UnderStandingAmericaStudy

UAS COVID-19 LONGITUDINAL FILES DATA DESCRIPTION

Contents

1. INTRODUCTION

This document describes the UAS Covid-19 National Sample Longitudinal File and the UAS Covid-19 Los Angeles County Sample Longitudinal File of the Understanding America Study (UAS). These files were created at the University of Southern California's Center for Economic and Social Research (CESR). Main support for this data set has been provided by the University of Southern California and the Bill & Melinda Gates Foundation.

The UAS Covid-19 National Sample Longitudinal File and the UAS Covid-19 Los Angeles County Sample Longitudinal File (known collectively as the Covid longitudinal files) are comprised of online survey data collected by CESR through its UAS internet panel. The UAS is a nationally representative panel of American households randomly recruited from United States Postal Service delivery sequence files. UAS respondents are individuals age 18 and older who answer surveys via an online interface that is technologically powerful yet friendly for the participants and quick to deliver results. More information about the methodology of the UAS is available from the UAS data pages.

The data in the Covid longitudinal files consist of appended data sets from the UAS *Understanding Coronavirus in America* ("Covid") tracking surveys of attitudes and behaviors around the Novel Coronavirus pandemic in the United States. The first 29 waves of these surveys were fielded on a rolling 14-day or 28-day schedule between Marc 10, 2020 and July 20, 2021, and included a national long-form questionnaire and a Los Angeles County short-form questionnaire administered in each wave[1]. Five follow-up surveys have also been administered: wave 30 on September 23, 2021, wave 31 on Februrary 1, 2022, wave 32 on July 20, 2022, and wave 33 on November 18, 2023, and wave 34 on May 12, 2023.

The Covid longitudinal files have a consistent variable naming structure so that the content associated with each variable is consistent from wave to wave in the data. Appending the individual UAS Covid survey data sets required renaming variables from their original names if they were used in different surveys to ask different questions. To assist users of the Covid longitudinal files in mapping original UAS Covid survey variables to their counterparts in the longitudinal files, the UAS Covid-19 Survey Longitudinal Files Crosswalk was created. This document connects variable names from each UAS Covid survey to their names in both the national sample longitudinal file and the Los Angeles County sample longitudinal file.

The details contained in this document pertain to both the UAS Covid-19 National Longitudinal Sample File and the UAS Covid-19 Los Angeles County Sample Longitudinal File, except where differences are explicitly stated.

---

[1] The structure of the first wave differs from later waves; see Section 2.3.1 for details

The Covid-19 longitudinal files can be linked with any other UAS survey. How to perform that linkage is described later in this document. An updated listing of all available UAS data sets is available on the UAS All Surveys page.

Please send all questions about UAS data sets or this data description to uas-l@mymaillists.usc.edu.

2. OVERVIEW OF THE UAS COVID-19 LONGITUDINAL FILES

2.1 Obtaining UAS Covid-19 Related Data

The data described in this document are based on UAS public data release files. These files are accessible through the UAS Covid-19 National Survey Data and the UAS Covid-19 Los Angeles County Survey Data webpages.

Before accessing any UAS data, permission must first be obtained by registering on the UAS site. Registering with UAS is considered to be agreeing to the "conditions of use" governing access and usage of the data. Completion and submission of an additional form is required to access UAS Covid-19 related data. These steps are described in detail on the UAS Covid-19 Data Access page.

2.2 Input Files

The current Covid longitudinal files are updated with data from the individual UAS Covid surveys as they are concluded, a new wave every four weeks. Table 1 on the next page lists the concluded waves through Jul 10, 2023.

