



UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO · SANTA BARBARA · SANTA CRUZ

The California Work and Health Survey
Institute for Health Policy Studies
Box 0920
San Francisco, CA 94143-0920

(415) 476-1001

A User's Guide to the California Work and Health Survey

2000 Cross-sectional File

1999-2000 Longitudinal File

(updated September 12, 2000)

Overview

The California Work and Health Survey (CWHS) is a telephone-based, longitudinal survey of California adults, designed by faculty and staff of the Work and Health Program at the University of California, San Francisco, with input from researchers and practitioners in the fields of health and economics. It is conducted by Field Research Corporation and funded by the California Wellness Foundation. The survey includes extensive coverage of employment status, recent job loss, working conditions and environment, and of physical and mental health status. Interviews are conducted in English and Spanish. Three annual interviews have now been completed. Preliminary findings from all three years, 1998-2000, are available at <http://medicine.ucsf.edu/programs/cwhs/>.

This codebook provides general guidelines for use of both the 2000 cross-sectional and the 1999-2000 longitudinal file. The cross-sectional dataset includes the responses from the 2,168 participants interviewed between May 1 and July 9 2000. The longitudinal dataset contains interviews with the 1,265 participants who responded to both 1999 & 2000 interviews.

Public Use Agreement

The California Work and Health Survey data are available for use by the general public, with the understanding that they will be used solely for research purposes. Users of the CWHS must agree to make no attempt to identify individual respondents on the datafile. Additionally, we ask that you provide us with a copy of any publication or report that you produce using the CWHS. Send such publications to us at:

The California Work and Health Survey
University of California, San Francisco
Box 0920
San Francisco, CA 94143-0920

Acknowledgments

The principal and co-principal investigators for the CWHS are Edward Yelin, Ph.D. and Laura Trupin, M.P.H. of the Institute for Health Policy Studies and Department of Medicine at the University of California, San Francisco. Irene Yen, Ph.D., of the Institute for Health and Aging at the University of California, San Francisco, assisted in the design of the survey. The project is funded by the Work and Health Initiative of The California Wellness Foundation. Ruth Brousseau, Ph.D. and Lucia Corral, J.D., are the Senior Program Officer and Program Officer, respectively, for the Work and Health Initiative at the Foundation. The survey was conducted by

Field Research Corporation, an independent public policy research organization in San Francisco, under the direction of Mark DiCamillo. Stephanie Rush prepared the public release file and this guide.

The following individuals provided additional expertise in the design of the survey: Steve Levy of the Center for the Continuing Study of the California Economy; David Dooley of the School of Social Ecology at the University of California, Irvine; E. Richard Brown and Thomas Rice of the Center for Health Policy Research at UCLA, Paul Ong, Director of the UCLA Lewis Center for Regional Policy Studies; David Hayes-Bautista, Director of the Center for the Study of Latino Health at UCLA; Nancy Adler, Director of the Health Psychology Program at the University of California, San Francisco; Amy Dean, President of the South Bay AFL-CIO Labor Council; James Head, President of the National Economic Development and Law Center in Oakland; Chris Benner of the Department of City and Regional Planning, University of California, Berkeley; Dan Friedland, George Kaplan, and Rick Price of the Michigan Prevention Research Center at the University of Michigan, Ann Arbor, and Harriet Presser of the University of Maryland.

Sample

Of the 2,168 respondents to the 2000 CWHS, 1,265 were been participants in the 1999 survey, representing 64% of the 1,964 1999 respondents who had agreed to be re-contacted. Among these continuing respondents, 627 were first interviewed in 1998 and 638 were first interviewed in 1999. The 903 baseline 2000 respondents were comprised of 441 obtained through random-digit dialing with the remainder obtained through oversampling procedures described below.

For new respondents, up to six attempts were made to contact an adult member of the household. If more than one adult was available, the interviewer requested the youngest adult male; if there were no men at home, the oldest female was asked to participate. This strategy is one that has been developed by the major public opinion research firms, in order to balance participation in phone surveys. The overall cooperation rate¹ for new respondents is 54.6%.

In order to increase the accuracy of survey estimates among selected groups of interest, an additional 462 interviews were conducted among African Americans, Asian/Pacific Islanders, and Latinos. The oversampling procedures differed for each of the three populations. To oversample African Americans, phone numbers were drawn at random from a list of phone exchanges covering areas known to have higher proportions of African Americans than the state average, based on Census block information. For Asian/Pacific Islanders, phone numbers were drawn at random from a list of Asian surnames and phone numbers compiled by an independent survey sampling firm. A screening question determined if the respondent was a member of the particular group of interest. Latinos were screened into the survey through a question about Hispanic ethnicity at the beginning of the interviews.

