,767	473,846	398,532	318,166
5.4 Sample size	2,510	2,510	2,268	2,007	1,428
5.5 Coefficient of variation	36.80	37.32	38.35	37.18	36.91
6. Multiple telephone adjustment					
6.1 Sum of weights before adjustment	503,952	532,767	473,846	398,532	318,166
6.2 Overall adjustment factor	0.97	0.98	0.96	0.98	0.96
6.3 Sum of weights	490,504	520,456	456,857	390,775	305,989
6.4 Sample size	2,510	2,510	2,268	2,007	1,428
6.5 Coefficient of variation	38.89	39.24	41.25	39.26	40.08
7. Supplemental child sample subsampling adjustment					
7.1 Sum of weights before adjustment					
a. Households-child supp. w/child-RDD w/o child	467,649	499,382	433,573	369,352	288,638
b. Households-child supp. w/o children	22,855	21,075	23,284	21,423	17,351
7.2 Sum of weights after adjustment					
a. Households-child supp. w/child-RDD w/o child	490,504	520,456	456,857	390,775	305,989
b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
7.3	Sum of weights	490,504	520,456	456,857	390,775	305,989
7.4	Sample size	2,367	2,384	2,127	1,873	1,324
7.5	Coefficient of variation	39.36	39.57	41.37	39.81	40.09
8.	Household poststratification					
8.1	Number of completed screeners	2,367	2,384	2,127	1,873	1,324
8.2	Sum of weights before adjustment	490,504	520,456	456,857	390,775	305,989
8.3	Overall adjustment factor	1.08	0.97	1.15	1.16	1.12
8.4	Sum of weights	528,594	506,218	523,366	453,602	344,129
8.5	Coefficient of variation	39.21	40.03	41.46	39.76	39.88
9.	Section G nonresponse adjustment					
9.1	Sum of weights before adjustment					
	a. Household with child 1st procedure	56,289	46,362	40,488	38,235	30,653
	b. Household w/o child 1st proc.-Sect. G comp.	270,769	263,136	308,489	262,045	201,070
	c. Household w/o child 1st proc.-Sect. G not comp.	201,536	196,720	174,390	153,323	112,407
9.2	Sum of weights after adjustment					
	a. Household with child 1st procedure	56,289	46,362	40,488	38,235	30,653
	b. Household w/o child 1st proc.-Sect. G comp.	472,305	459,856	482,878	415,367	313,476
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3	Sum of weights	528,594	506,218	523,366	453,602	344,129
9.4	Number HHs w/child 1st procedure	251	227	169	160	112
9.5	Number HHs w/o child 1st proc.–Sect. G comp	1,254	1,272	1,271	1,115	797
9.6	Coefficient of variation	43.04	42.64	42.99	42.19	39.67

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
1. Base weight					
1.1 Sample size	6,474	13,526	7,639	10,306	6,900
1.2 Sum of weight	663,300	1,078,200	633,000	746,200	531,400
1.3 Coefficient of variation (CV)	6.21	11.49	5.99	7.95	3.81
2. Refusal conversion subsampling adjustment					
2.1 Sum of weights before adjustment					
a. Household never refused	538,216	929,692	508,752	624,471	426,456
b. Household refused – selected for ref. conversion	74,596	87,933	59,875	53,497	41,862
c. Household refused – not selected for ref. conversion	50,488	60,575	64,373	68,232	63,081
2.2 Sum of weights after adjustment					
a. Household never refused	538,216	929,692	508,752	624,471	426,456
b. Household refused – selected for ref. conversion	125,084	148,508	124,248	121,729	104,944
c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3 Sum of weights	663,300	1,078,200	633,000	746,200	531,400
2.4 Sample size	5,982	12,754	6,862	9,358	6,082
2.5 Coefficient of variation	23.32	23.30	32.88	35.11	41.23
3. Adjusting for unknown residential					
3.1 Sum of weights before adjustment					
a. Residential – respondent	132,832	131,958	121,037	104,997	108,501
b. Residential – nonrespondent	77,135	116,842	81,934	90,246	67,014
c. Unknown residential status (NA, NM)	52,646	133,653	66,276	82,375	47,415
d. Nonresidential	400,687	695,747	363,754	468,582	308,470

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Fresno	San Francisco	Ventura	San Mateo	Kern
3.2	Sum of weights – allocating unknown residential					
	a. Residential – respondent	132,832	131,958	121,037	104,997	108,501
	b. Residential – nonrespondent	77,135	116,842	81,934	90,246	67,014
	c. (NA, NM)	21,899	56,438	28,162	34,639	20,688
3.3	Sum of weights before adjustment					
	a. Residential – respondent	132,832	131,958	121,037	104,997	108,501
	b. Residential – nonrespondent	99,085	173,421	110,092	124,943	87,738
	c. Estimated residential among unknown	0	0	0	0	0
3.3	Sum of weights	231,917	305,378	231,129	229,939	196,239
3.4	Sample size	1,624	2,596	1,780	1,891	1,587
3.5	Coefficient of Variation	39.39	53.48	52.96	63.49	62.11
4.	Supplemental list-sample eligibility adjustment					
4.1	Sum of weights before adjustment					
	a. RDD – household	230,502	293,525	229,946	226,889	195,764
	b. List eligible household	94	828	237	133	71
	c. List ineligible household	401	4,770	289	954	170
	d. Household with unknown list eligibility	920	6,257	657	1,962	234
4.2	Sum of weights before adjustment					
	a. RDD household	230,502	293,525	229,946	226,889	195,764
	b. List eligible household	229	1,980	557	314	164
	c. List ineligible household	952	11,453	700	2,290	415
	d. Household with unknown list eligibility	0	0	0	0	0
4.3	Sum of weights	231,683	306,958	231,203	229,494	196,343
4.4	Sample size	1,606	2,461	1,764	1,842	1,582
4.5	Coefficient of variation (CV)	38.24	47.58	51.98	61.29	61.76

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Fresno	San Francisco	Ventura	San Mateo	Kern
5.	Screener nonresponse adjustment					
5.1	Sum of weights before adjustment					
	a. Respondents	132,463	128,081	121,067	104,194	108,347
	b. Nonrespondents	99,220	178,877	110,136	125,300	87,997
5.2	Sum of weights after adjustment					
	a. Respondents	230,983	298,284	231,253	227,837	196,115
	b. Nonrespondents	0	0	0	0	0
5.3	Sum of weights	230,983	298,284	231,253	227,837	196,115
5.4	Sample size	1,105	1,390	1,197	1,157	1,126
5.5	Coefficient of variation	30.35	35.47	41.30	49.04	49.37
6.	Multiple telephone adjustment					
6.1	Sum of weights before adjustment	230,983	298,284	231,253	227,837	196,115
6.2	Overall adjustment factor	0.97	0.96	0.96	0.95	0.98
6.3	Sum of weights	225,084	286,490	222,645	217,278	192,114
6.4	Sample size	1,105	1,390	1,197	1,157	1,126
6.5	Coefficient of variation	32.78	38.64	43.66	52.16	50.21
7.	Supplemental child sample subsampling adjustment					
7.1	Sum of weights before adjustment					
	a. Households-child supp. w/child-RDD w/o child	213,472	269,809	212,824	208,548	184,921
	b. Households-child supp. w/o children	11,612	16,682	9,820	8,730	7,193
7.2	Sum of weights after adjustment					
	a. Households-child supp. w/child-RDD w/o child	225,084	286,490	222,645	217,278	192,114
	b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Fresno	San Francisco	Ventura	San Mateo	Kern
7.3	Sum of weights	225,084	286,490	222,645	217,278	192,114
7.4	Sample size	1,032	1,292	1,132	1,097	1,074
7.5	Coefficient of variation	32.81	39.13	44.37	52.84	50.93
8.	Household poststratification					
8.1	Number of completed screeners	1,032	1,292	1,132	1,097	1,074
8.2	Sum of weights before adjustment	225,084	286,490	222,645	217,278	192,114
8.3	Overall adjustment factor	1.12	1.15	1.09	1.17	1.09
8.4	Sum of weights	252,940	329,700	243,234	254,103	208,652
8.5	Coefficient of variation	32.74	42.35	44.43	53.78	50.90
9.	Section G nonresponse adjustment					
9.1	Sum of weights before adjustment					
	a. Household with child 1st procedure	21,393	11,551	21,524	24,213	19,267
	b. Household w/o child 1st proc.-Sect. G comp.	139,478	193,466	118,284	146,044	107,280
	c. Household w/o child 1st proc.-Sect. G not comp.	92,069	124,683	103,426	83,845	82,105
9.2	Sum of weights after adjustment					
	a. Household with child 1st procedure	21,393	11,551	21,524	24,213	19,267
	b. Household w/o child 1st proc.-Sect. G comp.	231,547	318,149	221,710	229,890	189,385
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3	Sum of weights	252,940	329,700	243,234	254,103	208,652
9.4	Number HHs w/child 1st procedure	92	59	105	113	104
9.5	Number HHs w/o child 1st proc.–Sect. G comp	579	757	590	620	580
9.6	Coefficient of variation	37.07	41.06	49.84	55.21	50.67

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
1. Base weight					
1.1 Sample size	4,246	4,230	3,765	6,353	11,902
1.2 Sum of weight	406,100	441,700	329,700	379,600	295,300
1.3 Coefficient of variation (CV)	8.70	5.30	3.02	2.45	2.63
2. Refusal conversion subsampling adjustment					
2.1 Sum of weights before adjustment					
a. Household never refused	310,071	355,456	254,982	313,897	227,670
b. Household refused – selected for ref. conversion	60,675	54,301	46,284	27,395	30,017
c. Household refused – not selected for ref. conversion	35,353	31,943	28,434	38,308	37,613
2.2 Sum of weights after adjustment					
a. Household never refused	310,071	355,456	254,982	313,897	227,670
b. Household refused – selected for ref. conversion	96,029	86,244	74,718	65,703	67,630
c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3 Sum of weights	406,100	441,700	329,700	379,600	295,300
2.4 Sample size	3,877	3,924	3,441	5,712	10,387
2.5 Coefficient of variation	23.24	20.98	21.80	36.44	35.93
3. Adjusting for unknown residential					
3.1 Sum of weights before adjustment					
a. Residential – respondent	96,680	93,521	83,015	65,416	65,552
b. Residential – nonrespondent	60,652	59,180	49,561	41,875	45,238
c. Unknown residential status (NA, NM)	37,228	47,580	27,379	37,351	29,487
d. Nonresidential	211,540	241,419	169,745	234,958	155,024

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
3.2	Sum of weights – allocating unknown residential					
	a. Residential – respondent	96,680	93,521	83,015	65,416	65,552
	b. Residential – nonrespondent	60,652	59,180	49,561	41,875	45,238
	c. (NA, NM)	16,296	25,766	14,198	18,162	15,797
3.3	Sum of weights before adjustment					
	a. Residential – respondent	96,680	93,521	83,015	65,416	65,552
	b. Residential – nonrespondent	76,931	84,944	63,691	60,075	61,017
	c. Estimated residential among unknown	0	0	0	0	0
3.3	Sum of weights	173,611	178,464	146,706	125,491	126,569
3.4	Sample size	1,323	1,187	1,212	1,220	3,101
3.5	Coefficient of Variation	36.43	40.81	35.65	63.22	53.19
4.	Supplemental list-sample eligibility adjustment					
4.1	Sum of weights before adjustment					
	a. RDD – household	172,154	177,509	146,118	125,048	126,320
	b. List eligible household	219	190	42	42	14
	c. List ineligible household	352	377	269	161	77
	d. Household with unknown list eligibility	885	389	277	241	158
4.2	Sum of weights before adjustment					
	a. RDD household	172,154	177,509	146,118	125,048	126,320
	b. List eligible household	540	455	96	101	32
	c. List ineligible household	849	905	653	393	182
	d. Household with unknown list eligibility	0	0	0	0	0
4.3	Sum of weights	173,543	178,869	146,867	125,542	126,535
4.4	Sample size	1,304	1,180	1,207	1,215	3,094
4.5	Coefficient of variation (CV)	34.71	39.79	35.15	62.85	53.06

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
5. Screener nonresponse adjustment					
5.1 Sum of weights before adjustment					
a. Respondents	96,561	93,410	82,713	65,315	65,443
b. Nonrespondents	76,982	85,459	64,154	60,227	61,092
5.2 Sum of weights after adjustment					
a. Respondents	173,112	178,478	146,322	125,427	126,385
b. Nonrespondents	0	0	0	0	0
5.3 Sum of weights	173,112	178,478	146,322	125,427	126,385
5.4 Sample size	863	787	823	857	2,110
5.5 Coefficient of variation	26.82	27.29	26.98	50.53	42.98
6. Multiple telephone adjustment					
6.1 Sum of weights before adjustment	173,112	178,478	146,322	125,427	126,385
6.2 Overall adjustment factor	0.98	0.96	0.97	0.97	0.97
6.3 Sum of weights	169,355	172,155	142,564	121,355	122,830
6.4 Sample size	863	787	823	857	2,110
6.5 Coefficient of variation	29.25	30.36	28.94	51.27	44.50
7. Supplemental child sample subsampling adjustment					
7.1 Sum of weights before adjustment					
a. Households-child supp. w/child-RDD w/o child	162,331	160,701	136,637	118,786	121,587
b. Households-child supp. w/o children	7,024	11,454	5,927	2,569	1,243
7.2 Sum of weights after adjustment					
a. Households-child supp. w/child-RDD w/o child	169,355	172,155	142,564	121,355	122,830
b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
7.3	Sum of weights	169,355	172,155	142,564	121,355	122,830
7.4	Sample size	818	734	777	827	2,078
7.5	Coefficient of variation	29.53	31.80	28.68	50.74	44.29
8.	Household poststratification					
8.1	Number of completed screeners	818	734	777	827	2,078
8.2	Sum of weights before adjustment	169,355	172,155	142,564	121,355	122,830
8.3	Overall adjustment factor	1.07	1.00	1.02	1.13	1.06
8.4	Sum of weights	181,629	172,403	145,146	136,622	130,403
8.5	Coefficient of variation	29.25	32.02	28.58	51.03	44.68
9.	Section G nonresponse adjustment					
9.1	Sum of weights before adjustment					
	a. Household with child 1st procedure	19,135	12,407	16,459	11,406	11,663
	b. Household w/o child 1st proc.-Sect. G comp.	96,377	105,769	81,112	71,310	69,315
	c. Household w/o child 1st proc.-Sect. G not comp.	66,117	54,228	47,575	53,906	49,425
9.2	Sum of weights after adjustment					
	a. Household with child 1st procedure	19,135	12,407	16,459	11,406	11,663
	b. Household w/o child 1st proc.-Sect. G comp.	162,494	159,996	128,687	125,216	118,740
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3	Sum of weights	181,629	172,403	145,146	136,622	130,403
9.4	Number HHs w/child 1st procedure	87	50	87	78	174
9.5	Number HHs w/o child 1st proc.–Sect. G comp	447	467	443	442	1156
9.6	Coefficient of variation	34.38	33.26	31.26	54.03	46.02