Table 1. Fielding Dates for UAS Covid Surveys in the Longitudinal Files

Wave	UAS Surveys	Launch Date	Closing Date
1	230	3/10/2020	3/31/2020
2	235, 236	4/1/2020	4/27/2020
3	240, 241	4/15/2020	5/11/2020
4	242, 243	4/29/2020	5/25/2020
5	244, 245	5/13/2020	6/8/2020
6	246, 247	5/27/2020	6/22/2020
7	248, 249	6/10/2020	7/6/2020
8	250, 251	6/24/2020	7/20/2020
9	252, 253	7/8/2020	8/3/2020
10	254, 255	7/22/2020	8/17/2020
11	256, 257	8/5/2020	8/31/2020
12	258, 259	8/19/2020	9/14/2020
13	260, 261	9/2/2020	9/28/2020
14	262, 263	9/16/2020	10/12/2020
15	264, 265	9/30/2020	10/26/2020
16	266, 267	10/14/2020	11/9/2020
17	268, 269	10/28/2020	11/23/2020
18	270, 271	11/11/2020	12/7/2020
19	272, 273	11/25/2020	12/21/2020
20	274, 275	12/9/2020	1/4/2021
21	276, 277	12/23/2020	1/18/2021
22	278, 279	1/6/2021	2/1/2021
23	280, 281	1/20/2021	2/15/2021
24	282, 283	2/3/2021	3/1/2021
25	340, 341	2/17/2021	3/29/2021
26	342, 343	3/17/2021	4/27/2021
27	344, 345	4/14/2021	5/25/2021
28	346, 347	5/12/2021	6/22/2021
29	348, 349	6/9/2021	7/20/2021
30	350	9/23/2021	10/31/2021
31	351	2/1/2022	3/30/2022

32	352	7/20/2022	9/9/2022
33	353	11/18/2022	1/15/2023
34	354	5/12/2023	7/9/2023

Each UAS Covid-19 survey is classified as either a long-form or a short-form survey. Long-form surveys were administered to the national sample of respondents (including Los Angeles County residents). During waves 1 - 25 they were administered on a biweekly basis, while during waves 26 - 29 they were administered on a four week schedule. Subsequent waves, 30 - 34, were not administered with a regular fielding schedule, but as standalone long-form surveys that were fielded several months after the previous wave had closed. Long-form surveys consist of more questions and more topics than short-form surveys and so generally take longer for respondents to complete. The data from the long-form surveys are used to create the UAS Covid-19 National Sample Longitudinal File.

Short-form surveys were administered to Los Angeles County residents only. During waves 2 - 25 they were administered on a weekly basis, while during waves 26 - 29 they were administered on a biweekly schedule. No short-form survey was administered in wave 1 or waves 30 - 34. The short-form surveys chiefly contain a subset of questions from the long-form surveys, though they may also contain questions not found in the long-form surveys which mainly pertain to topics related specifically to the greater Los Angeles area. The data from the short-form surveys, combined with the data from long-form surveys solely pertaining to Los Angeles County residents, are used to create the UAS Covid-19 Los Angeles County Sample Longitudinal File.

Included in each longitudinal file are variables that indicate the survey number that corresponds to the UAS survey source for each data observation (variable name: *survey_source*), as well as the wave number associated with all data from each specific UAS Covid survey (variable name: *wave*).

Los Angeles County sample respondents were typically administered two surveys during each wave of data collection (1 long-form and 1 short-form survey), therefore the structure of the wave variable in the Los Angeles County longitudinal file differs from the structure of the wave variable in the national sample longitudinal file. Tables 2A and 2B illustrate this difference. Section 2.3 of this document explains the differences in the two samples in greater detail.

Table 2A. National Sample Longitudinal File Information (Example Only[2])

Wave Number	UAS Survey	Survey Type
1	230	Long-form
2	235	Long-form
3	240	Long-form
4	242	Long-form

Table 2B. Los Angeles County Sample Longitudinal File Information (Example Only)

Wave Number	UAS Survey	Survey Type
1.1	230	Long-form
2.1	235	Long-form
2.2	236	Short-form
3.1	240	Long-form
3.2	241	Short-form
4.1	242	Long-form
4.2	243	Short-form

Only 1 survey was administered in Wave 1

Table 2A shows the numbering of the wave variable in the national sample longitudinal file as an integer that represents the one long-form survey collected during each wave. Table 2B shows that the numbering of the wave variable in the Los Angeles County longitudinal file is broken down into 2 parts:

- The number to the left of the decimal point represents the overall wave of data collection.

- The number to the right of the decimal point, in which a value of 1 indicates the long-form survey data from the wave, and a value of 2 indicates the short-form survey data from the wave.