Results from the total sample are subject to a sampling error of ± 2.2 percentage points. Users are cautioned against attempting to make estimates for small subsets of the data.

Weights

In order to have the sample of respondents align as closely as possible to the known characteristics of the California population (in terms of household size, age, gender, race/ethnicity, and region) and to account for the oversampling of the groups described above,

¹ Cooperation rate = # of completed interviews / (# completed+ # refused+# ineligible)

the survey firm, Field Research Corporation, calculated weights for each respondent. Two types of weights have been included in the file: a proportional weight (WGTTOT00)—using this weight does not inflate the sample size— and an expansion weight (WGTEXT00)— using this variable inflates the sample to the entire California adult population.

The proportional weight should be used for any statistical calculations. The sum of the weighted frequencies will be equal to the sample size. Use the expansion weight to make prevalence estimates for the California adult population.

For example, to estimate the number of California workers with more than 1 job, use WGTEXT00. However, to compare the likelihood of multiple jobholding for men and women using a chi-square test, use WGTTOT00, so as not to inflate the sample size for the statistical test. Sample SAS code for these two examples is included in Appendix I.

For the longitudinal file, there are also two weight variables, WGT9900L and WGT9900X, representing the proportional and expansion weight for the longitudinal dataset, respectively. These weights are based on the 1999 sampling weights, with an adjustment to account for attrition. Appendix III describes in detail how we developed that adjustment.

We strongly recommend using the weights. If you choose not to include the weight variables in your analysis, you should select only those respondents obtained through random digit-dialing. The variable SAMPLE00 indicates whether the household was selected through random digit-dialing or from one of the oversampling processes. Note, however, that this group cannot be considered a true random sample of California adults, because the individual respondent in a given household was not selected at random. Also, it is not possible to determine from the 2000 data how the 1998 or 1999 respondents were originally obtained. This would limit the sample of randomly selected respondents to the 441 baseline respondents in 2000.

Data

Accompanying this document is a questionnaire that has been annotated with the variable names corresponding to each question. The variables in the dataset are in the same order as the questions². In addition, there is an alphabetical listing of the variables in both datasets included in Appendix II.

The dataset can be provided in several ways: as a SAS file (in version 6.12 for Windows or as a transport dataset), as an SPSS "portable" file or as a text file. In the SAS and SPSS files, the question numbers are identified in the variable labels. Appendix I gives sample SAS code for converting the transport dataset to a working file. The text file contains one row per respondent, with a space between each variable. Depending on the individual responses, the record length varies from a minimum of 892 to a maximum of 1,158 characters. A second text file lists the variable names in order. We may be able to create datasets in other formats as well; contact the CWHS office with your specific request.

Masking for confidentiality

To ensure that no individual can be identified from these data, we do not include geographic information below the county level. Further, we have combined 23 small counties (defined here as having fewer than 100,000 residents) into a single code -- "36".

² Due to the process of revising the survey skip patterns, the questions are frequently out of numeric order.

Notes on variables

Most variable names in this 2000 dataset have "00" in the last 2 positions (e.g. WORK00). The only exceptions to this are demographic variables that will not change over time, the respondent's ID, and the year a respondent first entered the survey. Prior years' variables have "98" or "99" in the last two positions. We have recoded responses from hours and minutes to minutes, and from months and years to years, in order to have all information from a single question contained in a single variable.

The social network questions (q. 112-116) ask about numbers of friends or relatives. Responses to these questions range up to 98.

Not all variables are listed on the questionnaire. These include: the weight and sampling variables described above, two longitudinal variables that will be useful when pooling multiple years of data, a handful of calculated variables for the convenience of the user, and labor force status variables based on U.S. Bureau of Labor Statistics (BLS) definitions. All of these additional variables are defined below:

Longitudinal and sampling variables

1. ID = Respondent identifier; will not change from year to year.
2. YEAR00 = survey year (2000 for all respondents).
3. YENTER = the year the respondent entered the survey (1998 or 1999 for panel respondents, 2000 for new respondents).
4. INTRVW00: Identifies interview number for each respondent. New respondents have a 1, continuing 1999 respondents have a 2, and continuing 1998 respondents have a 3.
5. PANEL00: 1=Continuing respondents, 0=new respondent
6. SAMPLE00: 1=Cross section (2000 Baseline), 2=Asian augment (excluding panel), 3=Black augment (excluding panel), 4=Disabled augment (excluding panel), 5=Aging supplement (excluding panel), 6 Panel re-interviews
7. LANG00 = Language of interview: 0=English 1=Spanish

Weight variables in the cross-sectional file

1. WGTTOT00 = proportional weight for all 2000 respondents
2. WGTEXT00 = expansion weight for all 2000 respondents

Weight variables in the longitudinal file (see Appendix III for details)

1. WGT9900L= proportional weight for 1999-2000 respondents
2. WGT9900X = expansion weight for 1999-2000 respondents
3. WGTTOT99 = original 1999 sampling weight
4. ATTRADJ = adjustment to WGTTOT99 to account for loss to follow-up.