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer
1. Base weight					
1.1 Sample size	4,952	4,874	38,025	4,507	4,837
1.2 Sum of weight	295,700	262,400	328,400	237,400	293,100
1.3 Coefficient of variation (CV)	3.87	4.08	1.00	4.40	4.53
2. Refusal conversion subsampling adjustment					
2.1 Sum of weights before adjustment					
a. Household never refused	244,285	216,835	276,754	194,779	229,319
b. Household refused – selected for ref. conversion	26,223	26,713	24,855	23,332	33,902
c. Household refused – not selected for ref. conversion	25,192	18,852	26,792	19,289	29,880
2.2 Sum of weights after adjustment					
a. Household never refused	244,285	216,835	276,754	194,779	229,319
b. Household refused – selected for ref. conversion	51,415	45,565	51,646	42,621	63,781
c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3 Sum of weights	295,700	262,400	328,400	237,400	293,100
2.4 Sample size	4,530	4,523	34,921	4,141	4,344
2.5 Coefficient of variation	27.40	21.87	27.53	24.89	28.39
3. Adjusting for unknown residential					
3.1 Sum of weights before adjustment					
a. Residential – respondent	59,499	52,452	51,628	48,645	60,732
b. Residential – nonrespondent	33,632	28,986	35,862	26,656	39,503
c. Unknown residential status (NA, NM)	21,033	24,701	34,021	20,522	30,291
d. Nonresidential	181,536	156,261	206,888	141,578	162,574

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Tulare	Santa Cruz	Marin	San Luis Obispo	Placer
3.2	Sum of weights – allocating unknown residential					
	a. Residential – respondent	59,499	52,452	51,628	48,645	60,732
	b. Residential – nonrespondent	33,632	28,986	35,862	26,656	39,503
	c. (NA, NM)	10,327	13,232	17,865	11,122	15,123
3.3	Sum of weights before adjustment					
	a. Residential – respondent	59,499	52,452	51,628	48,645	60,732
	b. Residential – nonrespondent	43,967	42,185	53,724	37,747	54,616
	c. Estimated residential among unknown	0	0	0	0	0
3.3	Sum of weights	103,466	94,636	105,352	86,392	115,347
3.4	Sample size	1,181	1,197	7,311	1,105	1,225
3.5	Coefficient of Variation	46.81	44.14	53.35	46.37	47.64
4.	Supplemental list-sample eligibility adjustment					
4.1	Sum of weights before adjustment					
	a. RDD – household	103,369	94,451	105,191	86,129	114,941
	b. List eligible household	0	30	38	0	75
	c. List ineligible household	42	60	52	171	54
	d. Household with unknown list eligibility	55	95	70	92	277
4.2	Sum of weights before adjustment					
	a. RDD household	103,369	94,451	105,191	86,129	114,941
	b. List eligible household		71	92		177
	c. List ineligible household	102	142	123	417	128
	d. Household with unknown list eligibility	0	0	0	0	0
4.3	Sum of weights	103,471	94,664	105,407	86,546	115,245
4.4	Sample size	1,179	1,194	7,303	1,101	1,217
4.5	Coefficient of variation (CV)	46.61	43.86	53.25	45.98	47.07

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer
5. Screener nonresponse adjustment					
5.1 Sum of weights before adjustment					
a. Respondents	59,457	52,433	51,622	48,474	60,657
b. Nonrespondents	44,014	42,231	53,785	38,072	54,588
5.2 Sum of weights after adjustment					
a. Respondents	103,369	94,715	105,410	86,129	115,441
b. Nonrespondents	0	0	0	0	0
5.3 Sum of weights	103,369	94,715	105,410	86,129	115,441
5.4 Sample size	838	844	4,935	785	841
5.5 Coefficient of variation	35.91	30.36	40.88	31.22	35.64
6. Multiple telephone adjustment					
6.1 Sum of weights before adjustment	103,369	94,715	105,410	86,129	115,441
6.2 Overall adjustment factor	0.98	0.94	0.93	0.97	0.96
6.3 Sum of weights	101,641	89,447	98,403	83,233	111,344
6.4 Sample size	838	844	4,935	785	841
6.5 Coefficient of variation	37.21	35.30	46.05	34.44	38.81
7. Supplemental child sample subsampling adjustment					
7.1 Sum of weights before adjustment					
a. Households-child supp. w/child-RDD w/o child	99,253	87,627	97,790	80,843	107,157
b. Households-child supp. w/o children	2,389	1,819	613	2,390	4,187
7.2 Sum of weights after adjustment					
a. Households-child supp. w/child-RDD w/o child	101,641	89,447	98,403	83,233	111,344
b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Tulare	Santa Cruz	Marin	San Luis Obispo	Placer
7.3	Sum of weights	101,641	89,447	98,403	83,233	111,344
7.4	Sample size	815	822	4,888	756	795
7.5	Coefficient of variation	37.84	35.16	45.77	34.53	37.92
8.	Household poststratification					
8.1	Number of completed screeners	815	822	4,888	756	795
8.2	Sum of weights before adjustment	101,641	89,447	98,403	83,233	111,344
8.3	Overall adjustment factor	1.09	1.02	1.02	1.11	0.84
8.4	Sum of weights	110,385	91,139	100,650	92,739	93,382
8.5	Coefficient of variation	37.87	35.58	45.68	34.58	37.93
9.	Section G nonresponse adjustment					
9.1	Sum of weights before adjustment					
	a. Household with child 1st procedure	12,774	8,245	7,028	6,730	8,020
	b. Household w/o child 1st proc.-Sect. G comp.	57,636	53,428	59,428	57,032	51,694
	c. Household w/o child 1st proc.-Sect. G not comp.	39,975	29,467	34,194	28,976	33,667
9.2	Sum of weights after adjustment					
	a. Household with child 1st procedure	12,774	8,245	7,028	6,730	8,020
	b. Household w/o child 1st proc.-Sect. G comp.	97,611	82,894	93,622	86,009	85,362
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3	Sum of weights	110,385	91,139	100,650	92,739	93,382
9.4	Number HHs w/child 1st procedure	95	79	343	57	65
9.5	Number HHs w/o child 1st proc.–Sect. G comp	437	486	2990	471	457
9.6	Coefficient of variation	44.40	38.56	48.39	36.31	39.78

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Merced	Butte	Shasta	Yolo	El Dorado
1.	Base weight					
1.1	Sample size	3,891	3,021	3,394	4,104	4,228
1.2	Sum of weight	122,800	161,000	151,600	141,900	159,900
1.3	Coefficient of variation (CV)	2.39	3.81	3.71	4.76	3.09
2.	Refusal conversion subsampling adjustment					
2.1	Sum of weights before adjustment					
	a. Household never refused	96,674	122,767	119,947	117,221	123,491
	b. Household refused – selected for ref. conversion	16,340	23,574	19,903	13,161	18,854
	c. Household refused – not selected for ref. conversion	9,787	14,659	11,750	11,518	17,556
2.2	Sum of weights after adjustment					
	a. Household never refused	96,674	122,767	119,947	117,221	123,491
	b. Household refused – selected for ref. conversion	26,126	38,233	31,653	24,679	36,409
	c. Household refused – not selected for ref. conversion	0	0	0	0	0
2.3	Sum of weights	122,800	161,000	151,600	141,900	159,900
2.4	Sample size	3,581	2,746	3,131	3,772	3,764
2.5	Coefficient of variation	19.89	22.31	19.94	25.72	29.16
3.	Adjusting for unknown residential					
3.1	Sum of weights before adjustment					
	a. Residential – respondent	29,773	45,276	39,221	31,021	35,647
	b. Residential – nonrespondent	19,192	22,011	17,444	17,063	21,110
	c. Unknown residential status (NA, NM)	9,130	14,713	12,886	14,350	16,438
	d. Nonresidential	64,705	79,000	82,049	79,466	86,705

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Merced	Butte	Shasta	Yolo	El Dorado
3.2	Sum of weights – allocating unknown residential					
	a. Residential – respondent	29,773	45,276	39,221	31,021	35,647
	b. Residential – nonrespondent	19,192	22,011	17,444	17,063	21,110
	c. (NA, NM)	5,056	7,759	6,627	7,083	8,823
3.3	Sum of weights before adjustment					
	a. Residential – respondent	29,773	45,276	39,221	31,021	35,647
	b. Residential – nonrespondent	24,256	29,801	24,079	24,135	29,950
	c. Estimated residential among unknown	0	0	0	0	0
3.3	Sum of weights	54,029	75,078	63,300	55,156	65,596
3.4	Sample size	1,271	1,008	1,034	1,093	1,085
3.5	Coefficient of Variation	33.03	36.83	37.50	47.74	47.46
4.	Supplemental list-sample eligibility adjustment					
4.1	Sum of weights before adjustment					
	a. RDD – household	53,943	74,680	63,171	54,784	65,445
	b. List eligible household	0	0	0	27	0
	c. List ineligible household	42	202	86	142	45
	d. Household with unknown list eligibility	44	196	43	203	107
4.2	Sum of weights before adjustment					
	a. RDD household	53,943	74,680	63,171	54,784	65,445
	b. List eligible household				63	
	c. List ineligible household	105	470	209	341	105
	d. Household with unknown list eligibility	0	0	0	0	0
4.3	Sum of weights	54,048	75,150	63,380	55,188	65,550
4.4	Sample size	1,269	1,004	1,033	1,086	1,081
4.5	Coefficient of variation (CV)	33.19	38.38	37.26	47.26	47.21

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Merced	Butte	Shasta	Yolo	El Dorado
5. Screener nonresponse adjustment					
5.1 Sum of weights before adjustment					
a. Respondents	29,731	45,074	39,062	30,914	35,564
b. Nonrespondents	24,318	30,075	24,317	24,274	29,986
5.2 Sum of weights after adjustment					
a. Respondents	53,943	74,680	63,171	54,912	65,445
b. Nonrespondents	0	0	0	0	0
5.3 Sum of weights	53,943	74,680	63,171	54,912	65,445
5.4 Sample size	844	729	769	784	760
5.5 Coefficient of variation	21.90	25.87	24.71	30.26	35.93
6. Multiple telephone adjustment					
6.1 Sum of weights before adjustment	53,943	74,680	63,171	54,912	65,445
6.2 Overall adjustment factor	0.98	0.98	0.97	0.97	0.95
6.3 Sum of weights	53,024	73,218	61,551	53,384	62,490
6.4 Sample size	844	729	769	784	760
6.5 Coefficient of variation	23.93	28.19	26.91	32.77	39.60
7. Supplemental child sample subsampling adjustment					
7.1 Sum of weights before adjustment					
a. Households-child supp. w/child-RDD w/o child	52,173	71,800	59,844	52,327	60,914
b. Households-child supp. w/o children	851	1,419	1,707	1,057	1,576
7.2 Sum of weights after adjustment					
a. Households-child supp. w/child-RDD w/o child	53,024	73,218	61,551	53,384	62,490
b. Households-child supp. w/o children	0	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Merced	Butte	Shasta	Yolo	El Dorado
7.3	Sum of weights	53,024	73,218	61,551	53,384	62,490
7.4	Sample size	828	712	745	767	736
7.5	Coefficient of variation	23.90	28.21	26.82	33.07	39.88
8.	Household poststratification					
8.1	Number of completed screeners	828	712	745	767	736
8.2	Sum of weights before adjustment	53,024	73,218	61,551	53,384	62,490
8.3	Overall adjustment factor	1.20	1.09	1.03	1.11	0.94
8.4	Sum of weights	63,815	79,566	63,426	59,375	58,939
8.5	Coefficient of variation	23.90	28.40	27.24	34.37	39.88
9.	Section G nonresponse adjustment					
9.1	Sum of weights before adjustment					
	a. Household with child 1st procedure	7,513	4,158	4,754	5,223	4,899
	b. Household w/o child 1st proc.-Sect. G comp.	35,861	50,684	40,046	35,226	34,023
	c. Household w/o child 1st proc.-Sect. G not comp.	20,440	24,724	18,626	18,926	20,017
9.2	Sum of weights after adjustment					
	a. Household with child 1st procedure	7,513	4,158	4,754	5,223	4,899
	b. Household w/o child 1st proc.-Sect. G comp.	56,302	75,408	58,672	54,152	54,040
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0	0
9.3	Sum of weights	63,815	79,566	63,426	59,375	58,939
9.4	Number HHs w/child 1st procedure	96	35	49	77	63
9.5	Number HHs w/o child 1st proc.–Sect. G comp	471	460	482	454	433
9.6	Coefficient of variation	28.83	30.03	28.99	38.56	46.45

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Imperial	Napa	Kings	Madera
1. Base weight				
1.1 Sample size	4,692	5,299	3,750	3,961
1.2 Sum of weight	95,300	116,200	71,400	88,200
1.3 Coefficient of variation (CV)	2.30	2.15	2.15	4.00
2. Refusal conversion subsampling adjustment				
2.1 Sum of weights before adjustment				
a. Household never refused	78,023	93,824	54,726	68,939
b. Household refused – selected for ref. conversion	7,588	11,023	10,525	11,472
c. Household refused – not selected for ref. conversion	9,689	11,353	6,149	7,789
2.2 Sum of weights after adjustment				
a. Household never refused	78,023	93,824	54,726	68,939
b. Household refused – selected for ref. conversion	17,277	22,376	16,674	19,261
c. Household refused – not selected for ref. conversion	0	0	0	0
2.3 Sum of weights	95,300	116,200	71,400	88,200
2.4 Sample size	4,215	4,781	3,427	3,611
2.5 Coefficient of variation	33.73	29.20	20.04	22.44
3. Adjusting for unknown residential				
3.1 Sum of weights before adjustment				
a. Residential – respondent	18,494	20,801	18,212	20,900
b. Residential – nonrespondent	13,519	15,961	10,497	12,590
c. Unknown residential status (NA, NM)	8,838	14,119	5,108	6,031
d. Nonresidential	54,449	65,319	37,584	48,679

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Imperial	Napa	Kings	Madera
3.2	Sum of weights – allocating unknown residential				
	a. Residential – respondent	18,494	20,801	18,212	20,900
	b. Residential – nonrespondent	13,519	15,961	10,497	12,590
	c. (NA, NM)	4,068	7,204	2,304	2,893
3.3	Sum of weights before adjustment				
	a. Residential – respondent	18,494	20,801	18,212	20,900
	b. Residential – nonrespondent	17,589	23,158	12,803	15,483
	c. Estimated residential among unknown	0	0	0	0
3.3	Sum of weights	36,083	43,958	31,015	36,383
3.4	Sample size	1,150	1,202	1,219	1,179
3.5	Coefficient of Variation	52.58	52.73	31.51	34.89
4.	Supplemental list-sample eligibility adjustment				
4.1	Sum of weights before adjustment				
	a. RDD – household	36,062	43,890	31,003	36,383
	b. List eligible household	0	22	0	0
	c. List ineligible household	10	33	0	0
	d. Household with unknown list eligibility	11	13	12	0
4.2	Sum of weights before adjustment				
	a. RDD household	36,062	43,890	31,003	36,383
	b. List eligible household	0	54	0	0
	c. List ineligible household	22	81	0	0
	d. Household with unknown list eligibility	0	0	0	0
4.3	Sum of weights	36,084	44,025	31,003	36,383
4.4	Sample size	1,149	1,201	1,218	1,179
4.5	Coefficient of variation (CV)	52.48	52.57	31.48	34.89

