

Panel Study of Income Dynamics, Child Development Supplement 2020: User Guide

May 2022

This user guide was prepared by Narayan Sastry, Paula Fomby, and Yi-Miau Tsai

Abstract

The 2020 Child Development Supplement (CDS-2020) to the Panel Study of Income Dynamics (PSID) collected data for a nationally representative sample of children in the United States on their health and well-being within the family context. CDS-2020 conducted follow-up interviews in the Fall of 2020, during the Covid-19 pandemic, about children who participated in the 2019 wave of CDS (CDS-2019) with data collection largely completed before the onset of the pandemic. The CDS-2020 data thus provide a picture of children and family well-being in the midst of the pandemic among a nationally representative sample with pre-pandemic data. The CDS-2020 sample covered all children aged 2–18 years in families from CDS-2019 that had not completed a home visit before March 2020 when these visits were ended due to the Covid-19 pandemic. Approximately two-thirds of the CDS-2019 sample were eligible for CDS-2020. CDS-2020 included re-interviews with many children aged 6–18 years who participated in the first wave of the new, ongoing CDS in 2014. (Children from the original CDS, conducted between 1997 and 2007, had all reached adulthood by the time of the CDS relaunch in 2014.) The CDS-2020 interview content builds on the strengths of CDS-2019 and PSID, a genealogical study of U.S. families that began in 1968. Information was collected in CDS-2020 on family Covid-19 incidence and vaccination status, financial effects of the pandemic, food security, mental health of children and their primary caregivers, and children’s school attendance and activities. Many measures in CDS-2020 were comparable to those collected in CDS-2019, facilitating analyses of the effects of the pandemic on children and families. All CDS data are publicly available free of charge through the PSID Online Data Center (www.psidonline.org) and the CDS Online Data Center (www.cds-tas.org). This User Guide provides essential information to researchers planning or undertaking research using the CDS-2020 data.

Suggested citation of the CDS-2020 User Guide:

“Panel Study of Income Dynamics, Child Development Supplement 2020: User Guide,”
Institute for Social Research, University of Michigan, 2022.

Suggested citation of the CDS-2020 data:

Child Development Supplement to the Panel Study of Income Dynamics, public use
dataset [restricted use data, if appropriate]. Produced and distributed by the Survey
Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI
(year data were downloaded).

Suggested acknowledgement of the CDS-2020 data:

The collection of data used in this study was supported by the *Eunice Kennedy Shriver*
National Institute of Child Health and Human Development under grant number R01
HD052646.

Preface

The 2020 Child Development Supplement (CDS) to the Panel Study of Income Dynamics (PSID) was funded by Grant R01 HD052646 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD). Additional funding was provided by the U.S. Department of Agriculture's Food and Nutrition Service. We are very grateful to these sponsors for their support of CDS-2020.

Contents

Abstract.....	ii
Preface.....	iv
Contents.....	v
Acknowledgements.....	vi
1. INTRODUCTION.....	1
<i>Background of CDS and PSID.....</i>	<i>2</i>
<i>CDS-2019 Overview.....</i>	<i>3</i>
<i>CDS-2020 Overview.....</i>	<i>4</i>
<i>Overview of the CDS-2020 User Guide.....</i>	<i>4</i>
2. THE CDS-2019 QUESTIONNAIRE, MEASURES, AND VARIABLES.....	5
<i>General Principles for the CDS-2020 Questionnaire.....</i>	<i>5</i>
<i>Questionnaire Content Domains.....</i>	<i>6</i>
<i>Variable Naming Conventions.....</i>	<i>6</i>
<i>Description of Questionnaire Sections.....</i>	<i>7</i>
Primary Caregiver Household Items.....	7
Primary Caregiver Child Items.....	8
3. THE CDS-2020 SAMPLE.....	10
4. THE CDS-2020 DATA FILE STRUCTURE.....	12
5. THE CDS-2020 WEIGHTS.....	20
<i>Cross-Sectional Child Weight.....</i>	<i>20</i>
<i>Method to Construct Child Weight.....</i>	<i>21</i>
<i>Method to Construct PCG Weight.....</i>	<i>23</i>
<i>Child Longitudinal Weight.....</i>	<i>23</i>
<i>Summary of Weights.....</i>	<i>23</i>
<i>Recommendations for Using the Weights.....</i>	<i>23</i>
Appendix.....	25

Acknowledgements

The 2020 wave of the Child Development Supplement (CDS-2020) to the Panel Study of Income Dynamics (PSID) was based at the Survey Research Center (SRC) in the Institute for Social Research at the University of Michigan.

CDS-2020 was directed by Narayan Sastry, with Paula Fomby serving as the associate director. Yi-Miau Tsai was the CDS-2020 project manager. PSID investigators who provide input at key stages of the project include Katherine McGonagle and David Johnson.

Development of the questionnaires for CDS-2020, which were based on those used in prior waves of CDS, was led by Narayan Sastry, Paula Fomby, and Yi-Miau Tsai. The questionnaire development process was ably managed and implemented by Mary Dascola and Rose McAloon-Fernando, with additional input from Allison Mageli and Rachel Carter.

The design and implementation of the CDS-2020 weights was undertaken by Wen Chang, Raphael Nishimura, and James Wagner.

Within the Survey Research Operations (SRO) unit, which undertook the fieldwork for CDS-2020, Rachel Orlowski served as the SRO Project Manager under the direction of Shonda Kruger-Ndiaye (PSID Survey Director) and Stephanie Chardoul (PSID Senior Project Advisor and SRO Director). Camila Kendall helped manage the project, including taking lead responsibility for the project at SRO for several months. Among the many other key SRO staff members who contributed significantly to CDS-2020 were Maryam Buagelila, Dianne Casey, Minako Edgar, Megan Gomez-Mesquita, Brad Goodwin, and Youhong Liu.

PSID staff members Flannery Campbell and Kimberly Schlegel were responsible for data processing activities.

Special thanks go to other members of PSID who contributed to CDS-2020, including April Beale, Vivian Burgett, Noura Insolera, Dennis Kloska, Rhonda Moats, Mohammad Mushtaq, Beth Simmert, and Kathryn Snider.

Finally, we are very grateful to Regina Bures at the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development for her guidance, advice, and support of the CDS project.

1. INTRODUCTION

The Child Development Supplement (CDS) to the Panel Study of Income Dynamics (PSID) was designed to support research on the health, development, and well-being of children within their family and neighborhood context. By 2020, four waves of CDS had been completed in the 23 years since the study was launched in 1997 and a fifth wave—CDS-2019—was underway. CDS-2019 sought to collect information on all PSID children aged 0–17 years. Data collection had begun on 1 October 2019 and was scheduled to continue through mid-2020. As in previous waves of CDS, the 2019 wave included a major home visit data collection component to collect information that could not be obtained reliably by telephone, the mode used to collect information from children’s primary caregivers (PCGs) and from older children themselves. The home visit components included reading and math assessments for children (and reading assessments for PCGs), time diaries for children, interviews with children aged 8–11 years, anthropometric measurements, and collection of saliva samples for subsequent genetic sequencing for children, primary caregivers, and other adults in the household.

On 14 March 2020, due to the Covid-19 pandemic, CDS-2019 halted all home visits.¹ At that time, telephone interviews had been completed with about three-quarters of primary caregivers and adolescents aged 12–17 years. Home visits had been completed with approximately 900 families. Because of the ongoing pandemic, there was no opportunity to restart the home visits. Instead, only telephone interviewing of primary caregivers and adolescents was continued, through May 2020 when all CDS-2019 fieldwork ended.

To replace the home visits from CDS-2019 that could not be undertaken due to the Covid-19 pandemic, the project instead planned and implemented a follow-up effort for the fall of 2020 that is called CDS-2020. The goal of CDS-2020 was to complete the collection of most, but not all, items originally included in the CDS-2019 home visit through a telephone interview and a mail-out/mail-back protocol. The targeted items for remote collection included weekday and weekend time diaries for children, saliva samples from children, PCGs, and other adults for subsequent genetic analysis, anthropometric measurements, and record linkage consent forms. In addition, we designed a short new Covid-19 telephone questionnaire module for PCGs. This module collected information about the disease incidence of Covid-19 among family members, financial effects of the pandemic, and the consequences for child and family well-being—including food insecurity, mental health, summer activities, and home schooling. Excluded from CDS-2020 were the assessments of reading and math skills and interviews with children aged 8–11 years. These items were excluded because it was infeasible to collect them through a telephone or mail-out/mail-back protocol. Fieldwork for CDS-2020 began on 18 September 2020 and ended on 31 December 2020.

This user guide provides information about the CDS-2020 study design, questionnaire instrument and measures, fieldwork outcomes, data structure and relationship with Core PSID and other components of PSID, and the data structure. A separate data release, box-and-arrows questionnaire, and user guide are available for CDS-2019.

In this chapter, we provide background on CDS and PSID, an overview of CDS-2020, information about CDS-2020, and an outline of this user guide.

¹ See Sastry, N., McGonagle, K. and Fomby, P. (2020) Effects of the Covid-19 crisis on survey fieldwork: Experience and lessons from two major supplements to the US Panel Study of Income Dynamics. *Survey Research Methods*, 14(2), 241–245.

Background of CDS and PSID

CDS collects data on psychological and social well-being, health status and behavior, family environment, education, child care, time use, sibling relationships, caregiver social and psychological resources, non-coresident parents, future work and schooling expectations, and religiosity. CDS data support studies of health, development, and well-being in childhood; the relationship between children's characteristics and contemporaneous family decisionmaking and behavior; and the effects of childhood factors on subsequent social, demographic, economic, and health outcomes over the entire life course for these individuals as they are followed into the future as part of the ongoing Core PSID.