---

[2] The current UAS Covid-19 longitudinal files contain more than 4 waves of data. Tables 2A and 2B demonstrate the wave numbering and corresponding survey information for descriptive purposes only.

2.3 Sampling Frame and Fielding Process of the UAS Covid-19 Surveys

2.3.1   National Sample

The UAS Covid-19 National Sample Longitudinal File consists of data from respondents of the Covid long-form surveys.

The first survey was fielded on March 10, 2020 and was in the field until April 1, 2020. In contrast to subsequent survey waves 2 - 29, all respondents for this first survey were invited on March 10. The distribution of responses over the survey period is therefore not random and concentrated in the first part of the survey period (see the Survey Methods tab on the Covid-19 Pulse site).

For survey waves that started April 1, 2020 through February 3, 2021 (waves 2 - 24) a new survey was fielded every two weeks. Each day one fourteenth of the respondents were invited to take the survey. Since respondents had two weeks to answer the survey, the total field period was 4 weeks, so that responses during the last two weeks of a field period of one survey overlapped with responses in the first two weeks of the subsequent survey. Respondents were paid an additional incentive to complete their surveys on their assigned day, and as a result, more than 90% of responses to each wave were completed in the first 14 days.

For surveys that started February 17, 2021 through June 9, 2021 (waves 25 - 29), a new survey was fielded every four weeks. Each day 1/28th of the respondents were invited to take the survey. Since respondents still had exactly two weeks to answer the survey, the total field period was three weeks. Respondents were paid an additional incentive to complete their surveys on their assigned day, and as a result, more than 90% of responses to each wave were completed in the first 28 days.

The four follow-up surveys (waves 30 - 34) that began on September 23, 2021, February 1, 2022, July 20, 2022, November 18, 2022, and May 9, 2023 , respectively, had invite schedules similar to that of wave 1, with all respondents receiving invites near the beginning of the fielding period.

2.3.2   Los Angeles County Sample

The UAS Covid-19 Los Angeles County Sample Longitudinal File consists of data from respondents of the UAS Covid short-form surveys, and from long-form survey respondents who reside in Los Angeles County.

The first wave and last two waves of data for Los Angeles County (wave 1 and waves 30 - 34) are simply subsets of the national data, only including Los Angeles County residents. Waves 2 - 29 of the data collection for Los Angeles County consisted of two components, the national survey (long-form) that was also administered to the national sample and a shorter survey (short-form) administered to Los Angeles County residents only.

For survey waves that started April 1, 2020 through February 3, 2021 (waves 2 - 24), the long-form survey was administered once every fourteen days. The short-form survey had the same 14

day periodicity, but was administered in the alternate weeks. To give an example, if a Los Angeles County respondent was invited to a survey on Monday, April 6, for the national survey then the same respondent was invited for a shorter survey on Monday, April 13. Conversely, a respondent who was invited for a national survey on Monday, April 13, was invited for a short survey on Monday, April 6, etc. Since the content of the short survey is mainly a subset of the questions in the national survey, the result is that for the short survey, we collected data at a weekly frequency in Los Angeles County, rather than every other week. This set-up implies that a respondent in Los Angeles County had only one week to answer a survey, rather than two weeks as in the rest of the country.

For surveys that started February 17, 2021 through June 9, 2021 (waves 25 - 29) the periodicity was expanded to 4 weeks. The long-form survey was administered once every 28 days, while the short-form survey had the same 28 days periodicity but was administered either two weeks earlier or two week later.

2.4 UAS Covid-19 Longitudinal Files Data Formats

The longitudinal files are available for download in STATA, CSV[3], and SAS format from the National Survey Data page and Los Angeles County Survey Data page. If you would like the data in a different format, please email your request to uas-l@mymaillists.usc.edu.

2.5 Helpful Links

For easy reference, the following are links to pages and files that are mentioned in this document:

UASDATA Covid-19 home page

Data Access page: Lists the steps for how to gain access to UAS survey data, and also specifically, UAS Covid-19 survey data.

National Survey Data page: Contains links to the national sample longitudinal file download and documentation, each individual Covid survey questionnaire, codebook, and file download, as well as links to reports that use national survey data.