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Imperial	Napa	Kings	Madera
5. Screener nonresponse adjustment				
5.1 Sum of weights before adjustment				
a. Respondents	18,485	20,800	18,212	20,900
b. Nonrespondents	17,600	23,225	12,791	15,483
5.2 Sum of weights after adjustment				
a. Respondents	36,062	44,002	31,003	36,383
b. Nonrespondents	0	0	0	0
5.3 Sum of weights	36,062	44,002	31,003	36,383
5.4 Sample size	762	783	840	812
5.5 Coefficient of variation	41.18	41.03	24.89	26.66
6. Multiple telephone adjustment				
6.1 Sum of weights before adjustment	36,062	44,002	31,003	36,383
6.2 Overall adjustment factor	0.98	0.96	0.98	0.98
6.3 Sum of weights	35,360	42,239	30,419	35,833
6.4 Sample size	762	783	840	812
6.5 Coefficient of variation	43.21	43.64	26.90	28.27
7. Supplemental child sample subsampling adjustment				
7.1 Sum of weights before adjustment				
a. Households-child supp. w/child-RDD w/o child	35,136	41,611	30,163	35,412
b. Households-child supp. w/o children	224	628	256	421
7.2 Sum of weights after adjustment				
a. Households-child supp. w/child-RDD w/o child	35,360	42,239	30,419	35,833
b. Households-child supp. w/o children	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Imperial	Napa	Kings	Madera
7.3	Sum of weights	35,360	42,239	30,419	35,833
7.4	Sample size	756	770	832	801
7.5	Coefficient of variation	43.25	44.27	26.97	28.34
8.	Household poststratification				
8.1	Number of completed screeners	756	770	832	801
8.2	Sum of weights before adjustment	35,360	42,239	30,419	35,833
8.3	Overall adjustment factor	1.11	1.07	1.13	1.01
8.4	Sum of weights	39,384	45,402	34,418	36,155
8.5	Coefficient of variation	43.73	44.30	27.12	30.49
9.	Section G nonresponse adjustment				
9.1	Sum of weights before adjustment				
	a. Household with child 1st procedure	4,532	4,078	3,796	3,789
	b. Household w/o child 1st proc.-Sect. G comp.	20,899	25,940	17,806	19,095
	c. Household w/o child 1st proc.-Sect. G not comp.	13,954	15,384	12,816	13,271
9.2	Sum of weights after adjustment				
	a. Household with child 1st procedure	4,532	4,078	3,796	3,789
	b. Household w/o child 1st proc.-Sect. G comp.	34,852	41,324	30,622	32,366
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0
9.3	Sum of weights	39,384	45,402	34,418	36,155
9.4	Number HHs w/child 1st procedure	77	71	89	73
9.5	Number HHs w/o child 1st proc.–Sect. G comp	405	457	442	446
9.6	Coefficient of variation	44.57	46.73	30.83	33.83

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Monterey	Humboldt	Nevada	Mendocino
1. Base weight				
1.1 Sample size	7,705	6,288	2,942	3,266
1.2 Sum of weight	360,800	129,000	101,000	86,700
1.3 Coefficient of variation (CV)	3.11	2.35	2.66	3.18
2. Refusal conversion subsampling adjustment				
2.1 Sum of weights before adjustment				
a. Household never refused	303,429	107,699	77,660	70,888
b. Household refused – selected for ref. conversion	23,727	10,411	14,386	9,989
c. Household refused – not selected for ref. conversion	33,644	10,889	8,954	5,823
2.2 Sum of weights after adjustment				
a. Household never refused	303,429	107,699	77,660	70,888
b. Household refused – selected for ref. conversion	57,371	21,301	23,340	15,812
c. Household refused – not selected for ref. conversion	0	0	0	0
2.3 Sum of weights	360,800	129,000	101,000	86,700
2.4 Sample size	6,986	5,757	2,681	3,047
2.5 Coefficient of variation	35.33	27.74	21.23	18.63
3. Adjusting for unknown residential				
3.1 Sum of weights before adjustment				
a. Residential – respondent	54,792	28,723	24,376	18,129
b. Residential – nonrespondent	42,035	12,947	14,404	11,839
c. Unknown residential status (NA, NM)	37,266	10,436	12,018	9,181
d. Nonresidential	226,707	76,894	50,203	47,551

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Monterey	Humboldt	Nevada	Mendocino
3.2	Sum of weights – allocating unknown residential				
	a. Residential – respondent	54,792	28,723	24,376	18,129
	b. Residential – nonrespondent	42,035	12,947	14,404	11,839
	c. (NA, NM)	18,530	5,456	6,549	5,093
3.3	Sum of weights before adjustment				
	a. Residential – respondent	54,792	28,723	24,376	18,129
	b. Residential – nonrespondent	60,552	18,407	20,965	16,937
	c. Estimated residential among unknown	0	0	0	0
3.3	Sum of weights	115,344	47,130	45,341	35,066
3.4	Sample size	1,432	1,558	895	923
3.5	Coefficient of Variation	66.25	51.84	39.40	38.61
4.	Supplemental list-sample eligibility adjustment				
4.1	Sum of weights before adjustment				
	a. RDD – household	114,991	47,083	45,223	34,966
	b. List eligible household	183	0	0	0
	c. List ineligible household	34	16	42	26
	d. Household with unknown list eligibility	136	31	75	74
4.2	Sum of weights before adjustment				
	a. RDD household	114,991	47,083	45,223	34,966
	b. List eligible household	438	0	0	0
	c. List ineligible household	79	37	99	64
	d. Household with unknown list eligibility	0	0	0	0
4.3	Sum of weights	115,508	47,120	45,322	35,030
4.4	Sample size	1,428	1,556	892	921
4.5	Coefficient of variation (CV)	65.85	51.78	39.15	38.52

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Monterey	Humboldt	Nevada	Mendocino
5. Screener nonresponse adjustment				
5.1 Sum of weights before adjustment				
a. Respondents	54,966	28,707	24,278	18,059
b. Nonrespondents	60,542	18,413	21,045	16,971
5.2 Sum of weights after adjustment				
a. Respondents	116,379	47,083	45,223	34,966
b. Nonrespondents	0	0	0	0
5.3 Sum of weights	116,379	47,083	45,223	34,966
5.4 Sample size	973	1,195	613	619
5.5 Coefficient of variation	46.76	35.53	25.12	24.87
6. Multiple telephone adjustment				
6.1 Sum of weights before adjustment	116,379	47,083	45,223	34,966
6.2 Overall adjustment factor	0.97	0.97	0.96	0.95
6.3 Sum of weights	113,270	45,598	43,211	33,361
6.4 Sample size	973	1,195	613	619
6.5 Coefficient of variation	48.92	37.74	29.68	29.53
7. Supplemental child sample subsampling adjustment				
7.1 Sum of weights before adjustment				
a. Households-child supp. w/child-RDD w/o child	110,220	45,055	42,228	32,813
b. Households-child supp. w/o children	3,050	543	983	548
7.2 Sum of weights after adjustment				
a. Households-child supp. w/child-RDD w/o child	113,270	45,598	43,211	33,361
b. Households-child supp. w/o children	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Monterey	Humboldt	Nevada	Mendocino
7.3	Sum of weights	113,270	45,598	43,211	33,361
7.4	Sample size	941	1,178	597	608
7.5	Coefficient of variation	49.86	37.73	29.81	29.79
8.	Household poststratification				
8.1	Number of completed screeners	941	1,178	597	608
8.2	Sum of weights before adjustment	113,270	45,598	43,211	33,361
8.3	Overall adjustment factor	1.07	1.12	0.85	1.00
8.4	Sum of weights	121,236	51,238	36,894	33,266
8.5	Coefficient of variation	49.69	37.82	29.73	29.78
9.	Section G nonresponse adjustment				
9.1	Sum of weights before adjustment				
	a. Household with child 1st procedure	11,571	3,136	2,090	2,471
	b. Household w/o child 1st proc.-Sect. G comp.	63,391	33,720	24,237	21,070
	c. Household w/o child 1st proc.-Sect. G not comp.	46,274	14,382	10,567	9,725
9.2	Sum of weights after adjustment				
	a. Household with child 1st procedure	11,571	3,136	2,090	2,471
	b. Household w/o child 1st proc.-Sect. G comp.	109,665	48,102	34,804	30,795
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0
9.3	Sum of weights	121,236	51,238	36,894	33,266
9.4	Number HHs w/child 1st procedure	95	68	34	45
9.5	Number HHs w/o child 1st proc.–Sect. G comp	504	802	396	395
9.6	Coefficient of variation	52.38	37.34	30.26	30.95

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Sutter	Yuba	Lake	San Benito
1. Base weight				
1.1 Sample size	3,480	3,092	3,107	3,825
1.2 Sum of weight	60,000	51,600	60,500	37,100
1.3 Coefficient of variation (CV)	2.83	2.20	1.90	1.62
2. Refusal conversion subsampling adjustment				
2.1 Sum of weights before adjustment				
a. Household never refused	46,118	40,382	48,610	29,086
b. Household refused – selected for ref. conversion	4,877	5,968	6,968	3,640
c. Household refused – not selected for ref. conversion	9,005	5,250	4,921	4,374
2.2 Sum of weights after adjustment				
a. Household never refused	46,118	40,382	48,610	29,086
b. Household refused – selected for ref. conversion	13,882	11,218	11,890	8,014
c. Household refused – not selected for ref. conversion	0	0	0	0
2.3 Sum of weights	60,000	51,600	60,500	37,100
2.4 Sample size	2,958	2,778	2,855	3,374
2.5 Coefficient of variation	46.62	26.94	21.82	33.65
3. Adjusting for unknown residential				
3.1 Sum of weights before adjustment				
a. Residential – respondent	14,871	12,056	13,995	8,032
b. Residential – nonrespondent	9,488	6,904	7,939	4,940
c. Unknown residential status (NA, NM)	4,916	3,833	6,102	3,483
d. Nonresidential	30,724	28,808	32,464	20,645

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Sutter	Yuba	Lake	San Benito
3.2	Sum of weights – allocating unknown residential				
	a. Residential – respondent	14,871	12,056	13,995	8,032
	b. Residential – nonrespondent	9,488	6,904	7,939	4,940
	c. (NA, NM)	2,480	2,058	3,548	1,681
3.3	Sum of weights before adjustment				
	a. Residential – respondent	14,871	12,056	13,995	8,032
	b. Residential – nonrespondent	11,970	8,966	11,487	6,621
	c. Estimated residential among unknown	0	0	0	0
3.3	Sum of weights	26,842	21,021	25,482	14,653
3.4	Sample size	933	860	895	938
3.5	Coefficient of Variation	62.92	42.28	41.64	53.19
4.	Supplemental list-sample eligibility adjustment				
4.1	Sum of weights before adjustment				
	a. RDD – household	26,810	20,988	25,482	14,647
	b. List eligible household	12	0	0	0
	c. List ineligible household	8	9	0	7
	d. Household with unknown list eligibility	12	24	0	0
4.2	Sum of weights before adjustment				
	a. RDD household	26,810	20,988	25,482	14,647
	b. List eligible household	29	0	0	0
	c. List ineligible household	18	21	0	16
	d. Household with unknown list eligibility	0	0	0	0
4.3	Sum of weights	26,857	21,009	25,482	14,662
4.4	Sample size	932	859	895	938
4.5	Coefficient of variation (CV)	62.76	42.23	41.64	53.13

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Sutter	Yuba	Lake	San Benito
5. Screener nonresponse adjustment				
5.1 Sum of weights before adjustment				
a. Respondents	14,880	12,047	13,961	8,026
b. Nonrespondents	11,976	8,962	11,521	6,637
5.2 Sum of weights after adjustment				
a. Respondents	26,918	20,988	25,482	14,647
b. Nonrespondents	0	0	0	0
5.3 Sum of weights	26,918	20,988	25,482	14,647
5.4 Sample size	666	611	629	668
5.5 Coefficient of variation	53.60	30.64	26.74	40.79
6. Multiple telephone adjustment				
6.1 Sum of weights before adjustment	26,918	20,988	25,482	14,647
6.2 Overall adjustment factor	0.98	0.98	0.97	0.97
6.3 Sum of weights	26,311	20,502	24,642	14,139
6.4 Sample size	666	611	629	668
6.5 Coefficient of variation	54.99	32.57	29.03	43.17
7. Supplemental child sample subsampling adjustment				
7.1 Sum of weights before adjustment				
a. Households-child supp. w/child-RDD w/o child	26,034	20,363	24,305	14,056
b. Households-child supp. w/o children	278	139	338	83
7.2 Sum of weights after adjustment				
a. Households-child supp. w/child-RDD w/o child	26,311	20,502	24,642	14,139
b. Households-child supp. w/o children	0	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Sutter	Yuba	Lake	San Benito
7.3	Sum of weights	26,311	20,502	24,642	14,139
7.4	Sample size	657	606	619	663
7.5	Coefficient of variation	55.20	32.73	29.06	43.27
8.	Household poststratification				
8.1	Number of completed screeners	657	606	619	663
8.2	Sum of weights before adjustment	26,311	20,502	24,642	14,139
8.3	Overall adjustment factor	1.03	1.00	0.97	1.12
8.4	Sum of weights	27,033	20,535	23,974	15,885
8.5	Coefficient of variation	55.70	32.66	29.87	43.36
9.	Section G nonresponse adjustment				
9.1	Sum of weights before adjustment				
	a. Household with child 1st procedure	2,069	2,429	1,617	1,972
	b. Household w/o child 1st proc.-Sect. G comp.	14,701	11,926	14,180	7,367
	c. Household w/o child 1st proc.-Sect. G not comp.	10,263	6,180	8,178	6,546
9.2	Sum of weights after adjustment				
	a. Household with child 1st procedure	2,069	2,429	1,617	1,972
	b. Household w/o child 1st proc.-Sect. G comp.	24,964	18,106	22,357	13,913
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0	0
9.3	Sum of weights	27,033	20,535	23,974	15,885
9.4	Number HHs w/child 1st procedure	46	71	35	79
9.5	Number HHs w/o child 1st proc.–Sect. G comp	373	358	380	332
9.6	Coefficient of variation	57.03	35.99	28.58	43.25

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
1. Base weight			
1.1 Sample size	2,978	4,123	3,927
1.2 Sum of weight	78,000	194,800	230,300
1.3 Coefficient of variation (CV)	2.22	1.29	3.01
2. Refusal conversion subsampling adjustment			
2.1 Sum of weights before adjustment			
a. Household never refused	63,043	167,889	190,599
b. Household refused – selected for ref. conversion	8,481	15,054	22,074
c. Household refused – not selected for ref. conversion	6,476	11,857	17,627
2.2 Sum of weights after adjustment			
a. Household never refused	63,043	167,889	190,599
b. Household refused – selected for ref. conversion	14,957	26,911	39,701
c. Household refused – not selected for ref. conversion	0	0	0
2.3 Sum of weights	78,000	194,800	230,300
2.4 Sample size	2,731	3,872	3,627
2.5 Coefficient of variation	23.12	21.16	23.50
3. Adjusting for unknown residential			
3.1 Sum of weights before adjustment			
a. Residential – respondent	18,895	31,341	43,326
b. Residential – nonrespondent	10,314	13,374	25,701
c. Unknown residential status (NA, NM)	7,668	14,066	23,865
d. Nonresidential	41,123	136,020	137,407