Calculated variables

1. HASJOB00 = combination of q. 1, q.2, and q. 9 -- 1=has job, 0=does not have job
Respondents with a 1 in this variable are asked the set of questions about current job(s).
Note that this is **not** a BLS labor force status variable, although it is quite similar to the employment/population ratio (see below).
2. TOTADU00 = Total number of adults in household
3. TOTCHI00 = Total number of children in household

4. DEPRES00 = The Short Geriatric Depression Scale (S-GDS) score ^{1,2}. This is a count of all the depressive symptoms identified in questions 127a - 127o. The score is calculated by converting the responses so that 1 signifies a depressive symptom for all 15 items. A cut-point of 7 (out of 15 responses) has been suggested as an indicator of a high degree of depressive symptomology³.

Note that all respondents who answered *any* of the items in question 127 have a GDS score. Users may wish to omit respondents who skipped items, or to make the cut-point for depression be based on the total number of items answered.

References

1. Sheikh JI & Yesavage JA. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clin Gerontol* 5:165-173.
2. Rule BG, Harvey HZ & Dobbs AR. (1989). Reliability of the Geriatric Depression Scale for younger adults. *Clin Gerontol* 9:37-43.
3. Cwikel J & Ritchie K. (1989). Screening for depression among the elderly in Israel: An assessment of the short Geriatric Depression Scale (S-GDS). *Isr J Med Sci* 25:131-137.

5. BMI00 = Body mass index (kg/m²)
6. RACE = a variable identifying 1 racial category for each person.
 - 1 = White
 - 2 = African American
 - 3 = Asian/Pacific Islander
 - 4 = Other

Note: In keeping with current US Census guidelines, the CWHHS questionnaire allows for persons to identify themselves as a member of multiple racial groups. In cases of multiple responses, non-“other” (i.e. white, African American, or Asian/Pacific Islander) takes precedence over “other,” and “non-white” (i.e. African American or Asian/Pacific Islander) takes precedence over white. There were no respondents who identified themselves as both African American and Asian/Pacific Islander. The original responses have been retained in the file, as well.

7. RURAL00 = whether the person resides in a rural county (1) or not (0); taken from the California Rural Health Policy Council definition which is based on population density (<http://www.ruralhealth.cahwnet.gov/>). Rural counties include all those coded “36” in the COUNTY00 variable, and a number of geographically large, sparsely-populated counties.
8. SOCA00 = resides in Southern (1) or Northern California (0). Southern California includes the 9 most southerly counties, and Northern California includes all other counties.

Longest job variables

Two variables identify the occupation and industry of the longest job held: LONGOC00 and LONGIN00. These variables are coded in the same way as the current occupation and industry variables. The information is gathered from several questions, depending upon the respondent's current and past employment situation. Responses from 1999 participants for these variable were brought forward into the 2000 dataset. Therefore, users will not be able to replicate this variable using the raw 2000 data.

Labor force variables

Several variables help users identify labor force status in keeping with the Bureau of Labor Statistics (BLS) definitions. The CWHS employment status questions were patterned after the Current Population Survey (CPS), and the LFCPS00 variable was coded following the CPS guidelines. To be unemployed, a person must be on temporary lay-off and awaiting a callback, actively looking for work and available for work, or waiting to start a new job in the next 30 days. All other persons without jobs are considered to be out of the labor force. The SAS code used to identify the BLS labor force status is included in Appendix I.

1. LFCPS00 = Labor force status, modeled after the BLS' Current Population Survey
 - 1 = Working
 - 2 = With a job, but not at work
 - 3 = Unemployed
 - 4 = Out of the labor force
2. LFPRT00 = labor force participation rate:
 - 1 = In labor force (LFCPS00= 1,2,3), 0 = out of labor force (LFCPS00 = 4)
3. EMPRT00 = employment/population ratio:
 - 1 = Has job (LFCPS00 = 1,2), 0 = No job (LFCPS00 = 3,4)
4. UNEMPR00 = unemployment rate:
 - 1 = Unemployed (LFCPS00= 3), 0 = Has job (LFCPS00 = 1,2)Note: persons out of the labor force are excluded from this variable entirely.