CDS provides rich, comprehensive, and up-to-date panel data on a large, nationally representative sample of children in the United States that includes an over-sample of African American children and a representative sample of immigrant children. Public use data are available free of charge through the CDS Online Data Center (www.cds-tas.org) and the PSID Online Data Center (www.psidonline.org), which provide customized extracts and codebooks using a detailed index of variables. Sensitive information from CDS adolescent interviews are available through a special application procedure that requires a brief research plan and documentation of IRB review and approval (<https://simba.isr.umich.edu/restricted/ProcessReqSensData.aspx>). Restricted data, which include school identifiers and geocoded data about residential locations, are available to researchers through a data contract. Visit the PSID website for more information on obtaining access to restricted data from CDS (<http://simba.isr.umich.edu/restricted/RestrictedUse.aspx>).

There are several unique features of CDS-2020 that offer many important research opportunities to analysts. First, CDS-2020 is one of the few studies of children in the U.S. that has nationally representative data on children during the Fall 2020 phase of the Covid-19 pandemic—along with pre-pandemic information for these same children from CDS-2019. The unique data from CDS-2019 and CDS-2020 provide researchers with an important opportunity to examine the effects of the pandemic on children's well-being within the family context. Second, because the CDS children's parents are also participants in PSID, there is an enormous amount of data available from previous waves of Core PSID on many aspects of their lives. Information is available in CDS-2020 on siblings and cousins, providing unique research opportunities to control for family background and examine within-family differences in outcomes. Third, as CDS-2020 children move into adulthood, they will be interviewed in the PSID Transition into Adulthood Supplement in 2021 and beyond, and will also become primary PSID respondents. The information collected in CDS-2020 will provide invaluable insights into the effects of Covid-19 pandemic experiences and childhood circumstances on later adult development and on adult social, demographic, economic, and health outcomes.

CDS is part of the Panel Study of Income Dynamics, a longitudinal survey of a nationally representative sample of U.S. families that began in 1968.² The original 1968 PSID sample came from two sources: a nationally representative sample of approximately 3,000 families designed by the Survey Research Center at the University of Michigan (the "SRC sample") and an over-sample of approximately 2,000 low-income families from the Survey of Economic Opportunity (the "SEO sample"). PSID interviewed individuals from families in these two samples every year from 1968 to 1996 and biennially thereafter—whether or not they were

² McGonagle, K., Schoeni, R., Sastry, N., and Freedman, V. (2012). The Panel Study of Income Dynamics: Overview, Recent Innovations, and Potential for Life Course Research. *Longitudinal and Life Course Studies*, 3, 268–284.

living together in the same dwelling. In 1997, because of the escalation in costs driven by the doubling of the sample size during its 30-year history, PSID was forced to drop some families from the study. The cuts were made from the SEO sample. In 1997 and again in 2017, representative samples of new immigrants to the U.S. were added to PSID.

The original CDS began in 1997 with a cohort of 3,563 children from 2,394 families. The cohort included up to two randomly selected children aged 0–12 years in each family. Interviews were conducted with the children’s primary caregivers (PCGs; usually the children’s mother). Eligible CDS participants in 1997 were descended from the original 1968 PSID sample or the 1997 PSID immigrant refresher sample. In most cases, this means that the child’s father or mother was the child or grandchild of an original PSID respondent. In 2002, CDS families who participated in the 2001 Core PSID were contacted for a second round of data collection. CDS-2002 successfully re-interviewed 2,019 families (91 percent) who provided data on 2,907 children and adolescents aged 5–18 years. During 2007, 1,506 children aged 10–19 years were successfully re-interviewed (90 percent) in the third and final wave of the original CDS cohort study.

By 2014, all children in the original 1997 CDS cohort had reached adulthood, and a new generation of children had replaced them in PSID families. CDS-2014 sought to collect information on all PSID children aged 0–17 years in this new generation. The CDS-2014 sample included all PSID families that completed a Core PSID interview in 2013 and had one or more resident children. All eligible PSID children in each family were selected for CDS-2014, in contrast to the limit of two children per family in the original CDS. CDS-2014 participants formed a nationally-representative sample of children descended from the original 1968 families and the 1997 new immigrant refresher sample. (The CDS-2014 sample did not cover children from families in which both parents are post-1997 immigrants to the U.S.) CDS-2014 was primarily a telephone interview; however, a random 50 percent of households were selected to receive a home visit to collect information that could not be obtained reliably by telephone, including reading and math assessments for children (and reading assessments for PCGs), time diaries for a random weekday and a random weekend day, and interviews with children aged 8–11 years. The home visits facilitated the collection of other study components that were otherwise collected using a mail-out/mail-back protocol, including saliva samples for subsequent genotyping and anthropometric measurements.

CDS-2019 Overview

CDS-2019 continued the shift in orientation for the collection of information on PSID children from a study of a single cohort (from the original CDS) to a study that obtained information on the childhood experiences of all children in PSID families at regular intervals. The CDS-2019 sample included all age-eligible children from CDS-2014 (i.e., those aged 5–17 years in 2019) whose families participated in the 2019 wave of PSID, whether or not they participated in CDS-2014, and also added newly age-eligible children (i.e., those aged 0–4 years in 2019).

The CDS-2019 interview design and content drew heavily on previous waves of CDS. The study focused initially on completing telephone interviews with primary caregivers and older adolescents aged 12–17 years. After the telephone interviews were completed, all families were eligible for a home visit that included reading and math assessments for children, time diaries, collection of saliva samples, and other components.

A team of approximately 100 interviewers located around the country was recruited and trained for CDS-2019 and data collection began on 1 October 2019. Fieldwork began with

approximately 1,600 PSID families that had completed their 2019 Core PSID interview earlier in the year. Additional cases were released to the field for CDS-2019 as they completed their Core PSID interviews. Home visits to collect the in-person components of CDS-2019 began with CDS cases that completed their telephone interviews and were located near to interviewers or in large clusters if travel by interviewers was required. Following the completion of fieldwork for Core PSID on 31 December 2019, the final set of cases for CDS-2019 were released to the field in early February 2020.

The home visits were halted in mid-March 2020 as a result of the Covid-19 pandemic. Telephone interviewing of primary caregivers and adolescents continued through May 2020. In CDS-2019, information was collected on a total of 4,629 children from an eligible sample of 6,435 children. The overall, unconditional response rate at the child level for CDS-2019 was $4,629 / 6,435 = 72$ percent.

CDS-2020 Overview

CDS-2020 focused on the sample members from CDS-2019 who had not completed their home visit by 14 March 2020 when all such visits ended due to the Covid-19 pandemic. Study components included a telephone interview with PCGs to collect information on child well-being and family circumstances during the Fall 2020 phase of the pandemic. Other components included weekday and weekend time diaries for children, saliva samples, anthropometric measurements, and record linkage consent forms.

Development work for CDS-2020 was conducted in the Spring and Summer of 2020. A total of 23 interviewers were recruited and trained to work on CDS-2020, including several bilingual interviewers to conduct interviews in Spanish. Fieldwork began on 18 September 2020 and ended on 31 December 2020. Reminder calling for the mail-out/mail-back components continued through the end of January 2021.

CDS-2020 targeted 1,675 PCGs for telephone interviews, of which 1,153 were completed for a response rate of 68.9 percent. The mail-out/mail-back response rates were low, which is due to many factors including the lack of in-person interaction with interviewers, problems with mail delivery, and the Covid-19 pandemic circumstances. The low response rate for the time diaries from CDS-2020 presented a challenge for data release. The 489 completed time diaries from CDS-2020 cover a distinct period from the time diaries collected as part of CDS-2019 and hence we strongly recommend against pooling the two sets of time diaries. At the same time, the completed time diaries represent just $(489 / 2,629 =)$ 10.6 percent of the complete sample that represents the basis of the CDS-2020 weights—with accompanying concerns regarding coverage of all sample segments. We are planning to prepare an unofficial weight for the CDS-2020 time diaries that can be used by analysts who understand and accept its potential limitations and shortcomings but are nevertheless interested in analyzing the CDS-2020 time diary data.

Overview of the CDS-2020 User Guide

In Chapter 2, we provide a brief description of the CDS-2020 questionnaire instrument content. In Chapter 3, we provide an outline of the CDS-2020 sample. In Chapter 4 we describe the CDS-2020 data file structure and the procedures for merging files. Finally, in Chapter 5 we describe the construction and use of the CDS-2020 weights.

2. THE CDS-2019 QUESTIONNAIRE, MEASURES, AND VARIABLES

In this chapter, we provide a brief overview of the CDS-2020 questionnaire, which comprised a single instrument administered to one primary caregiver (PCG) per family. The CDS-2020 questionnaire focused on the Covid-19 pandemic, including disease incidence among family members and effects of the pandemic on family finances, children's schooling, and mental health. Several key scales were repeated from CDS-2019, providing opportunities to examine the effects of the pandemic on child and family well-being.

See the CDS Cross-Wave Variable Index for a comprehensive list of questionnaire items and scales available in CDS-2020 and prior waves. This index is available in the documentation section of the PSID website (<https://psidonline.isr.umich.edu/Guide/documents.aspx>) and the CDS website (<https://psidonline.isr.umich.edu/CDS/Guide/Documents.aspx>).