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
3.2	Sum of weights – allocating unknown residential			
	a. Residential – respondent	18,895	31,341	43,326
	b. Residential – nonrespondent	10,314	13,374	25,701
	c. (NA, NM)	3,801	7,142	12,761
3.3	Sum of weights before adjustment			
	a. Residential – respondent	18,895	31,341	43,326
	b. Residential – nonrespondent	14,115	20,502	38,439
	c. Estimated residential among unknown	0	0	0
3.3	Sum of weights	33,011	51,843	81,765
3.4	Sample size	887	733	915
3.5	Coefficient of Variation	42.41	47.60	46.61
4.	Supplemental list-sample eligibility adjustment			
4.1	Sum of weights before adjustment			
	a. RDD – household	33,011	51,760	81,604
	b. List eligible household	0	0	27
	c. List ineligible household	0	0	66
	d. Household with unknown list eligibility	0	83	68
4.2	Sum of weights before adjustment			
	a. RDD household	33,011	51,760	81,604
	b. List eligible household	0	0	62
	c. List ineligible household	0	0	163
	d. Household with unknown list eligibility	0	0	0
4.3	Sum of weights	33,011	51,760	81,828
4.4	Sample size	887	732	913
4.5	Coefficient of variation (CV)	42.41	47.64	46.30

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
5. Screener nonresponse adjustment			
5.1 Sum of weights before adjustment			
a. Respondents	18,895	31,341	43,295
b. Nonrespondents	14,115	20,419	38,533
5.2 Sum of weights after adjustment			
a. Respondents	33,011	51,760	81,729
b. Nonrespondents	0	0	0
5.3 Sum of weights	33,011	51,760	81,729
5.4 Sample size	635	564	638
5.5 Coefficient of variation	27.57	34.76	31.29
6. Multiple telephone adjustment			
6.1 Sum of weights before adjustment	33,011	51,760	81,729
6.2 Overall adjustment factor	0.98	0.97	0.97
6.3 Sum of weights	32,373	50,374	79,173
6.4 Sample size	635	564	638
6.5 Coefficient of variation	29.44	36.74	34.41
7. Supplemental child sample subsampling adjustment			
7.1 Sum of weights before adjustment			
a. Households-child supp. w/child-RDD w/o child	31,869	48,823	77,021
b. Households-child supp. w/o children	504	1,551	2,152
7.2 Sum of weights after adjustment			
a. Households-child supp. w/child-RDD w/o child	32,373	50,374	79,173
b. Households-child supp. w/o children	0	0	0

Table B-1. Household weighting for the combined RDD, Korean and Vietnamese surname list samples by stratum (continued)

		Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
7.3	Sum of weights	32,373	50,374	79,173
7.4	Sample size	623	543	618
7.5	Coefficient of variation	29.38	36.95	34.56
8.	Household poststratification			
8.1	Number of completed screeners	623	543	618
8.2	Sum of weights before adjustment	32,373	50,374	79,173
8.3	Overall adjustment factor	1.12	1.14	0.89
8.4	Sum of weights	36,282	57,242	70,168
8.5	Coefficient of variation	29.40	37.00	34.98
9.	Section G nonresponse adjustment			
9.1	Sum of weights before adjustment			
	a. Household with child 1st procedure	3,426	3,175	4,148
	b. Household w/o child 1st proc.-Sect. G comp.	21,710	37,970	42,369
	c. Household w/o child 1st proc.-Sect. G not comp.	11,145	16,097	23,651
9.2	Sum of weights after adjustment			
	a. Household with child 1st procedure	3,426	3,175	4,148
	b. Household w/o child 1st proc.-Sect. G comp.	32,856	54,067	66,020
	c. Household w/o child 1st proc.-Sect. G not comp.	0	0	0
9.3	Sum of weights	36,282	57,242	70,168
9.4	Number HHs w/child 1st procedure	58	27	34
9.5	Number HHs w/o child 1st proc.–Sect. G comp	384	371	379
9.6	Coefficient of variation	33.39	38.40	35.75

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
1. Adult initial weights					
1.1 Number of sampled adults	73,775	16,154	6,979	4,534	2,496
1.2 Sum of weights	24,082,937	6,693,594	2,046,596	2,046,799	1,201,291
1.3 Coefficient of variation (CV)	92.44	79.08	85.15	86.23	67.15
2. Nonresponse adjustment					
2.1 Number of completed interviews	43,020	8,712	3,828	2,493	1,443
2.2 Sum of weights before adjustment	24,082,937	6,693,594	2,046,596	2,046,799	1,201,291
a. Eligible respondents	12,693,316	3,271,570	1,065,725	1,018,521	656,703
b. Ineligible	555,137	211,697	44,871	38,538	25,277
c. Nonrespondents	10,834,484	3,210,327	935,999	989,740	519,311
2.3 Sum of weights after adjustment	24,082,937	6,693,594	2,046,596	2,046,799	1,201,291
a. Eligible respondents	23,008,246	6,274,296	1,959,706	1,966,591	1,155,652
b. Ineligible	1,074,691	419,298	86,890	80,208	45,639
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	98.33	83.94	83.94	89.35	71.65
2.5 Mean adjustment factor	1.90	2.05	1.92	2.01	1.83
3. Trimming adjustment*					
3.1 Number of trimmed records	39	11	3	1	0
3.2 Sum of weights before trimming adjustment	23,008,246	6,274,296	1,959,706	1,966,591	1,155,652
3.3 Sum of weights after trimming adjustment	22,970,493	6,276,050	1,959,800	1,915,077	1,142,067
4. Raking adjustment*					
4.1 Number of completed interviews	43,020	8,722	3,828	2,447	1,445
4.2 Sum of weights after adjustment	26,036,973	7,178,943	2,149,000	2,187,687	1,274,937
4.3 Coefficient of Variation (CV)	114.86	97.93	97.06	107.23	89.16
4.4 Mean adjustment factor	1.13	1.14	1.10	1.14	1.12
4.5 Mean weight	605.23	823.08	561.39	894.03	882.31

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	26,387,565	7,244,412	2,169,794	2,205,768	1,287,482
5.2 Coefficient of Variation (CV)	114.86	97.93	97.06	107.23	89.16
5.3 Mean adjustment factor	1.01	1.01	1.01	1.01	1.01
5.4 Mean weight	613.38	830.59	566.82	901.42	890.99

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
1. Adult initial weights					
1.1 Number of sampled adults	2,367	2,384	2,127	1,873	1,324
1.2 Sum of weights	1,157,226	1,082,823	1,082,798	907,919	693,489
1.3 Coefficient of variation (CV)	59.93	59.35	65.33	60.93	60.43
2. Nonresponse adjustment					
2.1 Number of completed interviews	1,325	1,322	1,318	1,166	843
2.2 Sum of weights before adjustment	1,157,226	1,082,823	1,082,798	907,919	693,489
a. Eligible respondents	596,374	549,880	626,258	517,844	406,883
b. Ineligible	35,062	25,373	22,821	15,113	10,400
c. Nonrespondents	525,790	507,570	433,719	374,962	276,206
2.3 Sum of weights after adjustment	1,157,226	1,082,823	1,082,798	907,919	693,489
a. Eligible respondents	1,085,710	1,031,458	1,042,289	882,245	675,774
b. Ineligible	71,516	51,364	40,509	25,674	17,715
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	65.94	64.66	70.26	66.89	64.31
2.5 Mean adjustment factor	1.94	1.97	1.73	1.75	1.70
3. Trimming adjustment*					
3.1 Number of trimmed records	2	0	2	0	1
3.2 Sum of weights before trimming adjustment	1,085,710	1,031,458	1,042,289	882,245	675,774
3.3 Sum of weights after trimming adjustment	1,089,825	1,038,140	1,013,084	879,551	704,356
4. Raking adjustment*					
4.1 Number of completed interviews	1,331	1,337	1,285	1,171	894
4.2 Sum of weights after adjustment	1,296,807	1,311,083	1,112,446	964,075	747,338
4.3 Coefficient of Variation (CV)	80.41	77.55	87.07	84.60	86.08
4.4 Mean adjustment factor	1.19	1.26	1.10	1.10	1.06
4.5 Mean weight	974.31	980.62	865.72	823.29	835.95

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	1,331,316	1,368,912	1,117,305	979,561	756,645
5.2 Coefficient of Variation (CV)	80.42	77.55	87.07	84.60	86.08
5.3 Mean adjustment factor	1.03	1.04	1.00	1.02	1.01
5.4 Mean weight	1000.24	1023.87	869.50	836.52	846.36

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
1. Adult initial weights					
1.1 Number of sampled adults	1,032	1,292	1,132	1,097	1,074
1.2 Sum of weights	546,209	643,363	531,427	517,260	438,304
1.3 Coefficient of variation (CV)	57.25	63.27	65.30	68.76	72.08
2. Nonresponse adjustment					
2.1 Number of completed interviews	598	777	632	661	606
2.2 Sum of weights before adjustment	546,209	643,363	531,427	517,260	438,304
a. Eligible respondents	294,298	353,814	257,935	295,925	222,915
b. Ineligible	11,605	9,216	9,976	9,945	7,687
c. Nonrespondents	240,307	280,333	263,515	211,390	207,703
2.3 Sum of weights after adjustment	546,209	643,363	531,427	517,260	438,304
a. Eligible respondents	523,519	626,893	511,836	498,591	422,831
b. Ineligible	22,690	16,470	19,591	18,669	15,474
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	58.76	62.17	59.94	70.91	74.29
2.5 Mean adjustment factor	1.86	1.82	2.06	1.75	1.97
3. Trimming adjustment*					
3.1 Number of trimmed records	1	1	1	2	0
3.2 Sum of weights before trimming adjustment	523,519	626,893	511,836	498,591	422,831
3.3 Sum of weights after trimming adjustment	512,453	605,272	528,759	530,809	416,935
4. Raking adjustment*					
4.1 Number of completed interviews	595	763	654	695	601
4.2 Sum of weights after adjustment	594,100	660,333	577,308	547,203	486,421
4.3 Coefficient of Variation (CV)	91.13	77.55	82.33	96.37	87.53
4.4 Mean adjustment factor	1.16	1.09	1.09	1.03	1.17
4.5 Mean weight	998.49	865.44	882.73	787.34	809.35

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	605,191	662,879	582,380	549,978	502,111
5.2 Coefficient of Variation (CV)	91.13	77.55	82.33	96.37	87.53
5.3 Mean adjustment factor	1.02	1.00	1.01	1.01	1.03
5.4 Mean weight	1017.13	868.78	890.49	791.33	835.46

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
1. Adult initial weights					
1.1 Number of sampled adults	818	734	777	827	2,078
1.2 Sum of weights	381,306	331,198	310,568	277,383	277,014
1.3 Coefficient of variation (CV)	52.37	55.79	52.18	67.05	63.14
2. Nonresponse adjustment					
2.1 Number of completed interviews	469	488	467	472	1,216
2.2 Sum of weights before adjustment	381,306	331,198	310,568	277,383	277,014
a. Eligible respondents	197,835	203,514	172,950	145,788	145,958
b. Ineligible	6,249	6,361	3,591	4,944	6,328
c. Nonrespondents	177,221	121,323	134,028	126,651	124,729
2.3 Sum of weights after adjustment	381,306	331,198	310,568	277,383	277,014
a. Eligible respondents	368,274	320,465	303,333	268,016	265,002
b. Ineligible	13,032	10,733	7,235	9,367	12,012
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	60.11	57.89	55.80	67.88	62.06
2.5 Mean adjustment factor	1.93	1.63	1.80	1.90	1.90
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	2	1	0
3.2 Sum of weights before trimming adjustment	368,274	320,465	303,333	268,016	265,002
3.3 Sum of weights after trimming adjustment	366,805	318,694	282,251	266,959	259,667
4. Raking adjustment*					
4.1 Number of completed interviews	471	496	444	469	1,189
4.2 Sum of weights after adjustment	435,735	356,061	344,375	294,171	291,232
4.3 Coefficient of Variation (CV)	94.25	81.71	72.80	90.94	82.08
4.4 Mean adjustment factor	1.19	1.12	1.22	1.10	1.12
4.5 Mean weight	925.13	717.87	775.62	627.23	244.94

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	447,453	357,337	351,746	296,218	293,194
5.2 Coefficient of Variation (CV)	94.25	81.71	72.80	90.94	82.08
5.3 Mean adjustment factor	1.03	1.00	1.02	1.01	1.01
5.4 Mean weight	950.01	720.44	792.22	631.60	246.59

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
1. Adult initial weights						
1.1 Number of sampled adults	815	822	4,888	756	795	828
1.2 Sum of weights	237,928	181,939	186,369	178,579	182,184	134,337
1.3 Coefficient of variation (CV)	58.80	56.01	62.82	54.53	53.75	46.78
2. Nonresponse adjustment						
2.1 Number of completed interviews	473	517	3,111	491	474	490
2.2 Sum of weights before adjustment	237,928	181,939	186,369	178,579	182,184	134,337
a. Eligible respondents	127,191	107,162	108,134	109,334	102,078	75,576
b. Ineligible	6,137	2,768	3,162	2,558	1,923	3,391
c. Nonrespondents	104,600	72,010	75,073	66,687	78,182	55,370
2.3 Sum of weights after adjustment	237,928	181,939	186,369	178,579	182,184	134,337
a. Eligible respondents	225,770	177,243	180,205	174,498	178,638	128,640
b. Ineligible	12,158	4,696	6,164	4,081	3,546	5,697
c. Nonrespondents	0	0	0	0	0	0
2.4 Coefficient of Variation (CV)	63.67	60.39	65.96	60.72	56.40	52.02
2.5 Mean adjustment factor	1.87	1.70	1.72	1.63	1.78	1.78
3. Trimming adjustment*						
3.1 Number of trimmed records	2	0	2	0	0	1
3.2 Sum of weights before trimming adjustment	225,770	177,243	180,205	174,498	178,638	128,640
3.3 Sum of weights after trimming adjustment	234,815	173,371	182,124	174,548	178,295	149,014
4. Raking adjustment*						
4.1 Number of completed interviews	485	502	3,097	490	471	513
4.2 Sum of weights after adjustment	268,407	192,105	188,688	189,973	223,446	158,894
4.3 Coefficient of Variation (CV)	85.49	77.17	95.71	88.41	67.72	91.58
4.4 Mean adjustment factor	1.14	1.11	1.04	1.09	1.25	1.07
4.5 Mean weight	553.42	382.68	60.93	387.70	474.41	309.74

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
5. Benchmark to 2005 DOF provisional estimates adjustment*						
5.1 Sum of weights after adjustment	275,428	192,913	189,470	191,670	231,111	162,802
5.2 Coefficient of Variation (CV)	85.49	77.17	95.71	88.41	67.72	91.58
5.3 Mean adjustment factor	1.03	1.00	1.00	1.01	1.03	1.02
5.4 Mean weight	567.89	384.29	61.18	391.16	490.68	317.35