Missing values

There are three types of missing values on the file. If a question was irrelevant to the respondent and appropriately skipped, the variable has a value of "." in the SAS file, and -99 in the text file. Responses of don't know, refused, or not answered are coded with ".R" in the SAS file and -98 in the text file. There are two variables with special missing values, coded ".S" in the SAS file and -97 in the text file, indicating no set number of sick days (q. 32y - SICKLV00) and never worked 6 months at time (q. 79a - FRSTJB00). Users of the text file are advised to convert the missing values before running numerical procedures.

Feedback

We welcome your comments about the survey topics, the specific questionnaire items, and the documentation. If you think that you have identified a coding error, please contact us promptly. Thank you.

Appendix I -- Sample SAS code

Working with weights

- a. Estimating prevalence of male and female workers with more than 1 job:

```
Proc freq; tables ONEJOB00*GENDER; weight WGTEXT00; run;
```

- b. Testing gender difference in proportion of population with more than 1 job:

```
Proc freq; tables ONEJOB00*GENDER /chisq; weight WGTTOT00; run;
```

(Note that the percentages in both examples will be identical).

Sample SAS code to copy the transport file to a working dataset:

Note: These statements are provided for users who have requested the transport dataset for using SAS under a platform other than Windows. Refer to the documentation for your SAS platform for the correct way to refer to file and library names, etc.

```
*****Set up incoming and outgoing libraries;
libname frxpt XPORT 'C:\lwork\cwhs00pb.xpt'; *name/location of transport data;
libname tosas 'C:\lwork\cawork\cwhsdata'; *location for destination dataset;

*****Run proc copy, selecting CWHS dataset;
proc copy in=frxpt out=tosas; select cwhs00pb; run;

***Confirm successful transport;
proc contents data=tosas.cwhs00pb; run;
```

Appendix I -- Sample SAS code, continued

Labor force variables (using Bureau of Labor Statistics definitions)

```
LIBNAME SSD 'C:\1WORK\CAWORK';
Data one; set ssd.cwhs00pb;
**2000 Labor force statistics -- per approx. CPS definition;
  *LFCPS00: 1=working, 2=have job, not at work, 3=unemployed, 4=out of LF;
LFCPS00 = 4; ** Out of LF, unless specified below;
** Working;
If WORK00 = 1 or anywrk00 = 1 then LFCPS00 = 1;
If anywrk00 = .R and WORK00 = .R then LFCPS00 = .R;

**Have job but not at work (i.e. illness, vacation, bad weather, labor disp
  or other reason - - but clearly have a job, based on subsequent answers);
If 1 le REASO00 le 4 then LFCPS00=2;
If REASO00 eq 8 and LSTWK200 eq 1 then LFCPS00=2;
If ABSENT00 eq 1 and REASO00 eq .R then LFCPS00=.R;
If ABSENT00 eq 2 and LKING00 eq .R then LFCPS00=.R;

**Unemployed or Out of LF;
If REASO00 in (5,7) then LFCPS00=3; *temp00 layoff or waiting for job
to start;
If LKING00 eq 1 and TAKJOB00 eq 1 then LFCPS00=3; *on indef. layoff & avail;
If LOOKNO00 eq 2 then LFCPS00=4; *not actively lking;

*CPS rates;
*LFPR;
If 1 le LFCPS00 le 3 then LFPRT00=1; *in labor force;
Else if LFCPS00 eq 4 then LFPRT00=0; *out of labor force;

*Emp/Pop ratio;
If 1 le LFCPS00 le 2 then EMPRT00=1; *Employed;
Else if 3 le LFCPS00 le 4 then EMPRT00=0; *Not empl. (included unemp
& out of LF);
*Unemp;
If LFCPS00 eq 3 then UNEMPR00=1; *Unemployed;
Else if 1 le LFCPS00 le 2 then UNEMPR00=0; *Employed;

*Employ00 (per CWHS);
If ANYWRK00 eq 1 or WORK00 eq 1 then EMPLOY00=1;
Else EMPLOY00=0;
If ANYWRK00=.R and WORK00=.R then EMPLOY00=.R;

*HASJOB00 (in 2000, treated as "employed" for purposes of employment
questions);
If EMPLOY00=1 or REASO00 in (1,2,3,4) then HASJOB00=1;
Else HASJOB00=0;
```