General Principles for the CDS-2020 Questionnaire

- The 2020 PSID Child Development Supplement (CDS-2020) was designed to support research on children's health and well-being during the Covid-19 pandemic within family, neighborhood, and school context.
- The CDS-2020 questionnaire was designed to provide directly comparable measures to CDS-2019 and earlier waves of CDS in several key content areas including food security, mental health, and behavior problems.
- The CDS-2020 questionnaire collected period-specific measures of children's activities for Spring 2020 (between March and May 2020) and for Summer 2020 (between June and August 2020). For school-aged children, information was also collected on school closure and attendance for the 2019–2020 school year and the 2020–2021 school year.
- Changes to the CDS-2020 questionnaire were implemented where necessary. Key changes included:
 - A transition in CDS from using the Behavior Problems Index (BPI) to the Strengths and Difficulties Questionnaire (SDQ) for assessing children's personality and behavior was begun in CDS-2019. In each family in CDS-2019, one age-eligible child (aged 3 years or older) was randomly selected for the SDQ and the BPI to facilitate a shift between these two modules. In CDS-2020, only the SDQ was administered. In each family, the child who was randomly selected for the SDQ in CDS-2019 had the module readministered. The SDQ was also collected for up to two additional children in each family, with the children randomly selected if there were three or more additional children. Four SDQ scales were administered in CDS-2020, with the SDQ prosocial scale omitted from the instrument because the pandemic affected children's social interactions differentially in uncertain but potentially significant ways.
 - The family food security module administered in CDS-2020 was based on a 30-day reporting period, in contrast to the six month reporting period used in previous CDS waves.
- Minimizing respondent burden by limiting the average length of the CDS-2020 interview to approximately 20 minutes.

Questionnaire Content Domains

Table 2.1 summarizes the questionnaire content domains in CDS-2020 and their comparability to CDS-2019.

Table 2.1. CDS-2019 Content Domains

Content Domain	Description of content	Questionnaire Items	Comparability to CDS-2019
Covid-19 health	Covid-19 health series for PCG, children and other adults	CVH1–CVH18	New in CDS-2020
Covid-19 finances & employment	Effects of Covid-19 on family finances and employment	CVF1–CVF17	New in CDS-2020
Food security	Thirty-day family food security module for families with children	CVR1–CVR15	Same items as in CDS-2019; change in reporting period
Caregiver health and psychological resources	General health, Kessler K6 30-day psychological distress scale	CVN1–CVN9	Same items as in CDS-2019
Health	General health of child	CVA2	Same question as in CDS-2019
Child psychological & social well-being	Strengths and Difficulties Questionnaire (SDQ)	CVB1A–CVB1Z	Four SDQ subscales (prosocial scale omitted); same items as in CDS-2019; collected for up to three children per family
Education & employment	School closure, remote and online school, attendance, lessons, summer activities	CVQ1–CVQ26	CVQ10 is comparable to CDS-2019 (CC1) in Childcare block; CVQ24 is the same CDS-2019 school question CHGRADE

Variable Naming Conventions

Items from the CDS-2020 questionnaire have a general prefix “CV” and a section abbreviation that indicates the questionnaire section. For example, CVH1 is the first item in Section H (covering health) of the CDS-2020 questionnaire. The questionnaire can be downloaded here: <https://psidonline.isr.umich.edu/cds/questionnaires/cds-20/pcg.pdf>.

The data associated with these questionnaire items appear as variables in the PSID Online Data Center (www.psidonline.org) and the CDS Online Data Center (www.cds-tas.org). Variables appear in the 2020 Covid-19 Health File, 2020 Primary Caregiver Household File, and 2020 Primary Caregiver Child File using the following structure:

- The leading character(s) refers to the data file for which the questionnaire item is located:
 - V = Health File
 - H = PCG Household Interview
 - P = PCG Child Interview
 - D = Time Diary Questionnaire
 - WD/WE = Time Diary Aggregated Activity File
 - COLA–COLJ = Time Diary Activity File (disaggregated)
- The following two characters in the variable refer to the calendar year that data collection began. For variable names associated with CDS-2020, these characters are always “20.” This scheme has been used since CDS-2014, but was not used consistently in earlier waves of CDS.

3. The remaining characters in the variable name refer to the location of the item in the questionnaire.
4. Generated variables (i.e., constructed scale scores, interview information like calendar dates, and other variables produced by PSID staff) adopt naming conventions (1) and (2). For these variables, the remaining characters typically use a mnemonic device to help users identify the variable's content. For example, the generated variables constructed from the household food security items are: H20CVHHFD for household food security category. The list of generated variables is available on the CDS generated variables and scales tab in the CDS Cross-Wave Variable Index <https://psid.isr.umich.edu/cds/CDS-cross-wave-variable-index.xlsx>.

Description of Questionnaire Sections

Primary Caregiver Household Items

The primary caregiver household (PCG-HH) items focus on characteristics of a child's family and household and of the PCG.

PCG Covid-19 Health Series (Section CVH). The PCG Covid-19 Health Series collected information on diagnoses, tests, and treatment related to Covid-19. All PCGs were asked whether anyone in the household had Covid-19 (CVH1) and a series of questions on whether the PCGs themselves, CDS children, and/or adults in the household were ever tested for Covid-19 (CVH2, CVH8MO–CVH10), diagnosed with Covid-19 (CVH3–CVH5YR), and/or admitted to a hospital for Covid-19 (CVH11). Information was collected for every household member for the series and appears in the CVH2020 data file. The file includes a type variable (V20CVHTYPE) to indicate whether the information pertains to a child, a PCG, or another adult in the household.

Covid-19 Financial Series (Section CVF). Information on family financial effects of the Covid-19 pandemic was collected through a set of 17 questions (CVF1–CVF17) that covered experiences and management of job and income changes (CVF1–CVF4) and financial difficulties (CVF5–CVF17). The information applies to the household unit and thus is released at the PCG-Household level (H20CVF1–H20CVF17).

PCG Psychological Distress (Section CVN). The Kessler 6 (K-6) Non-Specific Psychological Distress Scale (CVN1–CVN6) was designed to discriminate cases of serious mental illness from non-cases in a general population survey. The K-6 is administered to respondents is also included in the National Health Interview Survey and the National Household Survey on Drug Abuse, as well as in prior waves of CDS and in Core PSID.

The K-6 includes six items about how the respondent felt during the past 30 days. Response items are based on a scale from 1 to 5, where 1 indicates "all of the time" and 5 indicates "none of the time." Individual items may be rescored to range from 0 to 4 and then summed to calculate a total score that is comparable to other studies. A summed score of 13 or higher indicates a potential for nonspecific distress. The generated distress scale H20K6_14 is a sum score computed for all cases with valid responses to all six items in the scale.

The scale includes three follow-up items about persistence and impairment associated with symptoms of nonspecific distress (CVN7–CVN9). These items are administered to respondents who endorse any of the items in the K-6 series. Responses to these additional items are not required in order to score the K-6.

In CDS-2020, besides psychological distress, PCGs self-reported their general health status on a scale from 1 to 5, where 1 indicates “excellent” and 5 indicates “poor” (CVN0).

Food Security (Section CVR). The PCG-HH Interview included an 18-item version of the U.S. Household Food Security Survey Module developed by the Economic Research Service at the U.S. Department of Agriculture (CVR1–CVR15). The module includes questions about various levels of food security such as worries about having enough food and enough healthy food, cutting back to conserve food, and running out of money for food. The module collects separate information about household food security (CVR1–CVR8) and child food security (CVR9–CVR15). These data allow the food security status of CDS-2020 families to be defined along a continuum extending from high food security to very low food security. Generated variables associated with this series include raw scores summing the number of endorsed items pertaining to the household overall (H20HHFOODR) and separately for adults (H20ADFOODR) and children (H20CHFOODR). A parallel set of items describes the food insecurity status of the household overall (H19HHFOOD) and of adults (H19ADFOOD) and children (H19CHFOOD). A raw sum score (H20FOOD6R) and a food insecurity status indicator (H20FOOD6) based on a six-item subset of questionnaire items (CVR2–CVR6) are also available.

The set of questions are identical to CDS-2019, however, the timeframe is different. CDS-2020 asks each statement “in the last 30 days”, instead of “in the past six months” reference period used in prior CDS waves.

Primary Caregiver Child Items

In the rest of the instrument, the questions were asked to the PCG for each of the CDS children based on the age and/or school grade range of the items. See the questionnaire for age and/or school grade ranges for each item and for rules governing skip patterns throughout the instrument.

Child Health (Section CVA). CDS-2020 included one question about the general health of each child (CVA2). This health question is drawn from the National Health Interview Survey and has appeared in all prior waves of CDS.

Strengths and Difficulties Questionnaire (Section CVB). Twenty items from the Strengths and Difficulties Questionnaire (SDQ) were included in CDS-2020 for children aged 3–17 years. In CDS-2019, PCGs responded to the questionnaire for one randomly selected child in each family. In CDS-2020, PCGs responded for the same child and up to two additional children per family (selected at random in families with more than two additional children).

Two programming issues affected SDQ administration in CDS-2020. First, children age 0–2 years were age-ineligible but were inadvertently included in the randomization to select two additional target children per family. The SDQ was administered for these children but item responses have been removed and no calculated scores were created for them. Second, PCGs responded to items CVB1S1 and CVB1S2 for all selected children despite age rules for the two questions. Responses for age-ineligible children on each item have been removed.