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
1. Adult initial weights					
1.1 Number of sampled adults	712	745	767	736	756
1.2 Sum of weights	151,721	120,346	119,124	120,269	88,455
1.3 Coefficient of variation (CV)	50.43	51.46	53.54	57.65	61.77
2. Nonresponse adjustment					
2.1 Number of completed interviews	467	502	479	459	426
2.2 Sum of weights before adjustment	151,721	120,346	119,124	120,269	88,455
a. Eligible respondents	91,436	75,938	70,075	71,886	48,243
b. Ineligible	2,413	2,011	1,376	1,860	1,599
c. Nonrespondents	57,872	42,397	47,673	46,524	38,614
2.3 Sum of weights after adjustment	151,721	120,346	119,124	120,269	88,455
a. Eligible respondents	147,540	117,211	116,721	117,102	85,546
b. Ineligible	4,181	3,136	2,402	3,168	2,910
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	52.77	53.86	56.62	59.91	63.90
2.5 Mean adjustment factor	1.66	1.58	1.70	1.67	1.83
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	0	0	0
3.2 Sum of weights before trimming adjustment	147,540	117,211	116,721	117,102	85,546
3.3 Sum of weights after trimming adjustment	148,119	114,058	116,519	118,131	85,364
4. Raking adjustment*					
4.1 Number of completed interviews	476	482	486	461	425
4.2 Sum of weights after adjustment	161,971	135,827	132,095	129,093	106,319
4.3 Coefficient of Variation (CV)	77.92	74.23	85.97	64.52	77.70
4.4 Mean adjustment factor	1.09	1.19	1.13	1.09	1.25
4.5 Mean weight	340.27	281.80	271.80	280.03	250.16

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	163,694	137,620	133,858	131,954	109,581
5.2 Coefficient of Variation (CV)	77.92	74.23	85.97	64.52	77.70
5.3 Mean adjustment factor	1.01	1.01	1.01	1.02	1.03
5.4 Mean weight	343.89	285.52	275.43	286.23	257.84

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
1. Adult initial weights					
1.1 Number of sampled adults	770	832	801	941	1,178
1.2 Sum of weights	87,782	72,304	79,438	257,516	97,020
1.3 Coefficient of variation (CV)	62.37	48.65	55.51	71.71	58.55
2. Nonresponse adjustment					
2.1 Number of completed interviews	477	469	478	539	822
2.2 Sum of weights before adjustment	87,782	72,304	79,438	257,516	97,020
a. Eligible respondents	49,070	37,460	43,445	134,208	62,552
b. Ineligible	1,387	1,125	2,312	4,438	683
c. Nonrespondents	37,325	33,720	33,681	118,871	33,784
2.3 Sum of weights after adjustment	87,782	72,304	79,438	257,516	97,020
a. Eligible respondents	84,863	69,983	75,030	249,251	95,820
b. Ineligible	2,919	2,321	4,408	8,266	1,199
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	67.29	54.70	57.00	78.81	63.78
2.5 Mean adjustment factor	1.79	1.93	1.83	1.92	1.55
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	0	2	0
3.2 Sum of weights before trimming adjustment	84,863	69,983	75,030	249,251	95,820
3.3 Sum of weights after trimming adjustment	87,506	70,575	78,137	252,362	96,587
4. Raking adjustment*					
4.1 Number of completed interviews	490	466	473	554	818
4.2 Sum of weights after adjustment	95,358	88,267	90,814	296,849	98,576
4.3 Coefficient of Variation (CV)	83.92	83.70	86.22	84.88	86.82
4.4 Mean adjustment factor	1.09	1.25	1.16	1.18	1.02
4.5 Mean weight	194.61	189.41	192.00	535.83	120.51

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	96,248	89,839	93,506	297,554	99,042
5.2 Coefficient of Variation (CV)	83.92	83.70	86.22	84.88	86.82
5.3 Mean adjustment factor	1.01	1.02	1.03	1.00	1.00
5.4 Mean weight	196.42	192.79	197.69	537.10	121.08

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
1. Adult initial weights					
1.1 Number of sampled adults	597	608	657	606	619
1.2 Sum of weights	68,781	61,160	55,438	41,189	44,382
1.3 Coefficient of variation (CV)	50.22	49.27	77.12	50.92	51.88
2. Nonresponse adjustment					
2.1 Number of completed interviews	403	417	385	378	384
2.2 Sum of weights before adjustment	68,781	61,160	55,438	41,189	44,382
a. Eligible respondents	43,522	40,100	30,612	23,890	25,252
b. Ineligible	796	956	1,033	1,098	1,107
c. Nonrespondents	24,463	20,104	23,793	16,200	18,024
2.3 Sum of weights after adjustment	68,781	61,160	55,438	41,189	44,382
a. Eligible respondents	67,253	59,629	53,470	39,244	42,132
b. Ineligible	1,528	1,531	1,968	1,945	2,250
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	56.67	49.59	80.32	57.30	55.05
2.5 Mean adjustment factor	1.58	1.53	1.81	1.72	1.76
3. Trimming adjustment*					
3.1 Number of trimmed records	1	0	0	0	2
3.2 Sum of weights before trimming adjustment	67,253	59,629	53,470	39,244	42,132
3.3 Sum of weights after trimming adjustment	67,335	58,535	52,715	38,525	41,221
4. Raking adjustment*					
4.1 Number of completed interviews	402	411	380	357	378
4.2 Sum of weights after adjustment	77,761	67,078	62,190	44,244	47,473
4.3 Coefficient of Variation (CV)	74.94	66.10	101.98	79.56	72.43
4.4 Mean adjustment factor	1.15	1.15	1.18	1.15	1.15
4.5 Mean weight	193.43	163.21	163.66	123.93	125.59

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	78,726	67,305	64,042	45,947	47,952
5.2 Coefficient of Variation (CV)	74.94	66.10	101.98	79.56	72.43
5.3 Mean adjustment factor	1.01	1.00	1.03	1.04	1.01
5.4 Mean weight	195.84	163.76	168.53	128.70	126.86

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
1. Adult initial weights				
1.1 Number of sampled adults	663	623	543	618
1.2 Sum of weights	35,684	72,379	106,756	135,287
1.3 Coefficient of variation (CV)	67.04	58.47	57.75	50.46
2. Nonresponse adjustment				
2.1 Number of completed interviews	351	412	383	397
2.2 Sum of weights before adjustment	35,684	72,379	106,756	135,287
a. Eligible respondents	16,769	44,998	71,399	82,295
b. Ineligible	758	1,844	1,271	2,078
c. Nonrespondents	18,158	25,537	34,086	50,914
2.3 Sum of weights after adjustment	35,684	72,379	106,756	135,287
a. Eligible respondents	33,990	69,394	104,697	131,898
b. Ineligible	1,695	2,985	2,059	3,389
c. Nonrespondents	0	0	0	0
2.4 Coefficient of Variation (CV)	64.72	58.40	62.98	56.17
2.5 Mean adjustment factor	2.13	1.61	1.50	1.64
3. Trimming adjustment*				
3.1 Number of trimmed records	0	0	1	0
3.2 Sum of weights before trimming adjustment	33,990	69,394	104,697	131,898
3.3 Sum of weights after trimming adjustment	35,639	74,361	105,951	130,131
4. Raking adjustment*				
4.1 Number of completed interviews	351	430	396	389
4.2 Sum of weights after adjustment	39,406	78,841	111,594	142,448
4.3 Coefficient of Variation (CV)	103.66	70.08	75.49	72.51
4.4 Mean adjustment factor	1.11	1.06	1.05	1.09
4.5 Mean weight	112.27	183.35	281.80	366.19

Table B-2. Extended interview weighting for adult interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
5. Benchmark to 2005 DOF provisional estimates adjustment*				
5.1 Sum of weights after adjustment	39,718	80,429	112,692	144,782
5.2 Coefficient of Variation (CV)	103.66	70.08	75.49	72.51
5.3 Mean adjustment factor	1.01	1.02	1.01	1.02
5.4 Mean weight	113.16	187.05	284.58	372.19

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
1. Child initial weights					
1.1 Number of sampled children	14,963	2,978	2,374	919	555
1.2 Sum of weights	6,023,115	1,740,559	482,761	489,344	278,196
1.3 Coefficient of Variation (CV)	134.95	127.28	98.57	113.42	91.88
2. Nonresponse adjustment					
2.1 Number of completed child interviews	11,358	2,145	1,797	659	426
2.2 Sum of weights before adjustment	6,023,115	1,740,559	482,761	489,344	278,196
a. Eligible respondents	4,475,666	1,235,386	358,662	354,376	214,037
b. Ineligible	67,903	26,155	3,014	4,705	2,268
c. Nonrespondents	1,479,546	479,018	121,085	130,262	61,892
2.3 Sum of weights after adjustment	6,023,115	1,740,559	482,761	489,344	278,196
a. Eligible respondents	5,932,396	1,704,975	478,640	482,799	275,257
b. Ineligible	90,719	35,584	4,121	6,545	2,939
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	135.08	124.83	96.58	113.87	94.40
2.5 Mean adjustment factor	1.35	1.41	1.35	1.38	1.30
3. Trimming adjustment*					
3.1 Number of trimmed records	39	6	3	1	0
3.2 Sum of weights before trimming adjustment	5,932,396	1,704,975	478,640	482,799	275,257
3.3 Sum of weights after trimming adjustment	5,862,275	1,678,994	478,426	472,538	269,144
4. Raking adjustment*					
4.1 Number of completed child interviews	11,358	2,150	1,797	648	423
4.2 Sum of weights after adjustment	6,311,596	1,845,820	516,032	532,598	302,462
4.3 Coefficient of Variation (CV)	128.10	115.96	99.34	115.67	95.60
4.4 Mean adjustment factor	1.08	1.10	1.08	1.13	1.12
4.5 Mean weight	555.70	858.52	287.16	821.91	715.04

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	6,400,320	1,862,653	521,025	537,000	305,438
5.2 Coefficient of Variation (CV)	128.08	115.96	99.34	115.67	95.60
5.3 Mean adjustment factor	1.01	1.01	1.01	1.01	1.01
5.4 Mean weight	563.51	866.35	289.94	828.70	722.07

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
1. Child initial weights					
1.1 Number of sampled children	530	496	412	354	259
1.2 Sum of weights	366,220	305,627	221,436	215,308	159,101
1.3 Coefficient of Variation (CV)	112.95	106.22	86.47	84.14	77.45
2. Nonresponse adjustment					
2.1 Number of completed child interviews	402	367	314	279	202
2.2 Sum of weights before adjustment	366,220	305,627	221,436	215,308	159,101
a. Eligible respondents	278,852	218,928	164,613	168,284	128,436
b. Ineligible	3,631	6,352	3,686	280	0
c. Nonrespondents	83,737	80,347	53,137	46,743	30,665
2.3 Sum of weights after adjustment	366,220	305,627	221,436	215,308	159,101
a. Eligible respondents	361,570	296,776	216,559	214,988	159,101
b. Ineligible	4,650	8,851	4,877	320	0
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	121.08	87.66	81.99	85.02	81.38
2.5 Mean adjustment factor	1.31	1.40	1.35	1.28	1.24
3. Trimming adjustment*					
3.1 Number of trimmed records	1	1	1	1	0
3.2 Sum of weights before trimming adjustment	361,570	296,776	216,559	214,988	159,101
3.3 Sum of weights after trimming adjustment	351,065	299,421	209,982	214,658	166,756
4. Raking adjustment*					
4.1 Number of completed child interviews	400	370	301	283	217
4.2 Sum of weights after adjustment	366,326	314,990	237,188	237,417	162,858
4.3 Coefficient of Variation (CV)	92.42	94.55	83.40	82.66	80.68
4.4 Mean adjustment factor	1.04	1.05	1.13	1.11	0.98
4.5 Mean weight	915.81	851.32	788.00	838.93	750.50

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	376,074	328,884	238,224	241,231	164,886
5.2 Coefficient of Variation (CV)	92.42	94.55	83.40	82.66	80.68
5.3 Mean adjustment factor	1.03	1.04	1.00	1.02	1.01
5.4 Mean weight	940.19	888.87	791.44	852.41	759.84

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
1. Child initial weights					
1.1 Number of sampled children	222	139	232	208	240
1.2 Sum of weights	155,182	63,509	147,219	108,318	146,771
1.3 Coefficient of Variation (CV)	111.72	89.41	89.57	113.74	113.69
2. Nonresponse adjustment					
2.1 Number of completed child interviews	168	104	175	164	182
2.2 Sum of weights before adjustment	155,182	63,509	147,219	108,318	146,771
a. Eligible respondents	121,486	46,494	114,567	82,182	116,299
b. Ineligible	3,101	0	1,412	554	893
c. Nonrespondents	30,595	17,015	31,240	25,582	29,578
2.3 Sum of weights after adjustment	155,182	63,509	147,219	108,318	146,771
a. Eligible respondents	151,365	63,509	145,329	107,380	145,613
b. Ineligible	3,817	0	1,890	937	1,157
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	121.03	77.84	91.49	98.99	111.97
2.5 Mean adjustment factor	1.28	1.37	1.29	1.32	1.26
3. Trimming adjustment*					
3.1 Number of trimmed records	1	0	1	0	3
3.2 Sum of weights before trimming adjustment	151,365	63,509	145,329	107,380	145,613
3.3 Sum of weights after trimming adjustment	146,789	62,658	148,927	112,319	141,389
4. Raking adjustment*					
4.1 Number of completed child interviews	172	102	179	170	183
4.2 Sum of weights after adjustment	170,285	77,889	140,159	107,081	162,736
4.3 Coefficient of Variation (CV)	77.71	73.02	84.08	125.32	83.22
4.4 Mean adjustment factor	1.16	1.24	0.94	0.95	1.15
4.5 Mean weight	990.03	763.62	783.01	629.89	889.27

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	173,464	78,189	141,391	107,624	167,985
5.2 Coefficient of Variation (CV)	77.71	73.02	84.08	125.32	83.22
5.3 Mean adjustment factor	1.02	1.00	1.01	1.01	1.03
5.4 Mean weight	1008.51	766.56	789.89	633.08	917.95

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
1. Child initial weights					
1.1 Number of sampled children	182	109	178	162	373
1.2 Sum of weights	122,813	70,546	88,838	70,799	68,175
1.3 Coefficient of Variation (CV)	102.97	84.22	78.28	92.34	92.61
2. Nonresponse adjustment					
2.1 Number of completed child interviews	141	83	128	126	296
2.2 Sum of weights before adjustment	122,813	70,546	88,838	70,799	68,175
a. Eligible respondents	95,368	54,817	59,060	53,816	52,526
b. Ineligible	1,283	327	802	602	2,067
c. Nonrespondents	26,162	15,402	28,976	16,382	13,581
2.3 Sum of weights after adjustment	122,813	70,546	88,838	70,799	68,175
a. Eligible respondents	121,089	70,148	87,650	69,907	65,641
b. Ineligible	1,724	398	1,189	892	2,533
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	101.42	85.38	65.60	97.17	84.87
2.5 Mean adjustment factor	1.29	1.29	1.50	1.32	1.30
3. Trimming adjustment*					
3.1 Number of trimmed records	2	1	1	0	0
3.2 Sum of weights before trimming adjustment	121,089	70,148	87,650	69,907	65,641
3.3 Sum of weights after trimming adjustment	115,437	68,951	82,181	68,968	64,457
4. Raking adjustment*					
4.1 Number of completed child interviews	141	83	121	124	291
4.2 Sum of weights after adjustment	124,634	69,352	93,342	68,737	70,314
4.3 Coefficient of Variation (CV)	96.35	67.35	69.59	88.04	96.67
4.4 Mean adjustment factor	1.08	1.01	1.14	1.00	1.09
4.5 Mean weight	883.93	835.57	771.42	554.33	241.63