Appendix II -- Alphabetical variable lists

2000 cross-sectional file

ABSENT00	DISETH00	INDUL200	MOHRS200	RURAL00	WONDER00
ACCESS00	DISHAN00	INDUST00	NCHILD00	SAMPLE00	WORK00
ADC00	DISRAC00	INS12M00	NEW00	SELFEM00	WORK1800
AFDC00	DISSEX00	INSCHG00	NOINS00	SHIFT00	WORKD00
AGE00	DROPAC00	INSTYC00	NOISE00	SHOPS00	WORKE00
AGE1800	EDUC00	INTRVW00	NONEG00	SICKLV00	WORKED00
AGE51700	EMPLOY00	JOBDEG00	NOTHIR00	SIT00	WORKER00
AGECAT00	EMPRT00	JOBINJ00	NOTPRO00	SLEEP00	WORKMO00
AGLES500	ENERGY00	KIDNEY00	NOWORK00	SMAGE00	WORKN00
ALCDNK00	EXERCI00	LAGEN00	NWREAS00	SMDAY00	WORKS00
ALCFIV00	EXMTH00	LAIDOF00	OCC100	SMOKE00	WORKTI00
ALCMTH00	EXTIME00	LANG00	OCC200	SMPAST00	WORKYR00
ALCOHO00	EXTRBW00	LANGHOME	OCCL200	SMYRS00	WORTH00
ALCWEEK00	EXTYPT00	LANGWK00	ONE00	SOCA00	YEAR00
ANYPLN00	FAST00	LASMED00	ONEJOB00	SOCIAL00	YENTER
ANYWRK00	FEETPB00	LASTYR00	ORIFIR00	SPIRIT00	YOURWA00
ARTHRI00	FIRE00	LEAVJB00	ORIGIN	STAIR00	
ASIAN	FIREIM00	LFCPS00	ORIHIR00	STIFFJ00	
ASTHMA00	FLXHRS00	LFPR00	ORINO00	STOOP00	
ATHOME00	FODSTP00	LIFEMP00	ORIPRO00	SUFINC00	
BACK00	FREDOM00	LIFSAT00	OVERCO00	SUPER00	
BACKPN00	FRIEJW00	LIFT00	OWN00	TAKJOB00	
BAD00	FRIENW00	LIGHT00	PANEL00	TALL	
BEDDAY00	FRSTJB00	LIMACT00	PARENT00	TANF00	
BEND00	FULPRT00	LIMHO200	PARTED00	TEMPAY00	
BETTER00	GENDER	LIMHOU00	PAYEND00	TIME00	
BLACK	GRPDOW00	LIMWR200	PAYPAR00	TIRED00	
BMI00	HAND00	LIMWRK00	PDAYCA00	TMPHLT00	
BORED00	HAPPY00	LKAD00	PENPLN00	TOOLS00	
BORN	HASJOB00	LKEMP00	PENTYP00	TOTADU00	
BREATH00	HBP00	LKFRND00	PERJOB00	TOTCHI00	
CALBAC00	HEALTH00	LKING00	PERSON00	TRASH00	
CANCER00	HEAR00	LKOTHE00	PREFER00	TYPEMC00	
CARHRS00	HEART00	LONGER00	PREMI00	ULCER00	
CARPAL00	HINC300	LONGJB00	PRESOC00	UNEMPR00	
CARRY00	HISPANIC	LOOK200	PROBLE00	UNFIRE00	
CHANGE00	HLPLSS00	LOOKNO00	PROJEC00	UNICOV00	
CHEKUP00	HLTINS00	LOSEJO00	PROMO00	UNIMEM00	
CIGAV100	HMO00	LOSTJC00	RACE	USAGE	
CIGAV200	HOPLES00	LSTWK200	RACE_DK	VISION00	
COMPUT00	HOSINC00	LSTYR00	RACE_OTH	WALK00	
CONCEN00	HOSP00	LUNG00	REASO00	WEIGHT00	
CONSLW00	HOURSC00	MANYJ00	REASO100	WEKWRK00	
CONTRC00	HOUSEH00	MARIED00	REDUC00	WELFAR00	
CONTRL00	HOWLG200	MEDOFC00	REGPAY00	WGTEXT00	
COUNTY00	HOWLN200	MEDREG00	RELCLS00	WGTTOT00	
CRIME00	HRDShP00	MEDVIS00	RELPHC00	WHITE	
DEPRES00	HRSWK00	MIGRAI00	RELVIS00	WHOME00	
DIABET00	ID	MISSWK00	RENT00	WIC00	
DISAGE00	INDINC00	MOHRS100	RTTIME00	WKSLOK00	