The SDQ administration in CDS-2020 excluded five items about children’s prosocial behavior but was otherwise identical to CDS-2019. The CDS-2020 administration included items assessing conduct problems (CVB1E, CVB1G, CVB1M, CVB1S1/CVB1S2, and CVB1W1/CVB2W2), hyperactivity/inattention (CVB1B, CVB1K, CVB1P, CVB1V, and CVB1Z), emotional problems (CVB1C, CVB1H, CVB1N, CVB1Q, and CVB1Y), and peer relationship problems

(CVB1F, CVB1L, CVB1O, CVB1T, and CVB1X). Seven calculated scales are provided based of the SDQ items and scoring instruction (www.sdqscore.org). These includes SDQ Conduct Scale (P20CVSDQC), SDQ Hyperactivity Scale (P20CVSDQH), SDQ Emotional Scale (P20CVSDQEM), SDQ Peer Relationships Scale (P20CVSDQP), as well as SDQ Externalizing Score (P20CVSDQE), SDQ Internalizing Score (P20CVSDQN) and SDQ Total Difficulties Score (P20CVSDQ).

School Year 2019–2020: School Closure and Attendance. The section collected retrospective information on school attendance for the 2019–2020 school year, focusing on remote schooling in the period between the closing of schools due to the Covid-19 pandemic and the end of the school year. Several types of adaptations to children’s learning environments were covered, such as remote modes of schooling, instruction, and homework assignment (CV1–CV3, CV5), as well as student attendance (CV4, CV6) and parental involvement in school activities (CV7, CV8). Finally, PCGs reported on the learning quality for children during this period (CV9).

Spring 2020: Activities. A series of six questions (CV10–CV15) asked directly about the impact of the Covid-19 pandemic on children’s contact with friends and family (CV11–CV12) and use of TV, video games, and social media (CV13–CV15). One item asked whether people outside of the household provided childcare for the child (CV10). And one item covered perceived changes in family relationships during this period (CV16).

Summer 2020: Activities. A series of seven questions (CV17–CV23) asked about the period between June and August 2020, including childcare arrangement (CV17), academic activities, sports and lessons, and trips and camps (CV18–CV22). Information on employment was collected for children aged 12 years and older (CV23).

School Year 2020–2021: Attendance. The section collected information on the current (2020–2021) school year, including whether the child was attending school in-person, online, or a hybrid of the two (CV24–CV26).

3. THE CDS-2020 SAMPLE

The CDS-2020 sample was designed as a follow-up to CDS-2019 in order to collect measures for children who did not participate in the in-home components of CDS-2019 prior to the halting of all CDS-2019 home visits on 14 March 2020 due to the Covid-19 pandemic.

The CDS-2019 sample—and the PSID sample more generally—was designed to be representative of the corresponding U.S. population of children and families. By design, PSID and CDS-2019 have certain gaps in coverage. The CDS-2019 sample eligibility criteria were defined as follows:

- Family participated in the 2019 Core PSID survey.
- Child’s reported birth year was 2002–2018.
- Child was classified as belonging to the PSID sample (i.e., has the “PSID gene”).
- Child was not classified as a household head or the spouse/partner of a household head.

The CDS-2019 User Guide describes the eligible sample and the fieldwork outcomes. A total of 4,629 children participated in CDS-2019.

The CDS-2020 sample eligibility criteria were defined as follows:

- Child participated in CDS-2019 through completion of the PCG-Child Interview or the Child Interview.
- Child did not complete in-home components for CDS-2019.

Of the 4,629 children who participated in CDS-2019, 1,685 completed their in-home components rendering them ineligible for CDS-2020. The final eligible CDS-2020 sample comprised (4,629 – 1,685 =) 2,944 children. Interviews were completed in CDS-2020 for 1,997 children, yielding a child-level response rate of (1,997 / 2,944 =) 67.8 percent.

In Table 3.1 we present information on the family-level fieldwork outcomes for CDS-2020 based on the disposition of the PCG interview. A total of 1,675 PCGs were eligible and contacted for a CDS-2020 interview. One PCG was deceased, leaving 1,674 eligible PCGs. Among these 1,674 eligible PCGs, interviews were completed with 1,153 PCGs, corresponding to a PCG response rate of (1,153 / 1,674 =) 68.9 percent. Among the 521 total non-respondent PCGs, 123 refused to participate (7.3 percent) and 398 could not be reached after multiple contact attempts (23.8 percent).

Table 3.1. CDS-2020 PCG-Level Fieldwork Outcomes

CDS-2020 Outcome	Count
Interview completed	1,153
Field period ended – respondent not reached	398
Refusal	123
Non-sample – respondent deceased	1
Total	1,675

Children in CDS-2020 ranged in age from 2 to 18 years, as shown in Table 3.2, based on their year of birth and the dates of CDS-2020 fieldwork. Table 3.1 also shows that the CDS-2020 sample was divided approximately evenly between males and females.

Table 3.2. Age and Sex of Children in CDS-2020

Birth year	Males	Females	Total	Percent
2002	43	44	87	4.36
2003	66	51	117	5.86
2004	54	72	126	6.31
2005	61	54	115	5.76
2006	63	55	118	5.91
2007	51	65	116	5.81
2008	47	65	112	5.61
2009	64	60	124	6.21
2010	70	50	120	6.01
2011	56	49	105	5.26
2012	52	64	116	5.81
2013	63	55	118	5.91
2014	67	66	133	6.66
2015	66	61	127	6.36
2016	78	51	129	6.46
2017	61	61	122	6.11
2018	58	54	112	5.61
Total	1,020	977	1,997	100.00

4. THE CDS-2020 DATA FILE STRUCTURE

The CDS-2020 data package includes the following files:

1. 2020 Primary Caregiver Household Interview File (one record per interviewed primary caregiver, N=1,153)
2. 2020 Primary Caregiver Child Interview File (one record per child interview, N=1,997)
3. 2020 Time Diary Files
 - a. 2020 Time Diary Activity File (one record per activity spell, N=18,956 activity spells)
 - b. 2020 Time Diary Aggregated Activity File (one record per child with a completed time diary, N=489)
 - c. 2020 Time Diary Questionnaire File (one record per child with a completed time diary, N=489)
4. 2020 Covid-19 Health File (one record per household member listed in CDS-2019 roster, N=4,187)
5. 2019 Demographic File (one record per child, N=4,629)
6. 2019 Household Roster File (one record per household member listed in CDS-2019 roster, N=10,361)
7. Cumulative CDS ID Map File 1997–2019 (one unique record across all waves per CDS-selected child, primary caregiver, or other caregiver, N=19,644)

Table 4.1 summarizes these files according to the CDS-2020 individual for whom data are available and lists the number of records in each component/file.

Table 4.1. CDS-2020 Study Component Completion by Individual Sample Member Type

Individual	CDS-2020 File							
	DEMOG	PCG-CHILD	TD_QN	TD_AGG	TD_ACT	PCG-HH	CVH	HHROSTER
Child								
Age 0–2 years	X	X	X	X	X		X	X
Age 3–4 years	X	X	X	X	X		X	X
Age 5–7 years	X	X	X	X	X		X	X
Age 8–11 years	X	X	X	X	X		X	X
Age 12–17 years	X	X	X	X	X		X	X
PCG						X	X	X
Other HH members							X	X
Num. of records								
Total	4,629	1,997	489	489	18,956	1,153	4,187	10,361
With associated PCG/child record		1,997				1,139*		

*In 14 households, all children who were otherwise eligible for CDS-2020 had moved out of the family unit where they were observed in CDS-2019. There are no child-level data associated with these households in CDS-2020.

Primary Caregiver Household Interview (PCGHH2020)

The Primary Caregiver Household Interview is provided at the primary caregiver (PCG) level, with one record per interviewed PCG (N=1,153).

Note that there are some CDS-2020 households that completed the Primary Caregiver Household Interview but did not complete the Primary Caregiver Child Interview for any child (N=14 primary caregivers). In these families, all eligible children had moved out of the household after CDS-2019 and prior to the CDS-2020 interview. As a result, merged file content between the PCGHH2020 interview file and PCGCHILD2020 will not yield a file with complete data for all fields for all individuals.

Primary Caregiver Child Interview (PCGCHILD2020)

The Primary Caregiver Child Interview is provided at the child level (N=1,997). The data file includes records for all children aged 1–19 years (born 2002–2018) for whom a primary caregiver provided a Primary Caregiver Child interview.

Covid-19 Health Interview (CVH2020)

The Covid-19 Health Interview includes one record for each person who was rostered in a CDS-2019 household and still present in the household at the CDS-2020 interview (N=4,187). The variable V20CVHTYPE indicates whether the record pertains to a CDS child (1), primary caregiver (2), or other family unit member (3).

Time Diary

Questionnaire Administration File (TD_QN2020). The time diary questionnaire administration file is provided at the child level (one record per child with a completed time diary, N=489). The file describes time diary characteristics, including day, date, and mode of completion.

Activity File (TD_ACT2020). The time diary activity file is provided at the activity spell level (one record per activity spell within each daily diary for each child, N=18,956). The data file is sorted by child, diary day (weekday/weekend day), and clock time, beginning at 12:00 am on the diary day). Within each diary day, the rows describe the sequence and characteristics of a child's primary and secondary activities over 24 hours, with time measured in seconds. The complete list of children's activities and their corresponding codes are included in the CDS-2019 Time Diary Coding Manual (available on the PSID website, www.psidonline.org), which was used for coding the time diaries in CDS-2020 as well.

Aggregated Activity File (TD_AGG2020). The time diary aggregated activity file is provided at the child level (one record per child with a completed time diary, N=489). For each coded activity, the total time a child spent in that activity in a 24-hour day is provided for weekday and weekend reports separately.

Household Roster (HHROSTER2019)

The Household Roster File includes one record for each person residing in a CDS-2019 household at the time of the completed coverscreen interview (N=10,361).