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	127,985	69,601	95,340	69,215	70,787
5.2 Coefficient of Variation (CV)	96.35	67.35	69.59	88.04	96.67
5.3 Mean adjustment factor	1.03	1.00	1.02	1.01	1.01
5.4 Mean weight	907.70	838.56	787.93	558.19	243.26

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
1. Child initial weights						
1.1 Number of sampled children	192	158	715	121	135	218
1.2 Sum of weights	91,331	36,052	33,063	32,531	43,595	48,832
1.3 Coefficient of Variation (CV)	137.56	70.96	82.65	70.62	95.40	78.64
2. Nonresponse adjustment						
2.1 Number of completed child interviews	145	122	573	98	110	159
2.2 Sum of weights before adjustment	91,331	36,052	33,063	32,531	43,595	48,832
a. Eligible respondents	62,595	27,686	26,215	26,763	36,864	36,001
b. Ineligible	899	354	384	209	700	69
c. Nonrespondents	27,838	8,012	6,465	5,559	6,032	12,762
2.3 Sum of weights after adjustment	91,331	36,052	33,063	32,531	43,595	48,832
a. Eligible respondents	90,071	35,614	32,583	32,297	42,792	48,742
b. Ineligible	1,260	438	480	235	803	89
c. Nonrespondents	0	0	0	0	0	0
2.4 Coefficient of Variation (CV)	95.51	73.74	83.98	72.87	99.67	79.27
2.5 Mean adjustment factor	1.46	1.30	1.26	1.22	1.18	1.36
3. Trimming adjustment*						
3.1 Number of trimmed records	1	1	0	1	3	0
3.2 Sum of weights before trimming adjustment	90,071	35,614	32,583	32,297	42,792	48,742
3.3 Sum of weights after trimming adjustment	88,446	36,015	32,583	32,805	40,708	51,073
4. Raking adjustment*						
4.1 Number of completed child interviews	146	115	573	100	112	163
4.2 Sum of weights after adjustment	86,404	36,811	32,477	31,112	47,578	49,077
4.3 Coefficient of Variation (CV)	83.71	93.59	94.78	79.11	78.36	78.94
4.4 Mean adjustment factor	0.98	1.02	1.00	0.95	1.17	0.96
4.5 Mean weight	591.81	320.10	56.68	311.12	424.81	301.09

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
5. Benchmark to 2005 DOF provisional estimates adjustment*						
5.1 Sum of weights after adjustment	88,664	36,966	32,612	31,390	49,211	50,284
5.2 Coefficient of Variation (CV)	83.71	93.59	94.78	79.11	78.36	78.94
5.3 Mean adjustment factor	1.03	1.00	1.00	1.01	1.03	1.02
5.4 Mean weight	607.29	321.45	56.91	313.90	439.38	308.49

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
1. Child initial weights					
1.1 Number of sampled children	94	124	161	140	165
1.2 Sum of weights	26,629	29,262	27,674	27,798	25,592
1.3 Coefficient of Variation (CV)	84.39	83.20	74.34	75.26	90.99
2. Nonresponse adjustment					
2.1 Number of completed child interviews	73	106	125	110	122
2.2 Sum of weights before adjustment	26,629	29,262	27,674	27,798	25,592
a. Eligible respondents	21,010	25,857	20,136	21,611	17,481
b. Ineligible	0	373	238	0	77
c. Nonrespondents	5,619	3,032	7,300	6,188	8,035
2.3 Sum of weights after adjustment	26,629	29,262	27,674	27,798	25,592
a. Eligible respondents	26,629	28,848	27,357	27,798	25,481
b. Ineligible	0	414	317	0	112
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	82.92	86.02	75.28	77.59	89.57
2.5 Mean adjustment factor	1.27	1.13	1.37	1.29	1.46
3. Trimming adjustment*					
3.1 Number of trimmed records	0	1	2	0	1
3.2 Sum of weights before trimming adjustment	26,629	28,848	27,357	27,798	25,481
3.3 Sum of weights after trimming adjustment	26,692	26,213	27,288	26,490	24,790
4. Raking adjustment*					
4.1 Number of completed child interviews	74	99	129	106	122
4.2 Sum of weights after adjustment	27,220	26,734	28,954	23,104	28,508
4.3 Coefficient of Variation (CV)	70.79	87.30	68.01	84.13	70.67
4.4 Mean adjustment factor	1.02	1.02	1.06	0.87	1.15
4.5 Mean weight	367.84	270.04	224.45	217.97	233.67

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	27,509	27,087	29,341	23,616	29,382
5.2 Coefficient of Variation (CV)	70.79	87.30	68.01	84.13	70.67
5.3 Mean adjustment factor	1.01	1.01	1.01	1.02	1.03
5.4 Mean weight	371.75	273.60	227.45	222.79	240.84

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
1. Child initial weights					
1.1 Number of sampled children	130	221	141	194	174
1.2 Sum of weights	16,965	26,070	22,222	76,195	20,924
1.3 Coefficient of Variation (CV)	65.48	78.96	78.84	102.04	92.31
2. Nonresponse adjustment					
2.1 Number of completed child interviews	101	178	114	152	151
2.2 Sum of weights before adjustment	16,965	26,070	22,222	76,195	20,924
a. Eligible respondents	13,748	21,110	17,750	58,454	17,594
b. Ineligible	0	144	66	0	0
c. Nonrespondents	3,217	4,816	4,406	17,742	3,331
2.3 Sum of weights after adjustment	16,965	26,070	22,222	76,195	20,924
a. Eligible respondents	16,965	25,887	22,136	76,195	20,924
b. Ineligible	0	183	86	0	0
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	65.41	72.76	80.49	103.17	88.63
2.5 Mean adjustment factor	1.23	1.23	1.25	1.30	1.19
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	0	2	1
3.2 Sum of weights before trimming adjustment	16,965	25,887	22,136	76,195	20,924
3.3 Sum of weights after trimming adjustment	18,153	25,832	21,916	75,961	20,711
4. Raking adjustment*					
4.1 Number of completed child interviews	106	177	112	158	151
4.2 Sum of weights after adjustment	19,772	26,632	26,260	74,800	18,391
4.3 Coefficient of Variation (CV)	70.24	81.89	91.56	85.05	84.79
4.4 Mean adjustment factor	1.09	1.03	1.20	0.98	0.89
4.5 Mean weight	186.53	150.46	234.47	473.42	121.80

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	19,957	27,106	27,039	74,978	18,478
5.2 Coefficient of Variation (CV)	70.24	81.89	91.56	85.05	84.79
5.3 Mean adjustment factor	1.01	1.02	1.03	1.00	1.00
5.4 Mean weight	188.27	153.14	241.42	474.54	122.37

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
1. Child initial weights					
1.1 Number of sampled children	85	96	119	145	79
1.2 Sum of weights	11,701	12,155	15,286	12,257	10,478
1.3 Coefficient of Variation (CV)	75.82	75.39	130.18	83.79	113.96
2. Nonresponse adjustment					
2.1 Number of completed child interviews	67	79	92	117	59
2.2 Sum of weights before adjustment	11,701	12,155	15,286	12,257	10,478
a. Eligible respondents	8,399	9,471	11,041	9,372	6,757
b. Ineligible	153	958	1,356	513	0
c. Nonrespondents	3,149	1,725	2,889	2,371	3,721
2.3 Sum of weights after adjustment	11,701	12,155	15,286	12,257	10,478
a. Eligible respondents	11,508	11,032	13,680	11,639	10,478
b. Ineligible	193	1,123	1,606	618	0
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	45.00	73.77	97.01	76.63	87.48
2.5 Mean adjustment factor	1.39	1.28	1.38	1.31	1.55
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	0	0	1
3.2 Sum of weights before trimming adjustment	11,508	11,032	13,680	11,639	10,478
3.3 Sum of weights after trimming adjustment	10,571	10,186	13,404	11,706	10,186
4. Raking adjustment*					
4.1 Number of completed child interviews	62	78	90	115	59
4.2 Sum of weights after adjustment	10,197	11,688	15,114	12,295	9,890
4.3 Coefficient of Variation (CV)	53.14	53.91	103.18	92.31	79.75
4.4 Mean adjustment factor	0.96	1.15	1.13	1.05	0.97
4.5 Mean weight	164.47	149.85	167.93	106.91	167.63

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	10,324	11,728	15,564	12,768	9,990
5.2 Coefficient of Variation (CV)	53.14	53.91	103.18	92.31	79.75
5.3 Mean adjustment factor	1.01	1.00	1.03	1.04	1.01
5.4 Mean weight	166.51	150.36	172.93	111.02	169.32

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
1. Child initial weights				
1.1 Number of sampled children	159	120	84	71
1.2 Sum of weights	11,061	18,082	22,103	24,534
1.3 Coefficient of Variation (CV)	87.84	63.93	67.14	90.13
2. Nonresponse adjustment				
2.1 Number of completed child interviews	113	100	75	59
2.2 Sum of weights before adjustment	11,061	18,082	22,103	24,534
a. Eligible respondents	7,463	14,122	20,076	19,902
b. Ineligible	0	0	0	278
c. Nonrespondents	3,598	3,960	2,027	4,353
2.3 Sum of weights after adjustment	11,061	18,082	22,103	24,534
a. Eligible respondents	11,061	18,082	22,103	24,196
b. Ineligible	0	0	0	338
c. Nonrespondents	0	0	0	0
2.4 Coefficient of Variation (CV)	77.36	66.33	66.83	94.52
2.5 Mean adjustment factor	1.48	1.28	1.10	1.23
3. Trimming adjustment*				
3.1 Number of trimmed records	0	0	1	1
3.2 Sum of weights before trimming adjustment	11,061	18,082	22,103	24,196
3.3 Sum of weights after trimming adjustment	11,230	19,944	22,951	23,365
4. Raking adjustment*				
4.1 Number of completed child interviews	114	105	78	59
4.2 Sum of weights after adjustment	11,317	17,392	17,563	22,087
4.3 Coefficient of Variation (CV)	85.46	67.38	80.50	81.60
4.4 Mean adjustment factor	1.01	0.87	0.77	0.95
4.5 Mean weight	99.27	165.64	225.16	374.35

Table B-3. Extended interview weighting for child interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
5. Benchmark to 2005 DOF provisional estimates adjustment*				
5.1 Sum of weights after adjustment	11,406	17,743	17,735	22,448
5.2 Coefficient of Variation (CV)	85.46	67.38	80.50	81.61
5.3 Mean adjustment factor	1.01	1.02	1.01	1.02
5.4 Mean weight	100.06	168.98	227.38	380.48

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
1. Adolescent initial weights					
1.1 Number of sampled adolescents	8,081	1,737	904	475	240
1.2 Sum of weights	3,277,330	921,593	257,038	262,679	120,569
1.3 Coefficient of Variation (CV)	115.34	100.47	108.88	115.45	72.58
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	4,029	793	411	222	125
2.2 Sum of weights before adjustment	3,277,330	921,593	257,038	262,679	120,569
a. Eligible respondents	1,575,346	401,037	117,615	125,879	64,539
b. Ineligible	31,283	6,049	5,499	83	238
c. Nonrespondents	1,670,701	514,508	133,925	136,717	55,792
2.3 Sum of weights after adjustment	3,277,330	921,593	257,038	262,679	120,569
a. Eligible respondents	3,213,593	908,007	245,788	262,515	120,115
b. Ineligible	63,737	13,587	11,250	164	453
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	114.87	98.04	98.67	97.32	71.48
2.5 Mean adjustment factor	2.08	2.30	2.19	2.09	1.87
3. Trimming adjustment*					
3.1 Number of trimmed records	55	9	4	1	0
3.2 Sum of weights before trimming adjustment	3,129,899	884,488	240,010	253,111	120,689
3.3 Sum of weights after trimming adjustment	3,310,822	922,254	264,790	267,347	134,332
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	4,029	792	412	217	125
4.2 Sum of weights after adjustment	3,310,822	922,254	264,790	267,347	134,332
4.3 Coefficient of Variation (CV)	111.35	96.29	102.10	101.23	90.27
4.4 Mean adjustment factor	1.058	1.043	1.103	1.056	1.113
4.5 Mean weight	821.75	1,164.46	642.69	1,232.01	1,074.66

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	All Strata	Los Angeles	San Diego	Orange	Santa Clara
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	3,358,659	930,665	267,352	269,557	135,654
5.2 Coefficient of Variation (CV)	111.40	96.29	102.10	101.23	90.27
5.3 Mean adjustment factor	1.01	1.01	1.01	1.01	1.01
5.4 Mean weight	833.62	1175.08	648.91	1242.20	1085.23

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
1. Adolescent initial weights					
1.1 Number of sampled adolescents	309	288	221	197	145
1.2 Sum of weights	213,093	173,447	121,734	122,907	90,774
1.3 Coefficient of Variation (CV)	85.44	105.91	70.16	103.95	81.64
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	150	137	106	108	76
2.2 Sum of weights before adjustment	213,093	173,447	121,734	122,907	90,774
a. Eligible respondents	106,040	85,659	53,950	68,752	48,613
b. Ineligible	1,088	0	2,685	0	0
c. Nonrespondents	105,965	87,788	65,099	54,155	42,161
2.3 Sum of weights after adjustment	213,093	173,447	121,734	122,907	90,774
a. Eligible respondents	210,343	173,447	115,696	122,907	90,774
b. Ineligible	2,750	0	6,038	0	0
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	86.64	116.87	63.67	80.90	66.24
2.5 Mean adjustment factor	2.01	2.02	2.26	1.79	1.87
3. Trimming adjustment*					
3.1 Number of trimmed records	2	2	0	0	2
3.2 Sum of weights before trimming adjustment	205,095	159,897	117,037	123,396	88,710
3.3 Sum of weights after trimming adjustment	215,390	181,151	119,002	129,189	91,388
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	151	138	105	110	78
4.2 Sum of weights after adjustment	215,390	181,151	119,002	129,189	91,388
4.3 Coefficient of Variation (CV)	80.84	78.80	75.91	89.82	59.97
4.4 Mean adjustment factor	1.050	1.133	1.017	1.047	1.030
4.5 Mean weight	1,426.42	1,312.69	1,133.36	1,174.44	1,171.64

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Bernardino	Riverside	Alameda	Sacramento	Contra Costa
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	221,121	189,141	119,522	131,264	92,526
5.2 Coefficient of Variation (CV)	80.84	78.80	75.91	89.82	59.97
5.3 Mean adjustment factor	1.03	1.04	1.00	1.02	1.01
5.4 Mean weight	1,464.38	1,370.59	1,138.31	1,193.31	1,186.23