Appendix II -- Alphabetical variable lists, continued

1999-2000 longitudinal file

ABSENT00	ATTRADJ	CONTRL99	FAST00	HMO99
ABSENT99	BACK00	COUNTY00	FAST99	HOPLES00
ACCES299	BACK99	COUNTY99	FEETPB00	HOPLES99
ACCESS00	BACKPN00	CRIM299	FINBET99	HOSINC00
ACCESS99	BAD00	CRIME00	FIRE00	HOSINC99
ADC00	BAD99	CRIME99	FIRE099	HOSP00
ADC99	BARDNK99	DEPRES00	FIREIM00	HOSP99
AFDC00	BEDDAY00	DEPRES99	FLXHRS00	HOURSC00
AFDC99	BEDDAY99	DIABET00	FLXHRS99	HOURSC99
AG51799	BEND00	DIABET99	FODSTP00	HOUSEH00
AGE00	BEND99	DISAGE00	FODSTP99	HOUSEH99
AGE99	BETTER00	DISAGE99	FREDOM00	HOWLG200
AGE1800	BETTER99	DISETH00	FREDOM99	HOWLG299
AGE1899	BLACK	DISETH99	FRIEJW00	HOWLN199
AGE51700	BMI00	DISHAN00	FRIEJW99	HOWLN200
AGECAT00	BMI99	DISHAN99	FRIENW00	HOWLN299
AGECAT99	BORED00	DISRAC00	FRIENW99	HRDShP00
AGERET99	BORED99	DISRAC99	FRSTJB00	HRDShP99
AGESOC99	BORN	DISSEX00	FULPRT00	HRSWK00
AGESUP99	BREATH00	DISSEX99	FULPRT99	HRSWK99
AGLES500	CALBAC00	DROPAC00	GAPPAY99	ID
AGLES599	CALBAK99	DROPAC99	GENDER	INDINC00
ALCDNK00	CANCER00	EDUC00	GRPDOW00	INDINC99
ALCDNK99	CANCER99	EDUC99	GRPDOW99	INDUL199
ALCFIV00	CARE99	ELICON99	HAND00	INDUL200
ALCFIV99	CARHRS00	ELIPLN99	HAND99	INDUL299
ALCMTH00	CARHRS99	EMPLAN99	HAPPY00	INDUST00
ALCMTH99	CARPAL00	EMPLOY00	HAPPY99	INDUST99
ALCOHO00	CARPAL99	EMPLOY99	HASJOB00	INS12M00
ALCOHO99	CARRY00	EMPRT00	HASJOB99	INS12M99
ALCWEK00	CARRY99	EMPRT99	HBP00	INSCHG00
ALCWEK99	CHANGE00	ENERGY00	HBP99	INSCHG99
ANYPLN00	CHANGE99	ENERGY99	HEALTH00	INSTYC00
ANYPLN99	CHEKUP00	EXERCI00	HEALTH99	INSTYC99
ANYWRK00	CHEKUP99	EXERCI99	HEAR00	INTRVW00
ANYWRK99	CIGAV100	EXMTH00	HEART00	INTRVW99
ARTHRI00	CIGAV199	EXMTH99	HEART99	JOB99
ARTHRI99	CIGAV200	EXTIME00	HINC300	JOBDEG00
ASIAN	CIGAV299	EXTIME99	HINC399	JOBINC99
ASSET99	COMPUT00	EXTRBW00	HISPANIC	JOBINJ00
ASSET299	CONCEN00	EXTRBW99	HLPLSS00	KIDNEY00
ASSIST99	CONSLT99	EXTYPT00	HLPLSS99	KIDNEY99
ASTHMA00	CONSLW00	EXTYPT99	HLTH99	LAGEN00
ASTHMA99	CONTRC00	FAM99	HLTINS00	LAGEN99
ATHOME00	CONTRC99	FAM299	HLTINS99	LAIDOF00
ATHOME99	CONTRL00	FAMHLY99	HMO00	LAIDOF99