Los Angeles County Survey Data page: Contains links to the Los Angeles County sample longitudinal file download and documentation, each individual Covid Los Angeles County survey questionnaire, codebook, and file download, as well as links to reports that use Los Angeles County survey data.

---

[3] The CSV output is split into three files to limit any issues caused by opening a very large CSV file: one containing waves 1 - 10, a second containing waves 11 - 20, a third with waves 21 – 30, and a fourth containing all subsequent waves.

UAS Covid-19 Survey Longitudinal Files Crosswalk

UAS All Surveys page: Contains the complete list of all UAS surveys that have been administered, and provides the codebook, data download link, and other information about each UAS survey.

My Household Survey Data page: Contains information about the My Household survey, which is demographic and socioeconomic information updated quarterly by each UAS participant. More information about My Household is included in Section 3.2.2.

UAS Standard Variables page: Lists the full set of descriptive variables that are included in each data set.

UAS Sample Weights page: Provides detailed information about UAS weighting procedures

2.6 Merging with other UAS Data Sets

The Covid longitudinal files can be merged with any other UAS data set by using the unique person identifier variable, *uasid*. This variable, assigned to a respondent at recruitment into the UAS panel, stays fixed for each survey taken. The longitudinal files are panel data sets, meaning that multiple records are likely for each uasid; one for each instance a respondent completed a Covid UAS survey.

For more information about default identification variables included in each survey, such as household identifier, *uashhid*, please visit the UAS standard variables page.

3. UAS COVID-19 LONGITUDINAL FILES DESCRIPTION

The UAS Covid-19 National Sample Longitudinal File and the UAS Covid-19 Los Angeles County Sample Longitudinal File both contain data from survey questions covering many aspects of how people are being affected by the coronavirus pandemic. In addition, derived variables, survey information variables, and sample weights are included in the files. This section describes the situations that require original UAS Covid survey variables to be renamed in the longitudinal files, and also provides some detail about the content included in the longitudinal files.

3.1 Variable Renaming

Efforts are made when creating each longitudinal file to preserve the original variable names from each UAS Covid data set. However, there are circumstances that require original variable names to be changed to assure consistent naming of identical questions across the individual surveys. These circumstances are:

1. The same question and response options have different original variable names in two or more UAS Covid surveys.

For example, original variable *cr005* from UAS 230 and original variable *cr023* from UAS 235 both ask:
"On a scale of 0 to 100 percent, what is the chance that you will get the coronavirus in the next three months? If you're not sure, please give your best guess."

Therefore, both variables have been renamed to *prisk_infection* in the longitudinal files.

2. Different question text and/or response options have the same original variable name in two or more UAS Covid surveys.

For example, *cr002* from UAS 230 asks:
"Has a doctor or another healthcare professional diagnosed you with the coronavirus (COVID-19)?"

While *cr002* in UAS 235 asks:
"Have you been tested for coronavirus? If so, what was the result?"

In this case, *cr002* from UAS 230 has been renamed to *covid_diagnosed* and *cr002* from UAS 235 has been renamed to *cor_tested* in the longitudinal files.

To quickly identify the longitudinal file variable name equivalent to an original Covid survey variable name, please use the UAS Covid-19 Survey Longitudinal Files Crosswalk. The crosswalk has separate sections for the national sample file and the Los Angeles County-only sample file. This document will be updated as each new wave of data is added to the longitudinal files.

When question text or response options differ only slightly, the original variable name is preserved in the longitudinal files. Each instance is reviewed by the UAS team manually and a decision is made on whether to preserve the original variable name or to rename the original variable into two new variable names that reflect the different content from each survey. For cases where the decision is made to preserve the original variable name despite a slight difference, a note is attached to the variable in the longitudinal data sets that explains the difference and identifies in which UAS survey it first occurred.

A complete list of variables in the longitudinal files with slight differences in crosswave question text or response options (but maintaining the same variable name) is given in Table A.1 of Appendix A. Table A.1 identifies the wave of data collection for UAS Covid surveys in which the change occurred. This includes situations where the sole difference of a variable between two waves is an added answer option. For example, the variable *beh_facemask* in the longitudinal files asks whether you have worn a facemask in the past seven days to keep yourself safe from the coronavirus. In UAS 230, the answer options are Yes or No. In UAS 235, the answer options are Yes, No, Unsure.