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
1. Adolescent initial weights					
1.1 Number of sampled adolescents	139	75	130	101	132
1.2 Sum of weights	107,355	46,919	79,150	46,349	65,923
1.3 Coefficient of Variation (CV)	75.63	67.24	86.16	93.22	77.83
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	72	33	64	57	67
2.2 Sum of weights before adjustment	107,355	46,919	79,150	46,349	65,923
a. Eligible respondents	53,668	21,664	37,067	24,374	37,826
b. Ineligible	3,731	0	1,149	0	581
c. Nonrespondents	49,956	25,255	40,935	21,974	27,516
2.3 Sum of weights after adjustment	107,355	46,919	79,150	46,349	65,923
a. Eligible respondents	100,376	46,919	77,024	46,349	64,926
b. Ineligible	6,978	0	2,126	0	997
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	63.93	68.39	96.37	94.08	69.36
2.5 Mean adjustment factor	2.00	2.17	2.14	1.90	1.74
3. Trimming adjustment*					
3.1 Number of trimmed records	0	2	2	2	2
3.2 Sum of weights before trimming adjustment	100,729	42,352	75,808	44,570	61,880
3.3 Sum of weights after trimming adjustment	93,628	33,973	77,300	52,397	73,762
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	73	32	66	59	66
4.2 Sum of weights after adjustment	93,628	33,973	77,300	52,397	73,762
4.3 Coefficient of Variation (CV)	75.66	67.14	69.01	90.04	63.25
4.4 Mean adjustment factor	0.930	0.802	1.020	1.176	1.192
4.5 Mean weight	1,282.58	1,061.65	1,171.21	888.08	1,117.60

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Fresno	San Francisco	Ventura	San Mateo	Kern
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	95,376	34,104	77,979	52,662	76,141
5.2 Coefficient of Variation (CV)	75.66	67.14	69.01	90.04	63.25
5.3 Mean adjustment factor	1.02	1.00	1.01	1.01	1.03
5.4 Mean weight	1,306.52	1,065.74	1,181.50	8,92.58	1,153.65

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
1. Adolescent initial weights					
1.1 Number of sampled adolescents	107	66	94	83	238
1.2 Sum of weights	64,987	38,645	54,648	35,130	48,054
1.3 Coefficient of Variation (CV)	74.01	51.32	95.01	79.14	97.46
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	48	34	47	47	114
2.2 Sum of weights before adjustment	64,987	38,645	54,648	35,130	48,054
a. Eligible respondents	30,593	18,640	29,356	20,943	21,374
b. Ineligible	2,460	490	243	0	519
c. Nonrespondents	31,934	19,516	25,049	14,187	26,161
2.3 Sum of weights after adjustment	64,987	38,645	54,648	35,130	48,054
a. Eligible respondents	60,150	37,656	54,199	35,130	46,902
b. Ineligible	4,837	990	449	0	1,152
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	67.23	44.39	113.63	69.49	80.96
2.5 Mean adjustment factor	2.12	2.07	1.86	1.68	2.25
3. Trimming adjustment*					
3.1 Number of trimmed records	0	0	1	1	0
3.2 Sum of weights before trimming adjustment	58,630	37,656	44,281	34,249	46,158
3.3 Sum of weights after trimming adjustment	69,461	39,944	53,427	35,028	39,660
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	46	34	41	46	114
4.2 Sum of weights after adjustment	69,461	39,944	53,427	35,028	39,660
4.3 Coefficient of Variation (CV)	64.36	46.44	88.48	68.46	85.03
4.4 Mean adjustment factor	1.185	1.061	1.207	1.023	0.859
4.5 Mean weight	1,510.03	1,174.81	1,303.10	761.48	347.90

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Joaquin	Sonoma	Stanislaus	Santa Barbara	Solano
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	71,329	40,087	54,571	35,272	39,928
5.2 Coefficient of Variation (CV)	64.36	46.44	88.48	68.46	85.03
5.3 Mean adjustment factor	1.03	1.00	1.02	1.01	1.01
5.4 Mean weight	1,550.63	1,179.02	1,331.00	766.78	350.24

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
1. Adolescent initial weights						
1.1 Number of sampled adolescents	101	95	447	74	78	111
1.2 Sum of weights	40,762	24,638	19,608	19,041	24,473	24,834
1.3 Coefficient of Variation (CV)	96.68	75.27	85.92	72.77	58.91	75.72
2. Nonresponse adjustment						
2.1 Number of completed adolescent interviews	50	53	246	44	43	50
2.2 Sum of weights before adjustment	40,762	24,638	19,608	19,041	24,473	24,834
a. Eligible respondents	18,692	13,587	10,692	10,476	12,405	11,212
b. Ineligible	757	589	98	0	0	0
c. Nonrespondents	21,313	10,463	8,818	8,565	12,068	13,623
2.3 Sum of weights after adjustment	40,762	24,638	19,608	19,041	24,473	24,834
a. Eligible respondents	39,176	23,615	19,436	19,041	24,473	24,834
b. Ineligible	1,586	1,023	171	0	0	0
c. Nonrespondents	0	0	0	0	0	0
2.4 Coefficient of Variation (CV)	70.33	60.91	89.29	51.04	59.96	71.24
2.5 Mean adjustment factor	2.18	1.81	1.83	1.82	1.97	2.22
3. Trimming adjustment*						
3.1 Number of trimmed records	2	1	0	2	0	1
3.2 Sum of weights before trimming adjustment	38,866	22,952	19,431	18,355	23,874	31,419
3.3 Sum of weights after trimming adjustment	44,634	21,380	17,155	20,598	28,877	28,279
4. Raking adjustment*						
4.1 Number of completed adolescent interviews	51	52	246	45	44	56
4.2 Sum of weights after adjustment	44,634	21,380	17,155	20,598	28,877	28,279
4.3 Coefficient of Variation (CV)	69.09	55.63	98.91	45.80	76.01	79.17
4.4 Mean adjustment factor	1.148	0.931	0.883	1.122	1.210	0.900
4.5 Mean weight	875.18	411.16	69.74	457.73	656.30	504.98

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Tulare	Santa Cruz	Marin	San Luis Obispo	Placer	Merced
5. Benchmark to 2005 DOF provisional estimates adjustment*						
5.1 Sum of weights after adjustment	45,802	21,470	17,226	20,782	29,868	28,974
5.2 Coefficient of Variation (CV)	69.09	55.63	98.91	45.80	76.01	79.17
5.3 Mean adjustment factor	1.03	1.00	1.00	1.01	1.03	1.02
5.4 Mean weight	898.08	412.89	70.03	461.82	678.81	517.40

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
1. Adolescent initial weights					
1.1 Number of sampled adolescents	64	80	95	64	111
1.2 Sum of weights	18,491	17,132	16,270	12,574	18,560
1.3 Coefficient of Variation (CV)	74.47	66.09	60.42	69.32	90.48
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	37	45	59	39	59
2.2 Sum of weights before adjustment	18,491	17,132	16,270	12,574	18,560
a. Eligible respondents	10,378	8,694	10,012	7,372	9,005
b. Ineligible	0	0	0	169	371
c. Nonrespondents	8,113	8,438	6,258	5,033	9,184
2.3 Sum of weights after adjustment	18,491	17,132	16,270	12,574	18,560
a. Eligible respondents	18,491	17,132	16,270	12,293	17,826
b. Ineligible	0	0	0	281	734
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	67.82	68.16	58.23	52.49	77.13
2.5 Mean adjustment factor	1.78	1.97	1.63	1.71	2.06
3. Trimming adjustment*					
3.1 Number of trimmed records	1	0	2	0	0
3.2 Sum of weights before trimming adjustment	18,152	13,930	14,925	12,293	17,826
3.3 Sum of weights after trimming adjustment	18,296	11,992	16,774	18,363	17,642
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	39	40	57	39	59
4.2 Sum of weights after adjustment	18,296	11,992	16,774	18,363	17,642
4.3 Coefficient of Variation (CV)	67.21	44.20	58.60	60.85	83.61
4.4 Mean adjustment factor	1.008	0.861	1.124	1.494	0.990
4.5 Mean weight	469.13	299.80	294.28	470.84	299.01

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Butte	Shasta	Yolo	El Dorado	Imperial
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	18,491	12,150	16,998	18,769	18,183
5.2 Coefficient of Variation (CV)	67.21	44.20	58.60	60.85	83.61
5.3 Mean adjustment factor	1.01	1.01	1.01	1.02	1.03
5.4 Mean weight	474.12	303.76	298.21	481.27	308.18

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
1. Adolescent initial weights					
1.1 Number of sampled adolescents	70	102	78	116	119
1.2 Sum of weights	10,340	12,177	11,778	38,497	13,603
1.3 Coefficient of Variation (CV)	83.72	66.09	76.90	85.69	199.01
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	32	50	46	59	69
2.2 Sum of weights before adjustment	10,340	12,177	11,778	38,497	13,603
a. Eligible respondents	4,114	5,590	6,488	17,734	5,995
b. Ineligible	502	222	751	361	51
c. Nonrespondents	5,724	6,365	4,539	20,401	7,557
2.3 Sum of weights after adjustment	10,340	12,177	11,778	38,497	13,603
a. Eligible respondents	9,215	11,711	10,555	37,728	13,467
b. Ineligible	1,125	466	1,223	769	136
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	88.16	56.45	63.75	82.85	83.58
2.5 Mean adjustment factor	2.51	2.18	1.82	2.17	2.27
3. Trimming adjustment*					
3.1 Number of trimmed records	2	1	2	2	1
3.2 Sum of weights before trimming adjustment	9,252	12,347	10,089	36,607	12,923
3.3 Sum of weights after trimming adjustment	11,161	13,677	12,671	38,205	10,499
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	33	51	45	62	69
4.2 Sum of weights after adjustment	11,161	13,677	12,671	38,205	10,499
4.3 Coefficient of Variation (CV)	56.54	64.12	63.37	66.18	78.19
4.4 Mean adjustment factor	1.206	1.108	1.256	1.044	0.812
4.5 Mean weight	338.20	268.17	281.59	616.20	152.16

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Napa	Kings	Madera	Monterey	Humboldt
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	11,265	13,920	13,047	38,295	10,549
5.2 Coefficient of Variation (CV)	56.54	64.12	63.37	66.18	78.19
5.3 Mean adjustment factor	1.01	1.02	1.03	1.00	1.00
5.4 Mean weight	341.35	272.95	289.93	617.67	152.88

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
1. Adolescent initial weights					
1.1 Number of sampled adolescents	62	55	70	77	55
1.2 Sum of weights	8,737	7,770	9,827	5,558	5,794
1.3 Coefficient of Variation (CV)	69.49	59.76	97.94	58.32	108.41
2. Nonresponse adjustment					
2.1 Number of completed adolescent interviews	34	31	43	45	32
2.2 Sum of weights before adjustment	8,737	7,770	9,827	5,558	5,794
a. Eligible respondents	4,141	4,310	5,825	3,138	3,048
b. Ineligible	265	518	430	121	0
c. Nonrespondents	4,331	2,943	3,572	2,300	2,746
2.3 Sum of weights after adjustment	8,737	7,770	9,827	5,558	5,794
a. Eligible respondents	8,211	6,937	9,151	5,352	5,794
b. Ineligible	526	833	676	206	0
c. Nonrespondents	0	0	0	0	0
2.4 Coefficient of Variation (CV)	51.15	60.33	99.96	57.96	52.41
2.5 Mean adjustment factor	2.11	1.80	1.69	1.77	1.90
3. Trimming adjustment*					
3.1 Number of trimmed records	0	1	2	0	0
3.2 Sum of weights before trimming adjustment	8,211	6,426	7,787	4,913	5,688
3.3 Sum of weights after trimming adjustment	10,110	9,306	8,870	7,881	5,054
4. Raking adjustment*					
4.1 Number of completed adolescent interviews	34	30	43	41	31
4.2 Sum of weights after adjustment	10,110	9,306	8,870	7,881	5,054
4.3 Coefficient of Variation (CV)	57.98	69.86	74.74	80.08	49.51
4.4 Mean adjustment factor	1.231	1.448	1.139	1.604	0.888
4.5 Mean weight	297.34	310.18	206.28	192.23	163.03

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	Nevada	Mendocino	Sutter	Yuba	Lake
5. Benchmark to 2005 DOF provisional estimates adjustment*					
5.1 Sum of weights after adjustment	10,235	9,337	9,134	8,185	5,105
5.2 Coefficient of Variation (CV)	57.98	69.86	74.74	80.09	49.51
5.3 Mean adjustment factor	1.01	1.00	1.03	1.04	1.01
5.4 Mean weight	301.03	311.23	212.42	199.62	164.67

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
1. Adolescent initial weights				
1.1 Number of sampled adolescents	95	68	58	55
1.2 Sum of weights	6,557	9,858	14,390	15,066
1.3 Coefficient of Variation (CV)	71.75	89.29	58.41	63.53
2. Nonresponse adjustment				
2.1 Number of completed adolescent interviews	54	38	34	26
2.2 Sum of weights before adjustment	6,557	9,858	14,390	15,066
a. Eligible respondents	3,822	5,288	8,550	7,190
b. Ineligible	0	102	560	602
c. Nonrespondents	2,735	4,468	5,280	7,274
2.3 Sum of weights after adjustment	6,557	9,858	15,876	13,580
a. Eligible respondents	6,557	9,672	14,900	12,531
b. Ineligible	0	186	976	1,050
c. Nonrespondents	0	0	0	0
2.4 Coefficient of Variation (CV)	58.70	73.03	59.91	47.43
2.5 Mean adjustment factor	1.72	1.86	1.86	1.89
3. Trimming adjustment*				
3.1 Number of trimmed records	0	2	1	2
3.2 Sum of weights before trimming adjustment	6,557	11,804	14,506	12,018
3.3 Sum of weights after trimming adjustment	6,263	10,368	14,132	15,216
4. Raking adjustment*				
4.1 Number of completed adolescent interviews	54	43	34	26
4.2 Sum of weights after adjustment	6,263	10,368	14,132	15,216
4.3 Coefficient of Variation (CV)	63.49	66.50	62.82	58.24
4.4 Mean adjustment factor	0.955	0.878	0.974	1.266
4.5 Mean weight	115.97	241.12	415.63	585.21

Table B-4. Extended interview weighting for adolescent interview by stratum (combined RDD and surname list samples) (continued)

	San Benito	Colusa, Glenn, Tehama	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne
5. Benchmark to 2005 DOF provisional estimates adjustment*				
5.1 Sum of weights after adjustment	6,312	10,577	14,271	15,465
5.2 Coefficient of Variation (CV)	63.49	66.50	62.82	58.24
5.3 Mean adjustment factor	1.01	1.02	1.01	1.02
5.4 Mean weight	116.89	245.98	419.72	594.80

* Counts of completed interviews and sums of weights in sections 3 and 4 are based on self-reported strata rather than sampling strata.