Appendix II -- Alphabetical variable lists, continued

1999-2000 longitudinal file, continued

LANG00	LONGJB00	NEW00	OWN00	REDUC00
LANG99	LOOK200	NEW99	OWN99	REDUC99
LANGHOME	LOOK299	NOINS00	PANEL00	REGPAY00
LANGWK00	LOOKNO00	NOINS99	PANEL99	RELCLS00
LANGWK99	LOOKNO99	NOIS299	PARENT00	RELCLS99
LASMED00	LOSEJO00	NOISE00	PARENT99	RELPHC00
LASMED99	LOSEJO99	NOISE99	PART99	RELPHC99
LASTYR00	LOSJOB99	NONEG00	PARTED00	RELVIS00
LASTYR99	LOSTJC00	NONEG99	PARTED99	RELVIS99
LEAVJB00	LOSTJC99	NOSKIL99	PAYEND00	RENT00
LEAVJB99	LSTWK200	NOTHIR00	PAYEND99	RENT99
LFCPS00	LSTWK299	NOTHIR99	PAYPAR00	RESTDK99
LFCPS99	LSTYR00	NOTPRO00	PAYPAR99	RETAGE99
LFPRT00	LSTYR99	NOTPRO99	PDAYCA00	RETIRE99
LFPRT99	LTERM99	NOWORK00	PDAYCA99	RTTIME00
LIFEMP00	LUNG00	NWREAS00	PEN99	RURAL00
LIFEMP99	LUNG99	OCC100	PEN299	RURAL99
LIFSAT00	MAIRET99	OCC199	PENADJ99	SAMPLE00
LIFSAT99	MANYJ00	OCC200	PENPLN00	SAMPLE99
LIFT00	MANYJ99	OCC299	PENPLN99	SAV99
LIFT99	MARIED00	OCCL199	PENTY299	SAV299
LIGH299	MARIED99	OCCL200	PENTYP00	SELFEM00
LIGHT00	MBAR99	OCCL299	PENTYP99	SELFEM99
LIGHT99	MBAR299	ONE00	PERJOB00	SHIFT00
LIMACT00	MDRINK99	ONE99	PERJOB99	SHIFT99
LIMACT99	MEDOFC00	ONEJOB00	PERSON00	SHOP299
LIMHO200	MEDOFC99	ONEJOB99	PERSON99	SHOPS00
LIMHO299	MEDREG00	ORIFIR00	PREFER00	SHOPS99
LIMHOU00	MEDREG99	ORIFIR99	PREFER99	SICKLV00
LIMHOU99	MEDVIS00	ORIG199	PREMI00	SIT00
LIMWR200	MEDVIS99	ORIG299	PREMI99	SIT99
LIMWR299	MGAP99	ORIG399	PRESOC00	SLEEP00
LIMWRK00	MIGRAI00	ORIG499	PRESOC99	SLEEP99
LIMWRK99	MIGRAI99	ORIG599	PROBLE00	SMAGE00
LKAD00	MISSWK00	ORIG699	PROBLE99	SMAGE99
LKAD99	MOHRS100	ORIG799	PROJEC00	SMBAR99
LKEMP00	MOHRS199	ORIG899	PROJEC99	SMDAY00
LKEMP99	MOHRS200	ORIGIN	PROMO00	SMDAY99
LKFRND00	MOHRS299	ORIHIR00	PROMO99	SMMORE99
LKFRND99	MOST99	ORIHIR99	RACE	SMOKE00
LKING00	MOST299	ORINO00	RACE_DK	SMOKE99
LKING99	MREST99	ORINO99	RACE_OTH	SMPAST00
LKOTHE00	MREST299	ORIPRO00	RAGE99	SMPAST99
LKOTHE99	MWORK99	ORIPRO99	REASO00	SMREST99
LONGEM99	MWORK299	OTHACT99	REASO99	SMWORK99
LONGER00	NCHILD00	OVERCO00	REASO100	SMYRS00
LONGER99	NCHILD99	OVERCO99	REASO199	SMYRS99