3.1.1 Specific Variable Renaming and Response Recoding Changes

- For original survey questions cr031a through cr031d, in UAS 235 there were six answer options. For these same variables, starting in UAS 240 the number of answer options was

reduced to four. As a result, the data in the longitudinal files for these questions from UAS 235 is stored in the variables named *cor_percp_uas235_1* through *cor_percp_uas235_4*, and for UAS 240 and forward the variables are named *cor_percp_1* through *cor_percp_4*.

- Also for questions cr031a through cr031d, a small number of respondents using the Spanish language questionnaire were inadvertently provided with a six-value answer option that had been replaced by a four-value option in the English language questionnaire, a change that was missed during translation. We were unable to reliably map the six values to the four values, so have removed and replaced them with a missing value of ".d Deleted value" in the published UAS data files as well as in the UAS Covid longitudinal files. Access to the removed responses is available upon request. Starting with UAS 248, all respondents see the four-value response option. These variables have the names *cor_percp_1* through *cor_percp_4* in the longitudinal files.

- The variables *r_cl005_ddd* and *cl005_ddd_X_* (where X identifies the index of the person asked about) were single-response questions in UAS 240 and UAS 242. They were not asked in UAS 244, UAS 246, and UAS 248. They were asked again in UAS 250, but instead as multiple-response (i.e., checkbox or "Select all that apply") questions. To distinguish between the single-response and multiple-response versions of these variables, the multiple-response versions were renamed in the data files. The new variable names are:

    - *r_cl005_ddd_checkall*: Textual list of all responses selected
    - *r_cl005_ddd_checkalls1* through *r_cl005_ddd_checkalls4*: Yes/No variables that indicate whether a specific response option was selected in *r_cl005_ddd_checkall* (s1 represents the first option, s2 the second, etc.)
    - *cl005_ddd_checkall_X_*: Textual list of all responses selected for person index "X"
    - *cl005_ddd__checkall_X_s1* through *cl005_ddd__checkall_X_s4*: Yes/No variables that indicate whether a specific response option was selected in *r_cl005_ddd_checkall_X_* (s1 represents the first option, s2 the second, etc.)

These renamed variables are included in all applicable files: the individual UAS Covid downloadable files as well as in the UAS longitudinal files.

- In Covid-19 surveys UAS 244 through UAS 251, response options were originally coded differently in the Spanish-version questionnaire than in the English version for the variable *lr003*. The Spanish-version responses for *lr003* have been successfully recoded to match the English version responses in all files: the individual UAS Covid downloadable files as well as in the UAS Covid longitudinal files.

- In Covid-19 long-form surveys UAS 244 through UAS 262, the response options for lr029_month and lr032_month included all months from February to the month of survey fielding. From UAS 264 (fielded from 9/16/2020 - 10/13/2020) through UAS 280

(fielded 1/20/2021 - 2/16/2021) the response options were inadvertently fixed between February and September. Beginning in UAS 282, the response options included all months January through December.

- In UAS 350 & 351, variable *md001b* (whether support/oppose vax requirement for attending a large public event) had no specification for whether the event was outdoors or indoors. In UAS 352, this question was split, with *md001_outdoorevent* (which retains the original name *md001b* in raw survey data) asking about outdoor events only and *md001_indoorevent* (*md001h* in raw survey data) asking about indoor events only.

- In UAS 350 & 351, variable *md004b* (whether wear a mask when attending a large public event) had no specification for whether the event was outdoors or indoors. In UAS 352, this question was split, with *md004_outdoorevent* (which retains the original name *md004b* in raw survey data) asking about outdoor events only and *md004_indoorevent* (*md004e* in raw survey data) asking about indoor events only.

3.2 Longitudinal Files Content

The variables included in each longitudinal file can be classified into 4 main groups: UAS Covid surveys questionnaire data, UAS My Household survey data, variables derived from UAS Covid survey questionnaire data, and information about the surveys themselves. An overview of each follows.