The variable R19C20IND indicates whether an individual who was observed in CDS-2019 remained in the household at the time of the CDS-2020 interview.

The roster file includes three sets of unique identifiers for each person: a CDS-2019 family household identifier (R19CDHID) and roster position (R19INST); a PSID 2019 Core interview family unit identifier (R19YRID) and sequence number (R19CYPSN); and a time-invariant family lineage identifier (R19ID68) and person number (R19PN). These unique identifiers may be used to merge together individual-level content files within CDS-2020 or between CDS-2020 and CDS-2019 or other components of the PSID suite of studies, using the merging instructions provided below.

Demographic Data (DEMOG2019)

The Demographic Data File is provided at the child level (one record per child, N=4,629). Records are included for all children who have a record on either the Primary Caregiver Child Interview (PCGCHILD2019), the Child Interview (CHILD2019), or both. This file may be used to link children's records from CDS-2020 to their records in CDS-2019 and to obtain information about their demographic characteristics.

The Demographic Data File includes the following information:

- Eligibility for and participation in CDS-2020,
- Children's unique identifiers in the 2019 PSID Core interview,
- For each CDS-2019 study component, an indicator of whether a record for the child is included,
- CDS-2019 and CDS-2020 sampling weights,
- Primary caregiver relationship to child,
- Child characteristics at birth reported by each known birth or adoptive parent, including birthweight, race (up to three mentions), and Hispanic ethnicity,
- For each rostered household member, their age, sex, and relationship to the child (unique identifiers for those household members appear on the Household Roster file).

See the CDS-2019 User Guide, for more information about the Demographic Data File:

https://psidonline.isr.umich.edu/cds/CDS2019_UserGuide.pdf.

Cumulative ID Map (CDSIND2019)

The purpose of the Cumulative ID Map is to provide unique identifiers for CDS children and their caregivers that allow users to merge data files within CDS, to merge in family- and person-level information from other PSID study components including the Core (main) interview, and to map characteristics from one person to another (e.g., to attach caregivers' individual-level characteristics to their children's records).

The Cumulative ID Map File includes rows for all children selected to participate in CDS (regardless of whether they actually participated) and all designated primary or other caregivers since 1997 (N=19,644 as of CDS-2020). The file is in a wide format. Each selected CDS child or designated PCG or other caregiver (OCG, included in CDS in 1997, 2002, and 2007 only) occupies one row. A selected CDS child in one wave who becomes a designated primary caregiver to another CDS child in a later wave remains on the same row. Change in status from selected child to designated caregiver is indicated through the CDS record type variable described below.

The CDSIND2019 file includes the following sets of unique identifiers for each person:

- Time-invariant 1968 ID (CDSCUMID68) and person number (CDSCUMPN),
- Core (main) family interview ID (CRFID**) and sequence number (CRSN**) from the PSID Core interview wave immediately prior to a given wave of CDS, and
- CDS household ID (CDS_HID**) and sequence number (CDS_SN**) in a given CDS wave.

The asterisks (**) stand in for a two-year suffix at the end of each variable name denoting survey year.

For CDS children, the same sets of identifiers are provided for primary caregivers:

- For all waves: ID68PCG**, PNPCG**, CRPCGFID**, and CRPCGSN**, and
- For CDS-2014 onward: CDSPCGSN** (note that CDS_HID** is the same for caregivers and children).

For CDS children in the original CDS only (1997, 2002, and 2007), parallel identifiers are included for other caregivers:

- ID68OCG**, PNOCG**, CROCGFID**, CROCGSN**

For CDS-2014 and later waves, another set of variables describes the interview components associated with each primary caregiver or child:

- DEMOG_**, PCGCH_**, CHILD_** (not applicable for CDS-2020), PCGHH_** (primary caregiver-level).

NOTE: When working with records from CDS-2020, use identifiers from CDS-2019 on the Cumulative ID Map File (identifiers with the digits “19” appearing in the suffix of the variable name). There are no identifiers specific to CDS-2020 on the Cumulative ID Map File.

File Merging

Users may wish to combine information from multiple components of CDS-2020 into a single data file or to incorporate information about children and their families from CDS-2019 or the PSID Core (main) interview, earlier waves of CDS, or other PSID studies. The PSID Online Data Center (www.psidonline.org) and CDS Online Data Center (www.cds-tas.org) will deliver data extracts already merged together from multiple files for records pertaining to the same person. However, users wishing to combine information on multiple persons (e.g., primary caregivers and children) or who are using CDS-2020 packaged data will need to merge records across files using the following guidance.

Below we describe the unique identifiers and steps required to conduct data merges. See Table 4.2 for a complete list of unique identifiers pertaining to CDS-2020.

Merging Data Files Within CDS-2020

Merging Records from the PCG-Household and PCG-Child Interviews

The guidance below is directed to users working with CDS-2020 packaged data or who use the PSID Online Data Center (www.psidonline.org) to create a data extract containing variables from the CDS-2020 PCG-Household and PCG-Child Interviews. Users who create a data extract containing variables from both files in the CDS Online Data Center (www.cds-tas.org) may request to receive a data file on which these records are already merged at either the child or primary caregiver level. Currently this enhanced integration is only available in the CDS Online Data Center.

Table 4.2 Unique Identifiers in CDS-2020
(Note: When Two variables are Listed in a Cell, the Variables are Jointly Unique)

	Record unit (Ego)	2019 PSID Core interview (Family ID, person ID)		CDS-2020 (Family ID, person ID)			Fixed (time-invariant) (Family ID, person ID)	
		Ego	Caregiver to Ego	Ego	Caregiver to Ego	Caregiver number	Ego	Caregiver to Ego
DEMOG2019	Child	X19YRID, X19CYPSN						
PCGCH2020	Child	P20YRID, P20CYPSN						
TD_QN2020	Child	D20YRID, D20CYPSN						
TDAGG2020	Child	AGGID20, AGGSN20						
TDACT2020	Activity spell*	TDID20, TDSN20						
PCGHH2020	Primary caregiver	H20YRID, H20CYPSN		H20CDSHID, H20INST		H20PCGHH		
CVH2020	Household member			V20CDSHID, V20INST				
HHROSTER2019	Household member	R19YRID, R19CYPSN		R19CDSHID, R19INST	R19CDSHID, R19CDSHPIN	R19PCGHH	R19ID68, R19PN	
CDSIND2019	CDS sample (selected children and designated caregivers)	CRFID19, CRSN19	CRPCGFID19, CRPCGSN19	CDS_HID19, CDS_SN19	CDS_HID19, CDSPCGSN19	PCGHHNO19	CDSCUMID68, CDSCUMPN	ID68PCG19, PNPCG19

*File includes multiple activity spell records per child

To merge data between the Primary Caregiver Household Interview file and the Primary Caregiver Child Interview, use the unique CDS household interview number and PCG household number—which children and their associated primary caregiver have in common. (In households with more than one primary caregiver, the PCG household number indicates whether the caregiver associated with a CDS child is the first or second primary caregiver in the household.)

Merging child and PCG records is most straightforward when using the CDS Cumulative ID Map (CDS2019IND) as a bridge between files. The Cumulative ID Map can be downloaded from the PSID packaged data page (<https://simba.isr.umich.edu/Zips/ZipMain.aspx>), and variables from the Cumulative ID Map can be added to data carts via the PSID Online Data Center (www.psidonline.org) and the CDS Online Data Center (www.cds-tas.org).

Use the following steps:

1. Conduct a one-to-one merge between the Primary Caregiver Child Interview file and CDSIND2019 using the two child ID variables (CRFID19 and CRSN19 in CDSIND2019; [z]YRID and [z]CYPSN in the child file where [z] is the three-character file-identifier prefix) as the unique identifiers. Prior to merging, rename the ID variables as needed so that they will match. This will merge the CDS household interview ID (CDS_HID19), the PCG household number (PCGHHNO19) and PCG PSID family identifiers (CRPCGFID19 and CRPCGSN19) on CDSIND2019 to the child-level file.
2. Conduct a one-to-many merge between the Primary Caregiver Household Interview file (PCGHH2020) and CDSIND2019 using the CDS household interview number and PCG household number (CDS_HID19 and PCGHHNO19 in CDS2019IND and H20CDSHID and H20PCGHH in PCGHH2020) as the unique identifiers; prior to merging, rename these variables as needed so that they match between the two files. This will put the Primary Caregiver Household Interview data and PCG identifiers at the child level.
3. Conduct a one-to-one merge using the child identifiers CRFID19 and CRSN19 to merge the files created in Steps 1 and 2. Users may wish to remove records which are in only the Primary Caregiver Household Interview file (e.g., PCGHH2020 records that do not have a corresponding child record).

Merging Records from the Covid-19 Health File to Other CDS-2020 Files

The Covid-19 Health File includes one record for each family member who was present in a CDS child's household in CDS-2019 and CDS-2020. The variable V20CVHTYPE indicates whether each record pertains to a CDS child (1), primary caregiver (2), or other family unit member (3). The identifier variables on the file may be used to link each record to the CDS-2019 household roster file (HHROSTER19).

To merge records between the Covid-19 Health File and HHROSTER19 use the variables V20CDSHID and V20INST on the Covid-19 Health File and R19CDSHID and R19INST on HHROSTER19. Prior to merging, rename the variables as needed so that they will match.

The merged file will include the full set of identifiers listed in Table 4.2 to facilitate further merges to other files in CDS-2020, CDS-2019, or other PSID studies. See below for more information.

