

README file for Public Update MIDUS 2 Project 1 Survey Data

December 2020

*** It is important to read through this document carefully ***
*** prior to using the revised data and documentation. ***

This document outlines a number of revisions, improvements, and updates that have been made to the MIDUS 2 Project 1 (M2P1) survey data since the last version published in 2011.

Note: The revised data and documentation are intended to REPLACE all the files associated with previous releases of the M2P1 data and documentation.

A. Changes to Datasets

- a. Updated: The M2P1 Survey dataset has been updated to reflect a number of fixes and improvements. The M2P1 dataset is a rectangular dataset comprised of Phone interview and Self-Administered Questionnaire (SAQ) data for 4,963 cases that includes 2,189 variables.

The updated dataset: *M2_P1_SURVEY_N4963_20200720*

- b. Updated: Final M2 Survey disposition codes detailing final outcome and reasons for non-response to the telephone interview at MIDUS 2.

The updated dataset: *M2_DispositionCodes_N7108_20201105* (For details, see *M2_P1_DispositionCodeDocumentation_20201103*). Researchers interested in the latest mortality data for the MIDUS Core sample should consult ICPSR (<https://www.icpsr.umich.edu/web/ICPSR/studies/37237>) or the MIDUS Portal (<https://midus.colectica.org/>).

- c. New: Qualitative text and open-ended response data has been processed into coded variables.

The new dataset: *M2_P1_CodedTextData_N4963_20201105* (For details, see *M2_P1_DocumentationOfCodedTextData_20201103*)

- d. New: SAQ data is available as a result of an effort to obtain SAQ data from M2 non-responders in 2006. This dataset contains 370 variables on 365 cases.

The new dataset: *M2_SAQxPhone_N365_20191206* (For details, see *M2_P1_SAQxPhone_README_20201103*)

B. Revisions in the Updated M2P1 Data

- a. The age (B1PAGE_M2) and birth year (B1PBYEAR) variables have been updated with new versions (B1PRAGE_2019 and B1PBYEAR_2019, respectively). Both the previous and updated versions of these variables are included in the updated M2 dataset. Two new

variables (M2BYEAR_FLAG, M2AGE_FLAG) summarize the small number of cases that experienced a change in age or birth year.

- b. The M2 post-stratification weight variables have been added to the M2P1 survey dataset – they previously were included in a separate dataset. Post-stratification weight variables were revised slightly to be consistent with how similar weight variables are derived and constructed in other MIDUS survey datasets. A new weight variable (accounting for marital status) was also introduced and the education weight's education levels were re-categorized to be consistent with other education weight variables. More details can be found in **M2_P1_DocumentationOfWeights_20201103.docx**.
- c. The variable labels for many of the M2 survey variables have been updated to provide more descriptive labels that take into account MIDUS metadata and coding conventions as well as current software capabilities.
 - The values and codes for missing values (INAPP, DON'T KNOW, REFUSED/MISSING) and for skipped, filtered, or branching values have been updated for fixed for many variables.
 - Previously, cases that did not complete the SAQ simply had empty cells for all their SAQ variables. Now, SAQ variables are coded -1 "RESPONDENT DOES NOT HAVE SAQ DATA" for those cases that did not complete the SAQ.
- d. Interview date B1PDATE was removed and replaced with two new variables: month B1PDATE_MO and year B1PDATE_YR.
- e. New constructed variables were added (B1PHRTRS B1PHRTDX B1PANGIN B1PCACRS B1PCACDX B1PPARTN) to be consistent with other waves of data.
- f. Standard occupational classification codes (SOC) were added for respondent's current occupation (B1PSOC).
 - Updated missing R's current occupation for a small number of cases (n=43).
- g. Two new variables (B1PKIDHH, B1PHHSIZE) were added to better reflect HH and family composition.
 - Removed a number of empty HH Roster variables.
 - Removed variable B1PC5 because it duplicated B1PC4.
- h. Removed administrative variable B1PDCSX.
- i. Changed value labels of B1SE11A to B1SE11AA from YES/NO to CHECKED/UNCHECKED.
 - Recoded B1SE11 Life Events effect variables scale to remove negative values.
- j. Restored original extreme values (+/- 4 SD) for B1SWSTHI.
- k. Fixed coding discrepancies for B1SA43 series of variables.
- l. Recoded all the age at adoption variables (B1PHHA1, B1PHHA2, B1PHHA3, etc.) to

INAPP if relation was not 'adopted child.'