3.2.1   UAS Covid Survey Questionnaire Data

The items asked in the UAS Covid surveys aim to measure the impact of the coronavirus pandemic on physical and mental health, personal finances, employment, education, outlook on the future, and many other aspects of daily life--as well as opinions and behaviors concerning vaccination, mask-wearing, and mandates. To view the individual questions asked in each Covid survey, please use the links to the survey codebooks questionnaires, and reports found on the National Survey Data page and Los Angeles County Survey Data page.

The content of the Los Angeles County short-form surveys is primarily a subset of the larger long-form national surveys. However, there are some items that are only asked in the short-form surveys, such as questions that pertain specifically to residents of Los Angeles County.

In both longitudinal files, if a variable is not included in a wave or (more rarely) if the question was skipped over for every respondent during that wave resulting in only missing vlaues, this variable's values have been set to a missing code of ".z var not in wv (or always skipped)" to simplify any missing data analysis performed by data users. Other missing codes found in the files, such as ".a" and ".e", should be treated as respondent non-response that is not caused by the question's absence from the survey questionnaire. For example, the cause could be

non-response due to questionnaire skip pattern logic, the survey was not fully completed, or the respondent did not know the answer or refused to answer the question.

Currently, no special outlier detection process is performed while creating the longitudinal files. Future versions may incorporate data cleaning techniques to identify and flag highly influential data observations.

3.2.2 My Household Survey Data

The My Household survey is administered quarterly to all UAS respondents. To provide background information about the respondent and household, the survey asks about key demographics, including age, ethnicity, education, marital status, work status, state of residence, family structure, and many others.

The most recent My Household survey response data at the time each respondent completes a UAS survey is included in the data set of that UAS survey. This data is included in the Covid longitudinal files corresponding to each wave in which the respondent has participated.

For more information about the My Household survey, please visit the My Household Survey Data page.

3.2.3 Derived Variables

This section describes the derived variables that were constructed specifically for the Covid longitudinal data sets by using original UAS Covid survey data.

- The individual UAS Covid surveys each include questions from the Patient Health Questionnaire-4 (PHQ-4). The PHQ-4 is a brief-screening tool used by clinicians for detecting both anxiety and depressive disorders (Kroenke, K. et al., 2009). The questions are presented in the UAS Covid surveys with answer choices 1 to 4. However, in the original PHQ-4 the answer choices are labeled 0 to 3 and then summed to arrive at a total score. To facilitate scoring of the UAS Covid PHQ-4 data, each response is first re-scaled by subtracting 1. These derived re-scaled variables have the names: *phq4_anxious_r*, *phq4_worry_r*, *phq4_depressed_r*, and *phq4_littleinterest_r*. These variables are then summed to create the derived variable, *phq4_score*.

- The variable *laborstatus_covidSurv* summarizes how responses to a set of labor force status questions (lr001[4], lr003[4], lr003aa[5], lr003bb[5], lr003cc[6], lr003dd[6], lr003ee[7], lr003a[8]) in each UAS Covid survey map to the following six labor force status categories: currently working, sick/other leave, unemployed-temporary layoff, unemployed-looking, retired (and so not in the labor force), and not in the labor force but not retired. The variable

---

[4] Introduced in wave 2.
[5] Introduced in wave 3.
[6] Introduced in wave 5.
[7] Introduced in wave 30.
[8] Question lr003a (introduced in wave 1) was not asked in waves 30 forward.

was calculated by both directly mapping the response options selected and, when respondents selected the "None of these" option, coding respondents' open-ended text.

There is also a seventh code "unknown labor force status" in *laborstatus_covidSurv*, which represents when the information provided by respondents could not be confidently coded among the other six categories. After inclusion of additional response options for question lr003 in wave 4 (UAS 242/243)[9] and a fix to allow open-ended responses for question lr001 in the wave 7 long-form survey (UAS 250)[10], this "unknown labor force status" code was only output in two circumstances: for respondents with ambiguous open-ended text, and for respondents whose only completed question among the labor force status question set was lr003a where their response to lr003a was that they did not not currently have a job.

The response option design in the questionnaire was finalized with the inclusion of a response option for "retired" in question lr001 during wave 8 (UAS 252/253). Thus, longitudinal comparisons between single categories of *laborstatus_covidSurv* across waves are most reliable from wave 8 forward.