Merging CDS-2020 records to CDS-2019

CDS-2020 is an extension to CDS-2019, and the primary caregivers, children, and other family members in CDS-2020 are a subset of individuals and families included in CDS-2019. As a result, file merges for same-level records across the studies are straightforward.

To merge records between the PCG-Household Interview Files in CDS-2020 and CDS-2019, use the variables H20CDSHID and H20INST and H19CDSHID and H19INST respectively (renaming variable names as needed prior to merging so they will match across files).

To merge records between the PCG-Child Interview Files in CDS-2020 and CDS-2019, use the variables P20YRID and P20CYPSN and P20YRID and P20CYPSN respectively (renaming variable names as needed prior to merging so they will match across files).

To merge records between the PCG-Child Interview File in CDS-2020 and the Child Interview File in CDS-2019, use the variables P20YRID and P20CYPSN and C19YRID and C10CYPSN respectively.

Merging CDS-2020 records to other PSID studies

Merging Individuals' Records

Users may merge a CDS-2020 content file directly to records from the PSID 2019 Core interview by using the Core 2019 family interview number and sequence number included on all CDS-2020 files ([z]YRID and [z]CYPSN, where [z] is the three-character file identifier prefix). The equivalent variables in the PSID cross-year individual-level file are ER34701 (family interview number, equivalent to [z]YRID in CDS-2019) and ER34702 (sequence number, equivalent to [z]CYPSN).

Users who wish to merge to records from other PSID Core interview waves or other PSID studies should use the time-invariant 1968 ID and person number instead. These variables are included on the Household Roster (HHROSTER2019, variables R19ID68 and R19PN) and the Cumulative ID Map (CDSIND2019, variables CDSCUMID68 and CDSCUMPN). The Household Roster includes records for all CDS-2020 household members. The Cumulative ID Map includes records for children who were selected for CDS-2019/CDS-2020 and their designated primary caregivers.

Merging Child and Primary Caregiver Records

Users may wish to attach information about a primary caregiver that was collected in the PSID Core interview or elsewhere to a child's record. Use the time-invariant unique identifiers for the focal child (CDSCUMID68 and CDSCUMPN) and caregiver (ID68PCG19 and PNPCG19) on the Cumulative ID Map (CDSIND2019) for this purpose. The equivalent variables included on data extracts from the PSID Online Data Center or CDS Online Data Center are ER30001 (1968 family interview ID) and ER30002 (person number).

Merge the Cumulative ID Map to any other content file using the *primary caregiver's* unique identifiers. (First rename the unique identifiers for the primary caregiver as needed in order to

facilitate a merge between the two files.) This will attach the primary caregiver's characteristics from the external file to the child's record on the Cumulative ID Map.

Use a one-to-many merge approach because the same caregiver may appear on multiple children's records in the Cumulative ID Map file but will only appear once on their own record in data files associated with the PSID Core interview.

Note that only a subset of records will be matched. Some records will appear only on the Cumulative ID Map. This includes records for children who were not CDS-2019 participants; primary caregivers themselves; and children who were in CDS-2019 but whose primary caregiver has no record on the content file. Other records will appear only on the content file. This includes all individuals who were not the primary caregiver to a child in CDS-2019. Users may wish to remove these unmatched records.

An alternative to this approach is to request a data extract from the CDS Online Data Center (<https://www.cds-tas.org>). In the data cart, include at least one child-level variable from CDS-2020. In addition, select individual-level characteristics from the Curated PSID Variables that are of interest with regard to a child's primary caregiver such as age or years of educational attainment. At checkout, check the box for "Child to Primary Caregiver Integration." This will add unique identifiers for the primary caregiver as well as the primary caregiver's values on all of the variables included in the cart. These variable names will include suffixes that refer to the primary caregiver.

5. THE CDS-2020 WEIGHTS

CDS-2020 includes a cross-sectional child weight in the initial final release and will include a longitudinal child weight in a revised final release along with a cross-sectional PCG weight. The cross-sectional child weight is described below. The weights to be added will be described in a revision to this user guide. This chapter describes the construction and use of the CDS-2020 weights. We recommend that researchers use the provided weights with all analyses.

Cross-Sectional Child Weight

The Child Weight (X19CVWGT) allows researchers to generalize their statistical results to the U.S. national population of children aged 1–17 years in 2020. Note that the CDS-2020 sample was conditioned on participation in CDS-2019—and the weights are designed to project backwards to the full CDS-2019 sample. The CDS-2020 sample thus represents the experiences in the fall of 2020 of a national sample of children that was originally formed in 2019. By design, the CDS-2020 sample excluded all children from CDS-2019 who completed a home visit before these visits ended on 14 March 2020 due to the Covid-19 pandemic. The CDS-2020 weights account for these excluded cases by design, in order to represent the full CDS-2019 sample. A preliminary analysis was conducted to verify that the CDS-2020 was sufficiently large to represent the entire CDS-2019 sample one year later and that it was sufficiently diverse to adequately represent all demographic subgroups in the sample.

The Child Weight includes a base component derived from the 2019 Core PSID weight that accounts for differential sample selection probabilities in the PSID sample design and attrition in Core PSID. The CDS-2020 Child Weight was derived from the CDS-2019 Child Weight, which incorporates differential patterns of non-response and corrects potential under-coverage of some demographic subgroups using post-stratification. The CDS-2020 Child Weight also accounts for differential non-response in CDS-2020.

The CDS-2020 Child Weight (X19CVWGT) is provided for 1,997 children. X19CVWGT should be used when analyzing CDS-2020 outcomes and when comparing survey outcomes collected in CDS across both the 2019 and 2020 waves.

The CDS-2020 cross-sectional child weight should be used for all child-level analyses based on the interview data from CDS-2020 that are undertaken with one observation for each child in the sample. This weight should not be used for analysis of the time diary data collected in the fall of 2020. For analyses at the PCG level, data users can construct a PCG weight as the mean of child-level cross-sectional weight across all of the children for which each PCG in CDS-2020 was responsible. A constructed PCG Weight will be provided in a revised final release of the CDS-2020 data.

Table 5.1 summarizes for CDS-2020 the released Child Weight, the deliberate omission of a Time Diary Weight, and the PCG Weight that can easily be constructed from the Child Weight.

Table 5.1 Use of CDS-2020 Weights for Analyzing Interview Data

Analysis sample	Recommended weight	Cases
Child-level interview data	X19CVWGT	1,997
Time diary completed in Fall 2020	– None –	489
PCG or household interview data	Mean of X19CVWGT over children by PCG	1,153

Method to Construct Child Weight

The CDS-2020 Child Weight was based on the CDS-2019 child weight, the construction of which is described in the CDS-2019 User Guide.³ The CDS-2020 child weight incorporated an adjustment for non-response based on a regression model that predicted which cases from CDS-2019 completed the CDS-2020 interview using a comprehensive set of covariates. The non-response adjusted weight was then post-stratified to 2019 population totals.

Step 1. Non-Response Adjustment

A non-response adjustment factor for the weight was obtained from a logistic regression model of the response outcome. All eligible CDS-2020 child cases were included in the model.

It is possible to reduce non-response bias without increasing sampling variance of the survey estimates by including in the non-response model covariates that are correlated with both the survey response and the study outcomes.⁴ For this reason, the following substantive measures from the CDS-2019 PCG-Child Interview or PCG-Household Interview were incorporated in the non-response models for the CDS-2020 Child Weight:

- Child Behavioral Problems Index (a scale, ranging from 0–27),
- Safety of the local neighborhood (four-category response),
- Child health status (five-category assessment),
- Household food security status (four-category variable, based on a scale), and
- PCG K-6 psychological distress (scale, ranging from 0–24).

Because participation in CDS-2020 is conditional on completing the CDS-2019 interview, these variables are available for both respondents and non-respondents. Additional model covariates were obtained from the 2019 Core PSID and from coding the proximity of the interviewer and the travel cluster size. The logistic regression model predicted a response indicator, y , with $y=0$ if the case was non-response and $y=1$ if the case had a completed CDS-2020 interview. The estimated coefficients and standard errors are reported in Appendix Table A.1.

The regression model results indicate that the probability of response in CDS-2020 was higher among children in households with a male reference person, in households in the North Central region, for children residing outside of the U.S., in households in metro areas, in households in travel clusters, and for children of PCGs with medium-high distress level. The probability of response was lower for children in households with lower family income, in households with more children, and for children with a higher level of behavior problems. Although a number of variables in the model are not statistically significant predictors of CDS-2020 response, they were all retained in the model used to derive predicted probabilities of response. Overall, the Hosmer-Lemeshow test of goodness of fit test ($\chi^2=9.13$, 8 df, $p=0.33$) suggests that the response model provides an acceptable fit.

Based on the estimated logistic regression model, predicted probabilities of response were computed for each case included in the model and grouped into deciles. These decile groups served as the classes within which a uniform non-response weighting adjustment was applied. Each CDS-2020 child response case was assigned a non-response adjustment factor equal to

³ See “Panel Study of Income Dynamics, Child Development Supplement 2019: User Guide,” Institute for Social Research, University of Michigan, 2022.

⁴ Little, R.J., & Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology*, 31(2), 161.

the inverse of the median predicted probability of successful completion of the CDS-2020 interview within its decile weighting class. The median response propensity and adjustment factor for each decile of the predicted probability response are shown in Table 5.2.

Table 5.2. Median Response Propensity and Weighting Adjustment Factor for CDS-2020 Child Weight

Response propensity decile	Median response propensity	Adjustment factor
1	0.237	4.212
2	0.297	3.366
3	0.335	2.988
4	0.368	2.715
5	0.401	2.497
6	0.431	2.321
7	0.467	2.141
8	0.511	1.958
9	0.574	1.742
10	0.679	1.472

The probability of selection weight for each CDS-2020 child response case was then multiplied by the non-response adjustment factor to produce an interim weight that adjusts for the non-response to the CDS-2020 interview.