- m. Income variable B1SG12 ceiling set at \$50,000 for consistency with other variables and datasets.
 - B1SG24A was split into two variables (amount money left or owed) B1SG24A and B1SG24B to be consistent with other waves.
- n. Corrected variable names for B1SIADL series (changed to B1SMADL) as well as labels ('Instrumental Activity of Daily Living' to 'Intermediate Activity of Daily Living').
- o. For the 'check all that apply' variables in SAQ, if R answered at least 1 of the items within a question list all other unanswered items were recoded from REFUSAL (8) to NO (2).

C. Changes in Status of Cases

- a. Two pairs of twins who shared the same family ID (M2FAMNUM) were discovered to be related by marriage but not by biology. One pair of the twins was given a different family ID to reflect this relationship (M2FAMNUM changed from 120166 to 120967).
- b. Another two pairs of twins were discovered to be related by marriage but not by biology. One pair of the twins was given a different family ID to reflect this relationship (M2FAMNUM changed from 120188 to 120968).
- c. The MIDUS Administrative Core suspects that in a small number of instances the wrong respondent was interviewed at M2 follow-up. These errors were discovered in subsequent data collection and likely due to individuals having the same or similar names (such as parents and children) as the original M1 respondent, and/or cases having incomplete baseline information (full DOB or SSN) to confirm identify. Researchers may want to exclude these cases from analysis: M2ID = 13088, 13757, 14061, 15723, 16469, 16687. While a few of these cases (M2ID = 13757, 15723, 16469, 16687) are also found in the M2P3 (BTACT) dataset, no other M2 projects were affected by these cases.

D. Updated MIDUS 2 Documentation

- a. **M1_P1_DocumentationOfUpdatedAge_20201103.** Describes updates and corrections to age and birth year information in the M1 baseline survey data as well as why these changes were necessary. These changes affected a small number of M2 cases as well.
- b. **M2_P1_DocumentationOfPsychosocialConstructsAndCompositeVariables_20201103.** Updated to reflect fixes in scales and constructed variables.
- c. **M1_M2_Milwaukee_SampleDescriptions_20201103.** Updated to fix small errors in N and to reflect the changes in family member status among a small number of twin pairs.

- d. **M2_P1_DocumentationOfWeights_20201103.** Describes the construction of most recent post-stratification weights for the M2 RDD sample and compares weighted and unweighted sociodemographic variables to the Current Population Survey.
- e. **M2_P1_DocumentationOfIndustryAndOccupation_20201103.** Updated format for document describing coding of industry and occupation variables.
- f. **M2_P1_FieldReport_20201103.** Updated format of document for M2 field report.
- g. **M2_P1_DocumentationOfCodedTextData_20201103.** Describes the process for coding the open-ended and other specify text responses obtained through the Phone interview and SAQ. The resultant variables are available in the Coded Text Dataset.
- h. **M2_P1_DispositionCodeDocumentation_20201103.** Documentation that accompanies and summarizes final survey disposition codes for the M2 Phone interview.
- i. **M2_P1_SAQxPhone_README_20201103.** Describes SAQxPhone project and resulting dataset.
- j. **MIDUS_NamingAndCodingConventions_20191209.** Describes current file and variable naming and coding conventions for MIDUS datasets.
- k. **Variable crosswalk and harmonization.** A variable crosswalk of all the MIDUS survey variables (as well as all other MIDUS Project variables) is available online in the MIDUS Portal (<http://midus.colectica.org/>). The portal contains detailed searchable metadata for all MIDUS variables (including all waves and projects) and provides harmonization information for reconciling differences that might exist across versions of a variable. Customized datasets can also be created and downloaded that merge variables from across multiple MIDUS Project datasets and MIDUS samples (Core, Refresher, MKE, MIDJA, etc.)

E. ID Systems

Respondent ID. A respondent identification system exists in the MIDUS data for the Core sample via a variable called M2ID. This ID system is used throughout MIDUS Core sample datasets.

Family ID. A family identification system exists in the MIDUS data exists in the MIDUS data for the Core sample via a variable called M2FAMNUM. Every respondent has a family number. Related respondents within the sub-samples (Main-Siblings and Twins) have the same family number.

Sample ID. The MIDUS Core National sample is comprised of four sub-samples of respondents: Main RDD, Sibling, and Twin samples, plus a City Oversample. The same aggregation of

samples exists in all MIDUS Core National sample datasets and the variable SAMPLMAJ identifies which of the four subsamples a respondent belongs to.

Please report any errors or inconsistencies you find in the data or documentation to
midus_help@aging.wisc.edu