Response options corresponding to "currently have a job" (whether asked directly in lr003a[11], or split between the two response options "currently working" and "sick/other leave") have always been fielded within the labor force status question set. Thus, longitudinal differences in whether or not respondents reported being employed are considered reliable across all waves.

Earlier releases of the longitudinal files included a variable *job_current* that represented whether a respondent was currently working or not. The variable *laborstatus_covidSurv* revises and expands *job_current*, with *job_current* = 1 being analogous to *laborstatus_covidSurv* = 1 "currently working" and *job_current* = 2 "not currently working" being analogous to the disjunction of *laborstatus_covidSurv* values 2 through 6.

3.2.4   Survey Information Variables
The following variables are included in both the national sample and the Los Angeles County sample longitudinal files as identifiers of each UAS Covid survey:

- wave: Indicator of the wave number associated with each UAS Covid survey. See Table 2A for information about how the wave number corresponds to the UAS survey number in

---

[9] See Appendix A for details.

[10] All open-ended responses were collected in the short-form surveys beginning with wave 3 (UAS 241) .

[11] A response of "yes, currently have a job" to the lr003a question gets mapped to the *laborstatus_covidSurv* code "currently working", though this question wording could be thought of as corresponding to the combination of "currently working" and "sick/other leave". Question lr003a was the entry question to the labor force status question set and was fielded to respondents once, at the time of their first Covid survey/wave.

the current national sample longitudinal file, and see [Table 2B](#) for this information pertaining to the Los Angeles County sample longitudinal file.

- survey_source: Indicates the UAS survey number associated with each UAS Covid survey. For example, if the data for an observation in either longitudinal file is from UAS 230, then the value of *survey_source* is 230.

- surveytype: Has a value of 1 if the UAS survey source for a data observation is a long-form survey, and a value of 2 if the UAS survey is a short-form survey. **This variable is only found in the UAS Covid-19 Los Angeles County Sample Longitudinal File.** Each observation in the national sample longitudinal file is from a long-form survey.

3.3 Sample Weights

The longitudinal files include weights in the variable *final_weight* to correct for differential nonresponse rates across demographic groups and align the sample to the U.S. adult population along key demographic variables: gender, race/ethnicity, age, education and census region. Values used to generate the sample weights are taken from the My Household survey information for each respondent in each wave of data.

The weighting procedure is implemented separately for the USA excluding California, California excluding Los Angeles County, and Los Angeles County. The resulting weights are then scaled up to the population size. Thus, when the entire sample is selected, the final weights allow to match the distributions of the aforementioned demographic variables in the entire U.S. adult population. If one of the three specific geographic sub-samples is selected – USA excluding California, California excluding Los Angeles County, Los Angeles County – the final weights allow to match the distributions of the aforementioned demographic variables in each of these geographies. A complete description of the UAS weighting procedure for the UAS Understanding Coronavirus in America study can be found [here](#).

In the longitudinal files weights are wave-specific. That is, they ensure representativeness of the sample in each wave. Because of that, they should be used for the cross-sectional analyses of single waves of data and to examine the evolution of variables of interest over time in repeated cross-sections. The provided final weights are not longitudinal weights. As such, they are not meant to make the sample of UAS members who ever participated in the COVID survey representative of the population. To request longitudinal weights and/or for any questions about the weights contact uas-weights-l@mymaillists.usc.edu.

REFERENCES

Kroenke, K. & Spitzer, R. & Williams, J. & Löwe, B. (2009). An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4. Psychosomatics. 50. 613-21. 10.1176/appi.psy.50.6.613.

APPENDIX A. VARIABLES WITH CROSSWAVE CONTENT DIFFERENCES

Table A.1 Variables with Cross-wave Content Differences in the Longitudinal Files