Step 2. Non-U.S. Cases

There were 37 children in CDS-2020 with interview data that resided outside the U.S. during the fieldwork period. Although interviews were attempted for all of these cases and completed among some of them, these cases are not included in the post-stratification because the control totals for the post-stratification process are based on the U.S. resident population. At this step, for the non-U.S. cases, the Child Weight is designated to be complete.

Step 3. Trimming of Weights

The distribution of the interim, attrition-adjusted weights was examined and a decision was made to trim extreme values at each end of the distribution. The trimming rule, applied to the attrition-adjusted Child Weight from Step 1, assigned all cases with weights in the top one percent and in the bottom one percent of the distribution to, respectively, values at the 99th and 1st percentiles.

Step 4. Post-Stratification to Population Control Totals

We next post-stratified the trimmed, attrition-adjusted weights from Step 3 to population control totals from the 2019 American Community Survey. Post-stratification cells were formed based on the following respondent characteristics:

- Child sex (male/female)
- Birth year of child (2002–2018)
- Child race/ethnicity (Hispanic, non-Hispanic Black, non-Hispanic White, or other)
- Census region (Northeast, Midwest, South, West)

Post-stratification cells defined by the full four-way cross-classification of these categorical variables were collapsed as needed to ensure a minimum count of approximately 15–20 individuals in each cell, except for some cells for birth year 2002 that were retained in order to keep the estimated proportion of children in that birth year close to the population estimates. The post-stratification adjustment factors were computed as the ratio of the ACS control totals to the CDS-2020 weighted population estimate (using the interim weight from Step 3). Appendix Table A.2 shows the CDS-2020 sample count, CDS-2020 weighted estimates, the ACS population estimates, and the post-stratification adjustment factors for each of the 109 cells defined by birth year, sex, race/ethnicity, and region.

The post-stratification adjustment factors were applied to the interim weight to produce a post-stratified weight.

Step 5. Combining the U.S. and Non-U.S. Cases

The final step in creating the Child Weight was to combine the weights from Step 2 for non-U.S. cases with the weights from Step 4 for cases in the U.S.

Method to Construct PCG Weight

A PCG Weight can be derived from the Child Weight. In particular, PCGs can be assigned the average weight over all children for who they were the responsible primary caregiver. The PCG Weight will be made available in a subsequent revised final release of the CDS-2020 data.

Child Longitudinal Weight

The Child Longitudinal Weight is still under construction and will be made available in a subsequent revised final release of the CDS-2020 data.

Summary of Weights

In Table 5.3 we list the cross-sectional Child Weight and present case counts and summary statistics. The estimated U.S. population of children born from 2002 to 2018 and not living in institutional group quarters for 2019 (the target year for the CDS-2020 sample) is 69,281,033. The sum of the Child Weight is slightly higher because of the inclusion of the respondents residing outside of the U.S. The weighted total population of PCGs in CDS-2019 is 39 million.

Table 5.3 Summary of CDS-2020 Cross-sectional Weights

Weight type (variable name)	Count	1st pct.	50th pct.	99th pct.	Mean	Std. Dev.	Coef. Var.	Sum total
Child Weight (X19CVWGT)	1,997	1,351.73	28,713.21	149,700.87	35,117.08	30,846.42	87.84	70,128,801

Recommendations for Using the Weights

In this section, we summarize our recommendations for using the CDS-2020 weights. Our basic recommendation is for data users to use the provided weights in all analyses. In addition, we recommend that, when calculating standard errors, data users should wherever possible account for the clustering of the CDS-2020 data. To account for the stratification and clustering in the Core PSID sample design, the analyst can use the sampling error stratum (ER31996) and sampling error cluster (ER31997) variables. Because CDS-2020 comprises a subset of the Core

PSID sample, users may encounter instances where a cluster includes a single observation when analyzing the CDS-2020 data. Several statistical software programs have options to handle the single cluster issue and we recommend reading the statistical software manual or consulting with a survey statistician when this arises. Analysts could also consider accounting for the clustering of the sample by family so that the standard errors reflect the fact that siblings are more likely to have similar outcomes and characteristics than children selected at random. Controlling for family-level clustering of siblings also provides an appropriate correction due to clustering of families by household or neighborhood and recognizes the fact that often it is only possible to control for a single level of clustering.

When analyses focus on a subset of children (from the full sample, for the in-home components, or the time diaries), data users should use an appropriate “sub-population” adjustment. Clustering-corrected standard errors and sub-population commands are available in most standard statistical software (including SAS and Stata).

Child Weight (X19CVWGT)

This weight should be used for all analyses of child data from CDS-2020 when the child is the unit of observation for the analysis. This weight should also be used for any analysis that compares outcomes for the same children between CDS-2019 and CDS-2020.

Finally, if users have questions about whether their analyses should be weighted or unweighted or about how to reflect the sampling design in their calculation of parameter estimates and standard errors, they should consult with a survey statistician.

Appendix

Table A.1. Logistic Regression Model Results for CDS-2020 Child Response

Variable	Estimate	Std. err.	P-value	Significance
PSID sample component				
SRC sample (ref.)	.	.	.	
SEO sample	0.120	0.124	0.334	
1997/1999 new immigrant sample	0.223	0.149	0.135	
2017/2019 new immigrant sample	0.213	0.153	0.165	
Child is male (0/1)	0.087	0.062	0.163	
Child birth year				
2002–2007 (ref.)	.	.	.	
2008–2012	-0.057	0.080	0.479	
2013–2016	0.094	0.089	0.292	
2017–2018	0.384	0.365	0.293	
Child race/ethnicity				
Non-Hispanic White (ref.)	.	.	.	
Hispanic	-0.102	0.125	0.412	
Non-Hispanic Black	-0.139	0.124	0.264	
Non-Hispanic Other	-0.182	0.245	0.459	
Age of household reference person				
≤30 years	-0.042	0.094	0.656	
31–45 years (ref.)	.	.	.	
≥46 years	0.050	0.090	0.581	
Household reference person is male (0/1)	0.352	0.087	<.0001	***
Education of household reference person				
≤11 years	-0.041	0.111	0.710	
12 years	0.165	0.094	0.079	
13–15 years	0.042	0.090	0.645	
≥16 years (ref.)	.	.	.	
Education unknown	0.133	0.292	0.650	
Household reference person is employed (0/1)	0.215	0.131	0.100	
Family income quartile				
1st quartile	-0.258	0.123	0.036	*
2nd quartile	-0.326	0.103	0.002	**
3rd quartile	-0.219	0.093	0.018	*
4th quartile (ref.)	.	.	.	
Region				
South (ref.)	.	.	.	
Northeast	-0.091	0.113	0.421	
North Central	0.165	0.081	0.041	*
West	-0.038	0.097	0.694	
Outside U.S.	2.696	0.623	<.0001	***
Metro area (0/1)	0.452	0.096	<.0001	***
Number of children in the family unit				
1	0.103	0.087	0.233	
2 (ref.)	.	.	.	
3	-0.159	0.080	0.047	*
4+	-0.158	0.096	0.101	
Proximity of field interviewer and cluster size				
Within interviewer's local area (ref.)	.	.	.	
Area requires interviewer travel – large clusters	0.266	0.101	0.009	**
Area requires interviewer travel – medium clusters	0.838	0.093	<.0001	***
Area requires interviewer travel – small clusters	1.143	0.112	<.0001	***
Child's individual weight from 2019 Core PSID	<.0001	<.0001	0.389	

Variable	Estimate	Std. err.	P-value	Significance
Child behavioral problems index (BPI, 0–27)				
0	0.156	0.108	0.150	
1–3 (ref.)	.	.	.	
4–6	-0.009	0.100	0.926	
7–10	0.067	0.102	0.511	
11–15	0.140	0.114	0.220	
16+	-0.389	0.137	0.005	*
Child age ≤2 years	-0.436	0.358	0.224	
Safety of local neighborhood				
Completely safe	0.076	0.070	0.279	
Fairly safe (ref.)	.	.	.	
Somewhat dangerous	0.134	0.102	0.191	
Extremely dangerous	-0.144	0.231	0.533	
Child health status				
Excellent (ref.)	.	.	.	
Very good/good/fair/poor	0.119	0.065	0.068	
Family food security status				
High food security				
Marginal food security	-0.034	0.105	0.746	
Low food security	-0.016	0.111	0.883	
Very low food security	-0.270	0.147	0.067	
PCG psychological distress scale (0–24)				
0 (ref.)	.	.	.	
1	-0.011	0.129	0.930	
2	0.022	0.121	0.856	
3	-0.052	0.122	0.674	
4	0.100	0.126	0.429	
5	0.130	0.137	0.342	
6	0.140	0.138	0.310	
7–9	0.335	0.130	0.010	*
10+	-0.005	0.149	0.975	
Hosmer and Lemeshow goodness-of-fit test	9.1311 (8 df) p=0.3314			

Note: * P<=0.05, ** P<=0.01, *** P<=0.001; N=4,629 (response=1,997, nonresponse=2,632).