Appendix C

Table C-1. Average *DEFF* and *DEFT* for estimates from the adult interview for continuous and categorical variables

	Categorical variables					Continuous variables				
	Design effect *(DEFF)*				*DEFT*	Design effect (*DEFF*)				*DEFT*
County/strata	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
State Total	1.83	1.90	2.71	0.12	1.33	2.47	2.20	3.92	1.37	1.55
Los Angeles	1.53	1.51	2.10	0.94	1.23	1.49	1.39	2.36	0.92	1.21
San Diego	1.65	1.64	2.35	0.94	1.27	1.66	1.53	2.19	1.26	1.28
Orange	1.76	1.73	3.75	0.38	1.31	1.70	1.30	3.06	1.25	1.28
Santa Clara	1.60	1.42	3.60	0.97	1.25	1.14	1.10	1.63	0.80	1.06
San Bernardino	1.45	1.48	1.96	0.89	1.20	1.45	1.54	1.77	0.92	1.20
Riverside	1.39	1.35	2.12	0.62	1.17	1.34	1.14	2.43	0.76	1.13
Alameda	1.50	1.49	2.32	0.73	1.21	1.77	1.60	3.37	0.84	1.30
Sacramento	1.38	1.33	2.19	0.75	1.17	1.64	1.24	3.84	0.78	1.23
Contra Costa	1.51	1.42	3.01	0.33	1.21	1.32	1.28	1.93	0.79	1.14
Fresno	1.46	1.50	2.29	0.61	1.20	1.51	1.60	2.29	0.73	1.20
San Francisco	1.41	1.40	2.16	0.88	1.18	1.32	1.34	1.73	0.76	1.14
Ventura	1.44	1.47	2.47	0.96	1.19	1.24	1.17	1.75	0.85	1.11
San Mateo	1.60	1.60	2.56	0.52	1.25	1.68	1.35	2.99	1.18	1.28
Kern	1.48	1.48	2.43	0.67	1.20	1.51	1.42	1.97	1.30	1.23
San Joaquin	1.55	1.53	3.48	0.53	1.22	1.40	1.33	1.90	0.96	1.18
Sonoma	1.40	1.44	2.48	0.75	1.17	1.42	1.53	1.84	0.88	1.18
Stanislaus	1.24	1.27	1.86	0.50	1.10	1.31	1.15	1.96	0.79	1.13
Santa Barbara	1.54	1.52	3.10	0.80	1.23	1.65	1.49	2.61	0.82	1.26
Solano	1.41	1.38	2.08	0.75	1.18	1.28	1.35	1.67	0.79	1.12
Tulare	1.45	1.50	2.31	0.76	1.19	1.38	1.34	1.97	0.91	1.16
Santa Cruz	1.25	1.20	1.87	0.66	1.11	1.40	0.95	3.39	0.88	1.14
Marin	1.88	1.89	2.96	0.87	1.36	1.49	1.48	2.48	0.79	1.20
San Luis Obispo	1.42	1.38	2.27	0.69	1.18	1.45	1.21	2.68	0.74	1.18
Placer	1.22	1.19	2.26	0.76	1.10	1.32	1.28	2.22	0.47	1.12

Table C-1. Average *DEFF* and *DEFT* for estimates from the adult interview for continuous and categorical variables (continued)

	Categorical variables					Continuous variables				
	Design effect *(DEFF)*				*DEFT*	Design effect (*DEFF*)				*DEFT*
County/strata	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
Merced	1.61	1.66	3.35	0.75	1.25	1.29	1.18	1.90	0.89	1.13
Butte	1.41	1.43	2.08	0.74	1.18	1.33	1.39	1.81	0.87	1.15
Shasta	1.30	1.34	1.97	0.61	1.13	1.15	1.27	1.41	0.74	1.06
Yolo	1.39	1.43	2.17	0.83	1.17	1.42	1.46	1.96	0.84	1.18
El Dorado	1.25	1.17	3.24	0.74	1.10	1.10	1.08	1.32	0.90	1.05
Imperial	1.37	1.30	2.39	0.88	1.16	1.54	1.44	2.65	0.68	1.22
Napa	1.42	1.45	2.19	0.41	1.18	1.33	1.17	2.26	0.77	1.13
Kings	1.33	1.33	2.16	0.81	1.14	1.02	1.06	1.50	0.60	1.00
Madera	1.44	1.51	2.09	0.30	1.18	1.21	1.07	2.25	0.69	1.08
Monterey	1.25	1.17	2.14	0.68	1.10	1.19	1.19	1.66	0.83	1.09
Humboldt	1.49	1.52	2.35	0.79	1.21	1.29	1.44	1.67	0.56	1.12
Nevada	1.29	1.30	2.18	0.50	1.12	1.21	1.14	1.66	0.87	1.09
Mendocino	1.20	1.19	1.72	0.63	1.09	1.11	1.12	1.48	0.73	1.05
Sutter	1.72	1.78	3.07	0.71	1.30	1.45	1.52	1.80	0.95	1.20
Yuba	1.41	1.36	2.37	0.82	1.18	1.45	1.47	2.12	0.90	1.19
Lake	1.45	1.38	2.51	0.71	1.19	1.21	1.17	1.55	1.01	1.10
San Benito	1.61	1.64	2.43	0.64	1.25	1.89	1.58	3.83	1.02	1.34
Colusa, Glenn, Tehama	1.25	1.25	1.92	0.74	1.11	1.07	1.08	1.61	0.68	1.02
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.34	1.35	1.93	0.92	1.15	1.29	1.16	1.99	0.99	1.12
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.23	1.27	1.95	0.68	1.10	1.38	1.32	1.78	1.23	1.17

Table C–2. Average *DEFF* and *DEFT* for estimates from the child interview for continuous and categorical variables

County/strata	Categorical variables					Continuous variables				
	Design effect (*DEFF*)				*DEFT*	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
State Total	2.71	2.75	4.66	1.98	1.64	2.58	2.19	4.18	1.83	1.59
Los Angeles	2.26	2.07	3.89	1.32	1.49	2.37	2.34	2.63	2.17	1.54
San Diego	2.03	2.01	2.92	1.35	1.42	1.39	1.45	1.85	0.72	1.16
Orange	2.35	2.21	3.55	1.55	1.52	2.18	1.54	4.39	1.38	1.44
Santa Clara	2.14	1.95	4.88	0.84	1.43	2.34	2.24	3.85	1.27	1.51
San Bernardino	1.77	1.74	3.59	0.79	1.31	1.49	1.42	1.66	1.36	1.22
Riverside	1.64	1.66	2.40	0.66	1.26	1.66	1.68	1.77	1.54	1.29
Alameda	1.53	1.52	2.87	0.80	1.22	1.40	1.39	1.62	1.22	1.18
Sacramento	1.56	1.54	2.64	0.97	1.24	1.25	1.22	1.51	1.01	1.11
Contra Costa	1.44	1.43	2.19	0.56	1.19	1.40	1.20	1.91	1.03	1.17
Fresno	1.34	1.32	2.05	0.59	1.15	1.15	1.20	1.44	0.83	1.07
San Francisco	1.38	1.33	2.25	0.56	1.16	1.50	1.48	1.78	1.31	1.22
Ventura	1.76	1.74	3.53	0.67	1.31	1.39	1.38	1.85	0.86	1.17
San Mateo	1.93	1.77	5.42	0.50	1.34	1.69	1.67	1.98	1.15	1.29
Kern	1.71	1.78	2.44	0.52	1.29	1.30	1.28	1.53	1.17	1.14
San Joaquin	1.58	1.37	4.78	0.51	1.21	1.77	1.66	2.79	1.03	1.31
Sonoma	1.22	1.25	1.79	0.26	1.09	1.13	1.30	1.35	0.61	1.05
Stanislaus	1.54	1.46	2.44	0.82	1.23	1.82	1.88	2.24	1.12	1.34
Santa Barbara	1.41	1.46	2.80	0.44	1.17	1.35	1.33	1.56	1.16	1.16
Solano	1.47	1.37	3.00	0.51	1.19	2.04	1.46	2.99	1.34	1.40
Tulare	1.62	1.63	2.62	0.89	1.26	1.50	1.46	1.79	1.21	1.22
Santa Cruz	1.20	1.21	1.70	0.40	1.08	1.27	1.28	1.53	0.88	1.12
Marin	1.87	1.63	4.32	0.94	1.35	2.43	2.05	3.71	1.39	1.54
San Luis Obispo	1.52	1.47	3.38	0.41	1.21	1.26	1.30	1.69	0.91	1.12
Placer	1.36	1.14	4.21	0.21	1.12	1.52	1.43	2.06	1.33	1.23

Table C–2. Average *DEFF* and *DEFT* for estimates from the child interview for continuous and categorical variables (continued)

County/strata	Categorical variables					Continuous variables				
	Design effect (*DEFF*)				*DEFT*	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
Merced	1.47	1.46	2.36	0.68	1.20	1.49	1.34	1.92	1.22	1.21
Butte	1.46	1.41	2.28	0.65	1.20	1.28	1.33	1.63	0.96	1.13
Shasta	1.54	1.57	2.82	0.58	1.22	1.15	1.24	1.90	0.30	1.04
Yolo	1.36	1.35	2.13	0.45	1.15	1.27	1.25	1.58	1.06	1.12
El Dorado	1.67	1.45	4.25	0.81	1.25	1.61	1.62	2.07	1.19	1.26
Imperial	1.55	1.49	3.19	0.78	1.23	1.51	1.33	2.22	0.80	1.21
Napa	1.22	1.27	1.78	0.54	1.09	1.22	1.20	1.42	1.01	1.10
Kings	1.55	1.53	3.00	0.65	1.23	1.74	1.77	2.32	1.18	1.31
Madera	1.48	1.48	2.41	0.76	1.20	1.48	1.46	1.66	1.29	1.22
Monterey	1.48	1.46	2.89	0.49	1.20	1.70	1.59	2.12	1.17	1.29
Humboldt	1.51	1.47	2.97	0.55	1.21	1.39	1.14	1.90	1.06	1.17
Nevada	1.43	1.42	3.10	0.55	1.17	1.02	1.01	1.37	0.81	1.00
Mendocino	1.22	1.15	1.92	0.77	1.10	1.40	1.46	1.52	1.10	1.18
Sutter	1.50	1.41	2.62	0.50	1.20	2.22	2.13	2.68	2.02	1.49
Yuba	1.21	1.17	2.38	0.15	1.07	1.44	1.40	1.99	1.11	1.19
Lake	1.46	1.49	3.15	0.40	1.18	1.36	1.44	1.54	1.19	1.17
San Benito	1.23	1.35	1.95	0.17	1.08	1.19	1.22	1.45	0.98	1.09
Colusa, Glenn, Tehama	1.15	1.16	1.91	0.77	1.07	1.37	1.32	1.68	1.08	1.17
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.20	1.25	1.85	0.46	1.08	1.38	1.27	1.87	1.09	1.17
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.55	1.44	3.37	0.00	1.18	1.67	1.95	1.96	1.18	1.28

Table C–3. Average *DEFF* and *DEFT* for estimates from the adolescent interview for continuous and categorical variables

County/strata	Categorical variables					Continuous variables				
	Design effect (*DEFF*)				*DEFT*	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
State Total	1.97	1.99	4.27	1.40	1.39	1.93	1.73	2.47	1.62	1.38
Los Angeles	1.86	1.83	3.51	0.89	1.35	1.67	1.72	2.10	0.89	1.28
San Diego	1.69	1.67	2.32	0.78	1.29	1.62	1.58	1.88	1.43	1.27
Orange	1.87	1.75	3.67	0.44	1.34	2.02	1.76	2.77	1.41	1.41
Santa Clara	1.53	1.46	3.94	0.42	1.21	1.21	1.30	1.47	0.78	1.09
San Bernardino	1.48	1.45	2.01	1.05	1.21	1.15	1.08	1.52	0.87	1.07
Riverside	1.38	1.27	2.85	0.74	1.16	1.44	1.32	2.79	0.78	1.18
Alameda	1.35	1.33	1.95	0.71	1.15	1.45	1.54	2.05	0.87	1.19
Sacramento	1.62	1.51	3.82	0.27	1.25	1.96	1.65	3.11	1.29	1.38
Contra Costa	1.15	1.23	1.70	0.37	1.06	1.17	1.11	1.68	0.94	1.08
Fresno	1.57	1.52	2.56	0.99	1.24	1.40	1.45	1.88	0.72	1.17
San Francisco	1.10	1.15	1.42	0.45	1.04	1.71	1.30	4.83	0.86	1.25
Ventura	1.50	1.51	2.56	0.60	1.21	1.72	2.06	2.31	0.93	1.29
San Mateo	1.37	1.33	2.96	0.44	1.14	1.99	1.79	4.09	0.96	1.38
Kern	1.15	1.15	2.12	0.14	1.05	1.36	1.39	2.07	0.86	1.15
San Joaquin	1.28	1.28	2.59	0.60	1.12	1.34	1.25	2.14	0.93	1.15
Sonoma	1.25	1.23	2.24	0.80	1.11	1.24	1.15	1.88	0.96	1.11
Stanislaus	1.22	1.19	2.07	0.51	1.09	1.22	0.99	2.52	0.45	1.07
Santa Barbara	1.33	1.39	1.74	0.00	1.12	1.69	1.63	3.20	1.05	1.28
Solano	1.70	1.62	3.64	0.92	1.29	1.55	1.42	2.41	1.07	1.24
Tulare	1.28	1.20	2.26	0.57	1.12	1.25	1.11	2.04	0.67	1.10
Santa Cruz	1.26	1.22	2.84	0.66	1.11	1.23	1.33	1.51	0.68	1.10
Marin	1.85	1.91	3.09	0.63	1.35	1.55	1.48	1.83	1.36	1.24
San Luis Obispo	1.10	1.08	1.48	0.85	1.05	1.05	1.04	1.35	0.89	1.02
Placer	1.57	1.52	2.98	0.71	1.23	1.33	1.25	2.09	0.78	1.14

Table C–3. Average *DEFF* and *DEFT* for estimates from the adolescent interview for continuous and categorical variables (continued)

County/strata	Categorical variables					Continuous variables				
	Design effect (*DEFF*)				*DEFT*	Design effect (*DEFF*)				*DEFT*
	Average	Median	Maximum	Minimum	Average	Average	Median	Maximum	Minimum	Average
Merced	1.48	1.46	2.22	0.47	1.20	1.35	1.39	1.61	0.90	1.16
Butte	1.11	1.10	1.65	0.68	1.05	1.23	1.29	1.77	0.49	1.09
Shasta	1.12	1.13	1.79	0.73	1.05	1.06	1.06	1.40	0.67	1.02
Yolo	1.08	1.16	1.95	0.20	1.01	1.13	1.06	1.57	0.88	1.06
El Dorado	1.29	1.17	2.76	0.80	1.13	1.38	1.19	2.19	1.09	1.16
Imperial	1.72	1.54	2.99	0.83	1.29	1.73	1.59	2.70	0.91	1.30
Napa	1.05	0.97	1.74	0.51	1.01	1.23	1.14	1.56	0.88	1.11
Kings	1.28	1.29	1.83	0.60	1.13	1.22	1.29	1.56	0.43	1.09
Madera	1.47	1.39	3.52	1.06	1.20	1.47	1.57	1.83	1.09	1.21
Monterey	1.29	1.28	1.77	0.65	1.13	1.27	1.31	1.87	0.45	1.11
Humboldt	1.27	1.24	2.04	0.81	1.12	1.34	1.34	1.83	0.97	1.15
Nevada	1.14	1.22	1.50	0.48	1.06	1.28	1.30	1.70	0.96	1.13
Mendocino	1.35	1.34	3.45	0.62	1.14	1.87	1.41	3.47	1.02	1.34
Sutter	1.39	1.43	2.30	0.47	1.17	1.11	1.33	1.66	0.51	1.03
Yuba	1.49	1.46	3.25	0.48	1.20	1.47	1.57	1.79	0.95	1.21
Lake	1.31	1.28	2.13	0.69	1.14	1.32	1.27	1.66	0.82	1.14
San Benito	1.19	1.24	2.01	0.00	1.03	1.56	1.62	2.14	0.83	1.24
Colusa, Glenn, Tehama	1.15	1.22	1.44	0.60	1.07	1.36	1.34	2.17	0.78	1.15
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	1.19	1.20	1.92	0.68	1.08	1.00	1.05	1.21	0.67	1.00
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1.25	1.37	1.59	0.00	1.09	1.19	1.29	2.13	0.51	1.07