Wave	UAS Survey	Content Changes
2	UAS 235 & UAS 236	• Addition of response option in the following variables: *beh_avoidhighrisk, beh_avoidpublic, beh_avoidrest, beh_canceldrappt, beh_facemask, beh_pleasuretravel, beh_pray, beh_socialactivities, beh_stockpilefood, beh_visitdr, beh_washhands, beh_worktravel, beh_workactivities, beh_workhome* The response option "Unsure" was added.
	UAS 235 only	• Change to question text in variable *coping_exercise*: The phrase "extra exercise" in the question text was changed to simply "exercise" in UAS 235.
3	UAS 240 & UAS 241	• Change to question text for the following variables: *cor_tested, covid_diagnosed, think_havecorona, cor_care, cor_care_hosp, cor_care_local, cor_care_pcp, cor_care_urgent, cor_care_where, cor_care_none, cor_care_other, cor_contact, cor_contact_comm, cor_contact_emp, cor_contact_fam, cor_contact_hosp, cor_contact_local, cor_contact_online, cor_contact_oth, cor_contact_pcp* These questions are concerned with whether respondents had experienced or informed others about experiencing being tested for, diagnosed with, infected with, or sought medical care for coronavirus. The time frame for these questions was changed to ask about any experiences since the previous interview. Previously, respondents were asked if they had ever in their lives had one of these experiences. • **Change to response option in variable *lr003*:** The text "please specify" was added to option 3 'none of these' and open-ended responses were now collected.
	UAS 240 only	• **Subtraction of response options in variable *ins_randomizer_specific*:** Options available now limited to 1 '5%' through 4 '20%'. In UAS 230, options 5 '30%' through 11 '90%' were also available. • **Change to question text in variable *lr009*:** This question asked if respondent's work hours had been reduced. The time frame for these question was changed to ask if hours had been reduced since the previous interview. Previously, respondents were asked if hours had been reduced since March 1, 2020.
	UAS 241 only	• **Change to response option in variable *lr001*:** The text "please specify" was added to option 3 "none of these" and open-ended responses were now collected. This would remain in place for all subsequent short-from surveys. Due to an oversite, the "please specify"and open-ended response collecting was not added to the long-form surveys until UAS 248.
4	UAS 242 & UAS 243	• **Addition of response options in variable *lr003*:** Added the following two options for this labor force status question: 4 'I am retired', 5 'I am not in the labor force (not currently working and not looking for work)'.

	UAS 242 only	• **Addition of response option in the following variables:** *cl006b_1_, cl006b_2_, cl006b_3_, cl006b_4_, cl010aa_1_, cl010aa_2_, cl010aa_3_, cl010bb_1_, cl010bb_2_, cl011b_1_, cl011b_2_, cl013_1_, cl013_2_, cl013_3_, r_cl006b, r_cl010aa, r_cl010bb, r_cl011b, r_cl013.* Added option 12 'institution does not plan for in-person enrollment in the fall' for each of these questions.
5	UAS 244 only	• **Addition of response option in variable *lr010*:** Added the option 6 'Twice a month' to the question about how often the respondent is paid by employer.
7	UAS 248 & UAS 249	• **Change to survey instrument routing condition for *lr016*:** Prior to wave 7, question *lr016* was only asked when *lr004* (currently employed or not) = 2 'not currently working'. Starting in wave 7, *lr016* was fielded to all respondents.
	UAS 248 only	• **Change to response option in variable *lr001*:** The text "please specify" was added to option 3 'none of these' and open-ended responses were now collected.
9	UAS 252 & UAS 253	• **Addition of response option in variable *lr001*:** Added option 7 'I am now retired' to this labor force status question.
	UAS 252 only	• **Addition of response options in variable *lr033*:** Added options: 6 'I went back to work', 7 'I am still being paid by employer', 8 'I did receive benefits at some point' to the question about receiving unemployment benefits.
	UAS 252 only	• **Addition of response options in variable *lr034*:** Added options: 7 'I am retired', 8 'I receive disability or Supplemental Security Income', 9 'I am temporally on leave from my job /summer break /seasonal' to the question about the reason for not applying for unemployment benefits.
	UAS 253 only	• **Addition and deletion of response options in variable *fd009a*:** Added options: 13 'Full service restaurant or café', 14 'Fast food restaurant', 15 'Food donations from a faith-based organization, or other community-based organization'. Removed optons: 6 'Restaurant or café', 7 'Home garden', 10 'CSA (Community Supported Agriculture)', 11 'Meal kit subscription' from the question about the most frequent sources of food.
32	UAS 352	• **Change to question text in variable *sl082*:** In UAS 352, the age range in the question text changed from "5 - 18" to "5 - 17."