Table A.2. Post-Stratification Cells for CDS-2020 Weight

Cell	Birth Year	Sex	Race/ethnicity	Region	CDS sample size	CDS weighted estimate	ACS population totals	Adjustment factor
1	2002	F	Hispanic	All Regions	11	485,063	475,647	0.98059
2	2002	F	NH White/Other	Midwest	9	383,082	318,056	0.83026
3	2002	F	NH White/Other	Not Midwest	12	769,940	872,170	1.13278
4	2002	M	Hispanic	All Regions	9	240,905	514,753	2.13675
5	2002	M	NH White/Other	All Regions	16	870,094	1,260,990	1.44926
6	2002-2003	F	NH Black	Not South	10	285,368	268,036	0.93926
7	2002-2003	F	NH Black	South	18	222,009	340,203	1.53238
8	2002-2003	M	NH Black	Not South	13	159,630	271,896	1.70328
9	2002-2003	M	NH Black	South	28	553,932	359,740	0.64943
10	2003-2004	F	Hispanic	Not West	13	659,865	591,406	0.89625
11	2003-2004	F	Hispanic	West	13	538,742	399,100	0.74080
12	2003-2004	F	NH White/Other	Midwest	27	1,002,677	633,359	0.63167
13	2003-2004	F	NH White/Other	South	14	619,990	845,534	1.36379
14	2003-2004	F	NH White/Other	West	12	577,338	524,036	0.90768
15	2003-2004	M	Hispanic	Not West	19	669,999	608,232	0.90781
16	2003-2004	M	Hispanic	West	10	308,803	402,897	1.30470
17	2003-2004	M	NH White/Other	Midwest	20	1,067,065	668,806	0.62677
18	2003-2004	M	NH White/Other	South	13	879,281	875,939	0.99620
19	2003-2004	M	NH White/Other	West	10	469,967	551,583	1.17366
20	2003-2006	F	NH White/Other	Northeast	10	632,915	875,169	1.38276
21	2003-2006	M	NH White/Other	Northeast	21	1,265,016	906,924	0.71693
22	2004-2006	F	NH Black	Not South	24	551,461	424,448	0.76968
23	2004-2006	F	NH Black	South	41	829,998	562,521	0.67774
24	2004-2006	M	NH Black	South	44	819,711	580,795	0.70854
25	2004-2008	M	NH Black	Not South	25	709,577	757,551	1.06761
26	2005-2006	F	Hispanic	Not West	10	462,012	655,029	1.41778
27	2005-2006	F	NH White/Other	South	20	1,047,750	857,334	0.81826
28	2005-2006	F	NH White/Other	West	14	514,339	531,805	1.03396
29	2005-2006	M	Hispanic	Not West	18	546,593	687,211	1.25726
30	2005-2006	M	Hispanic	West	11	337,047	442,686	1.31342
31	2005-2006	M	NH White/Other	Midwest	13	505,833	661,317	1.30738
32	2005-2006	M	NH White/Other	South	14	721,143	884,863	1.22703
33	2005-2008	F	Hispanic	West	17	902,897	852,104	0.94374
34	2005-2008	F	NH White/Other	Midwest	35	1,193,559	1,273,634	1.06709
35	2005-2008	M	NH White/Other	West	16	986,080	1,113,466	1.12918
36	2007-2008	F	Hispanic	Not West	11	800,190	678,119	0.84745
37	2007-2008	F	NH Black	Not South	20	396,891	287,591	0.72461
38	2007-2008	F	NH Black	South	25	377,719	388,714	1.02911
39	2007-2008	F	NH White/Other	South	10	650,079	850,819	1.30879
40	2007-2008	F	NH White/Other	West	20	747,351	531,328	0.71095
41	2007-2008	M	Hispanic	Not West	18	891,626	690,841	0.77481
42	2007-2008	M	Hispanic	West	10	255,124	451,009	1.76780
43	2007-2008	M	NH Black	South	17	305,180	421,054	1.37969
44	2007-2008	M	NH White/Other	Midwest	15	441,047	646,950	1.46685
45	2007-2008	M	NH White/Other	South	15	779,056	907,621	1.16503
46	2007-2010	F	NH White/Other	Northeast	10	601,525	809,706	1.34609
47	2007-2010	M	NH White/Other	Northeast	17	1,081,148	863,925	0.79908

Cell	Birth Year	Sex	Race/ethnicity	Region	CDS sample size	CDS weighted estimate	ACS population totals	Adjustment factor
48	2009-2010	F	NH Black	Not South	14	495,127	304,402	0.61480
49	2009-2010	F	NH Black	South	23	247,134	372,270	1.50635
50	2009-2010	F	NH White/Other	Midwest	18	597,185	597,582	1.00067
51	2009-2010	F	NH White/Other	South	13	808,360	800,309	0.99004
52	2009-2010	F	NH White/Other	West	13	568,769	513,035	0.90201
53	2009-2010	M	Hispanic	Not West	20	796,641	675,833	0.84835
54	2009-2010	M	Hispanic	West	12	503,280	444,264	0.88274
55	2009-2010	M	NH Black	South	25	453,593	376,979	0.83110
56	2009-2010	M	NH White/Other	Midwest	17	608,330	630,720	1.03681
57	2009-2010	M	NH White/Other	South	24	1,268,219	837,709	0.66054
58	2009-2012	F	Hispanic	Not West	22	1,060,575	1,252,649	1.18110
59	2009-2012	F	Hispanic	West	17	703,142	792,616	1.12725
60	2009-2012	M	NH Black	Not South	27	611,556	587,397	0.96050
61	2009-2012	M	NH White/Other	West	16	861,296	1,060,405	1.23117
62	2011-2012	F	NH Black	Not South	17	273,400	276,453	1.01117
63	2011-2012	F	NH Black	South	25	438,495	364,694	0.83169
64	2011-2012	F	NH White/Other	Midwest	15	608,021	577,688	0.95011
65	2011-2012	F	NH White/Other	South	14	635,229	776,918	1.22305
66	2011-2012	F	NH White/Other	West	12	575,169	491,051	0.85375
67	2011-2012	M	Hispanic	Not West	16	774,871	639,708	0.82557
68	2011-2012	M	NH Black	South	28	378,619	373,436	0.98631
69	2011-2012	M	NH White/Other	Midwest	13	444,114	620,502	1.39717
70	2011-2012	M	NH White/Other	South	20	1,024,130	807,539	0.78851
71	2011-2014	F	NH White/Other	Northeast	17	843,035	761,374	0.90313
72	2011-2014	M	Hispanic	West	16	544,305	777,998	1.42934
73	2011-2014	M	NH White/Other	Northeast	11	584,497	787,321	1.34701
74	2013-2014	F	Hispanic	Not West	12	462,965	576,895	1.24609
75	2013-2014	F	Hispanic	West	13	421,722	368,271	0.87326
76	2013-2014	F	NH Black	South	28	506,089	353,405	0.69831
77	2013-2014	F	NH White/Other	Midwest	12	438,294	582,992	1.33014
78	2013-2014	F	NH White/Other	South	13	603,292	760,276	1.26021
79	2013-2014	F	NH White/Other	West	19	593,048	498,543	0.84065
80	2013-2014	M	Hispanic	Not West	17	548,724	615,231	1.12120
81	2013-2014	M	NH Black	Not South	18	277,343	280,601	1.01175
82	2013-2014	M	NH Black	South	30	301,419	351,168	1.16505
83	2013-2014	M	NH White/Other	Midwest	15	557,223	616,656	1.10666
84	2013-2014	M	NH White/Other	South	20	1,046,449	803,301	0.76764
85	2013-2016	F	NH Black	Not South	24	598,006	549,632	0.91911
86	2013-2018	M	NH White/Other	West	29	1,446,093	1,554,085	1.07468
87	2015-2016	F	NH Black	South	17	245,405	348,937	1.42188
88	2015-2016	F	NH White/Other	Midwest	21	653,822	588,127	0.89952
89	2015-2016	F	NH White/Other	South	25	794,015	800,410	1.00805
90	2015-2016	M	Hispanic	Not West	10	498,542	656,360	1.31656
91	2015-2016	M	NH Black	Not South	20	260,240	276,433	1.06222
92	2015-2016	M	NH Black	South	34	366,939	370,436	1.00953
93	2015-2016	M	NH White/Other	Midwest	24	743,000	637,156	0.85754
94	2015-2016	M	NH White/Other	South	16	531,077	850,527	1.60151
95	2015-2018	F	Hispanic	Not West	21	1,177,613	1,246,383	1.05840

Cell	Birth Year	Sex	Race/ethnicity	Region	CDS sample size	CDS weighted estimate	ACS population totals	Adjustment factor
96	2015-2018	F	Hispanic	West	17	678,648	757,664	1.11643
97	2015-2018	F	NH White/Other	Northeast	17	989,579	761,914	0.76994
98	2015-2018	F	NH White/Other	West	19	964,631	999,934	1.03660
99	2015-2018	M	Hispanic	West	14	745,889	807,498	1.08260
100	2015-2018	M	NH White/Other	Northeast	28	1,296,104	813,500	0.62765
101	2017-2018	F	NH Black	Not South	11	158,218	274,735	1.73643
102	2017-2018	F	NH Black	South	23	290,807	336,290	1.15640
103	2017-2018	F	NH White/Other	Midwest	16	580,323	573,921	0.98897
104	2017-2018	F	NH White/Other	South	27	790,325	730,889	0.92480
105	2017-2018	M	Hispanic	Not West	15	646,086	651,843	1.00891
106	2017-2018	M	NH Black	Not South	14	230,241	276,429	1.20061
107	2017-2018	M	NH Black	South	24	390,738	351,496	0.89957
108	2017-2018	M	NH White/Other	Midwest	21	535,744	602,204	1.12405
109	2017-2018	M	NH White/Other	South	16	738,314	779,122	1.05527