Appendix II -- Alphabetical variable lists, continued

1999-2000 longitudinal file, continued

SOCA00	ULCER00	WORKN99
SOCA99	ULCER99	WORKS00
SOCIAL00	UNEMPR00	WORKS99
SOCSE299	UNEMPR99	WORKTI00
SOCSEC99	UNFIRE00	WORKTI99
SPIRIT00	UNICOV00	WORKYR00
SPIRIT99	UNICOV99	WORTH00
SPORIT99	UNIMEM00	WORTH99
SPOSOC99	UNIMEM99	YEAR00
SPOUS299	USAGE	YEAR99
SPOUSE99	VISION00	YENTER
SSI99	WALK00	YOURWA00
SSI299	WALK99	YOURWA99
SSOC99	WEIGHT00	
SSOC299	WEIGHT99	
STAIR00	WEKWRK00	
STAIR99	WEKWRK99	
STIFFJ00	WELFAR00	
STOOP00	WELFAR99	
STOOP99	WGT9900L	
SUFINC00	WGT9900X	
SUFINC99	WGTEX1K	
SUPER00	WGTTOT99	
SUPER99	WHITE	
TAKJOB00	WHOME00	
TAKJOB99	WHOME99	
TALL	WIC00	
TANF00	WIC99	
TANF99	WKDNK99	
TEMPAY00	WKSLOK00	
TEMPAY99	WKSLOK99	
TIME00	WONDER00	
TIME99	WONDER99	
TIRED00	WORK00	
TMPHLT00	WORK99	
TMPHLT99	WORK1800	
TOOLS00	WORK1899	
TOOLS99	WORKD00	
TOTADU00	WORKD99	
TOTADU99	WORKE00	
TOTCHI00	WORKE99	
TOTCHI99	WORKED00	
TRAS299	WORKED99	
TRASH00	WORKER00	
TRASH99	WORKER99	
TYPEMC00	WORKMO00	
TYPEMP99	WORKN00	

Appendix III – Notes on the 1999 – 2000 longitudinal file and development of the longitudinal weights

The 1999-2000 longitudinal CWSHS file contains all responses from both years' interviews from those participants who were interviewed both years. There are 1,265 such respondents. The records from the two years have been merged so that there is only one record per person, containing all responses from the two years.

In all, 62% of the 2,040 1999 participants responded to the 2000 survey. Most non-response was due to the inability to re-contact the participant, rather than an active refusal on the part of the participant. Of the 775 1999 respondents who were not followed in 2000, 135 (17%) were refusals at the time of re-contact in 2000; an additional 76 (10%) had indicated in 1999 that they did not wish to be interviewed again. The remaining 564 (73%) were lost to follow-up, including 4 deceased, 2 institutionalized, and 5 who moved out of state.

Continuing participants differed from non-continuing participants in many ways. They were older, more likely to be white, to own their home, to speak English in the home, or to have a college degree, and less likely to report high levels of perceived stress (see question 129 in questionnaire), or have household incomes under \$20,000 than those who were not interviewed in 2000. (There were, however, no differences based on follow-up status in the likelihood of employment, frequency of chronic illness or disability, or language of interview).

Because of these differences, an adjustment to the 1999 sampling weights is needed that accounts for the differing probability of being followed among different sub-groups of the population. This adjustment for attrition would increase the weight of a person who, based on his/her known characteristics, was unlikely to have been followed-up in 2000, and reduce the weight of someone who is likely to have been followed. For example, white, college-educated respondents in their 50's were fairly likely to be re-interviewed in 2000; in fact, 69% of this group remained in the survey. Therefore, their sampling weights would be adjusted downward, because they make up a relatively larger proportion of the longitudinal dataset than they do of the initial 1999 dataset. The reduction in their sampling weight by use of the adjustment variable described below prevents their characteristics from being overestimated in the longitudinal dataset. Without such an attrition adjustment, one would lose the ability to generalize the findings to the California adult population despite the use of the sampling weights.

To calculate this adjustment for attrition, we developed a logistic regression model that predicted survey participation in 2000. The model contained the following independent variables: gender, age, age², race/ethnicity (African American, Latino, Asian/Pacific Islander, vs. non-Latino White), and household income (below 125% of Federal Poverty Level, unknown household income, vs. above 125% FPL).

We took steps to determine that the model adequately fit the data; specific results of the logistic regression model and influence diagnostics are available upon request. Next, we output the predicted probability of being interviewed in 2000, and took its inverse as the attrition adjustment. Thus, persons in the longitudinal file who were more likely to have been followed based on their known characteristics have small values of the attrition adjustment, which reduces the size of their longitudinal weight relative to others in the longitudinal sample. The size of this adjustment ranges from 1.4 to 4.0.

Appendix III – Notes on the 1999 – 2000 longitudinal file and development of the longitudinal weights, continued

To create the longitudinal weight, we multiplied the attrition adjustment (ATTRADJ) by the original 1999 sampling weight (WGTTOT99). Finally, we normalized the longitudinal weights so that the sum of these values is equal to the total sample size of 1,265. The longitudinal file contains two 1999-2000 longitudinal weight variables, a proportional weight (WGT9900L) and an expansion weight (WGT9900X). In addition, the attrition adjustment variable (ATTRADJ) and the original 1999 sampling weight variable (WGTTOT99) are retained in the longitudinal file. In most cases, however, users will not need either the attrition adjustment nor the original 1999 sampling weight variables. They are included on the file for users experienced in longitudinal sampling issues who may wish to investigate the construction of the longitudinal weights.