INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER
PANEL STUDY OF INCOME DYNAMICS

Panel Study of Income Dynamics, Child Development Supplement 2021: User Guide

April 2024

This user guide was prepared by Noura Insolera, Narayan Sastry, and Beth Simmert

Abstract

The 2021 Child Development Supplement (CDS-2021) 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 and neighborhood context. CDS-2021 was designed as a two-year follow-up to a main wave of CDS conducted in 2019 in order to support research on the effects of a cohort of children aged 2–17 years in 2021 for whom prior interview data in CDS-2019 were collected just prior to the onset of the COVID-19 pandemic in March 2020. An additional round of CDS data collection occurred in 2020 for about two-thirds of children in the CDS-2019 sample, providing data for up to three points in time for many children. CDS-2021 builds on the strengths of PSID, a genealogical study of US families that began in 1968. The interview content is highly comparable across the three waves of CDS in 2019, 2020, and 2021 and prior waves of the study. All publicly available CDS data are free of charge through the PSID Online Data Center and the CDS Online Data Center. CDS-2021 sensitive data are available to researchers through a special application procedure and restricted data are available through a contract. This User Guide provides essential information to researchers planning or undertaking research using the CDS-2021 data.

Suggested citation of the CDS-2021 User Guide:

> "Panel Study of Income Dynamics, Child Development Supplement 2021: User Guide," Institute for Social Research, University of Michigan, 2024.

Suggested citation of the CDS-2021 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-2021 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 HD104671.

Preface

The 2021 Child Development Supplement (CDS) to the Panel Study of Income Dynamics (PSID) was funded by Grant R01 HD104671from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD). Additional funding was provided by the US Department of Agriculture's Food and Nutrition Service. We are very grateful to these sponsors for their support of CDS-2021.

Contents

Acknowledgements

The 2021 wave of the Child Development Supplement (CDS-2021) 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-2021 was directed by Narayan Sastry, with Paula Fomby serving as the associate director through June 2022. Yi-Miau Tsai was the CDS-2021 project manager.

Development of the questionnaires for CDS-2021, 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 led by Allison Mageli and Rose McAloon-Fernando.

The design and implementation of the CDS-2021 weights was undertaken by Wen Chang and Raphael Nishimura.

Within the Survey Research Operations (SRO) unit, which undertook the fieldwork for CDS-2021, Sarah Crane served as the SRO Project Lead with the consultation of Rachel Orlowski (PSID Survey Director) and Stephanie Chardoul (PSID Senior Project Advisor and SRO Director). Among the many other key SRO staff members who contributed significantly to CDS-2021 were Maryam Buageila, Barbara Aghababian-Homburg, Peggy Lavanger, Minako Edgar, Megan Gomez-Mesquita, Brad Goodwin, Youhong Liu, and Peter Sparks.

PSID staff members Flannery Campbell, Paul Juska, and Jacob Rainey were responsible for data processing activities.

Special thanks go to other members of PSID who contributed to CDS-2021, including April Beaule, Vivian Burgett, Jennifer Huntington, Noura Insolera, Dennis Kloska, Mohammad Mushtaq, and Beth Simmert.

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.

CHAPTER 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.

CDS-2021 was an "extra" wave of CDS that was conducted two years following the last main wave of CDS in 2019. Since its launch in 1997 and its relaunch in 2014, CDS waves have been repeated every five years. The reason for conducting an additional wave of CDS in 2021 was to collect follow-up information on the effects of the COVID-19 pandemic on age-eligible children (aged 2–17 years) who participated in CDS-2019. Interview data were collected on most of these children in CDS-2019 prior to the onset of the COVID-19 pandemic in March 2020. An additional round of CDS data collection occurred in 2020 for about two-thirds of children in the CDS-2019 sample, providing data for up to three points in time for many children.

CDS data support studies of health, development, and well-being in childhood; the relationship between children's characteristics and contemporaneous family decision-making 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.

Six full and follow-up waves of CDS have been completed since the launch of the Original CDS in 1997. This user guide provides information about the seventh wave, CDS-2021, its study design, questionnaire instruments and measures, fieldwork outcomes, its relationship with Core PSID and other components of PSID, and the data structure.

There are several features of CDS that provide substantively unique research opportunities. First, 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—as well as the lives of the parents' parents (the CDS-2021 children's grandparents). These data can be combined to study intergenerational transmission of human and social capital. Intra-generational analysis is also possible because CDS includes siblings and cousins. Second, the original CDS (1997–2007) and the ongoing CDS (2014 and beyond) allow researchers to study cohort differences in development between children born from 1985 to 1996 and those born from 2002 to 2019, as well as differences between younger and older members of these cohorts. Third, many ongoing CDS children were born to members of the original CDS cohort, providing opportunities to examine intergenerational connections in child development and behavior. Fourth, as CDS children move into adulthood, they will be interviewed in the PSID Transition into Adulthood Supplement (TAS) and will also become primary PSID respondents. The information collected in CDS provides invaluable insights into the effects of childhood experiences and circumstances on adult social, demographic, economic, and health outcomes. Fifth, genetic markers, collected from children, parents, and grandparents in CDS-2014 and CDS-2019, allow researchers to address a number of important scientific questions that span the interests of population geneticists and social scientists. Finally, CDS-2021, together with CDS-2019 and the CDS-2020 follow-up, offers an opportunity to analyze the effects of a global pandemic on child health and development and family dynamics and well-being at three distinct points in time for a nationally representative sample of children.

In this chapter, we provide background on CDS and PSID, an overview of CDS-2021, and an outline of this user guide.

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, philanthropy, and religiosity. CDS data support studies of health, development, and well-being in childhood; the relationship between children's characteristics and contemporaneous family decision-making 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. For more information on CDS, visit the PSID documentation page.

Public use data from CDS are available free of charge through the CDS Online Data Center and the PSID Online Data Center, 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. 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.

CDS is part of the Panel Study of Income Dynamics, a longitudinal survey of a nationally representative sample of US families that began in 1968.[1] 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 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 US were added to PSID.

Original CDS

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%) who provided data on 2,907 children and

[1] 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.

adolescents aged 5–18 years. During 2007, 1,506 children aged 10–19 years were successfully re-interviewed (90%) in the third and final wave of the original CDS cohort study.

Ongoing CDS

By 2014, all children in the original 1997 CDS cohort had reached age 18 years, 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 US. CDS-2014 interviews were conducted by telephone with PCGs and adolescents aged 12–17 years; 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 was a scheduled full-scale study, with both telephone interviews and home visits, and designed as the second wave of the ongoing CDS. It continued the shift in orientation for the collection of information on PSID children from a study of a single cohort (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).

Data collection for CDS-2019 began in October 2019 and was scheduled to continue through mid-2020. However, in mid- March of 2020, CDS-2019 home visits were halted due to the COVID-19 pandemic.[2] At that time, telephone interviews had been completed with about three-quarters of primary caregivers and adolescents aged 12–17 years (N=~900 families). There was no opportunity to restart the home visits because of the ongoing pandemic. 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

[2] 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.

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 in September 2020 and ended on 31 December 2020.

CDS-2021 Overview

CDS-2021 was designed to collect reinterview data on PCGs and age-eligible children (aged 2–17 years) who participated in CDS-2019. Data collection occurred between November 2021 and June 2022. Only families who completed an interview for PSID-2021 were eligible to participate in CDS-2021. To reduce respondent burden, in families with four or more eligible children, three children were randomly selected for CDS-2021; seventy-two children who were otherwise eligible were not selected. The CDS-2021 interview content closely parallels CDS-2019 and CDS-2020, although new items were added on the effects of the COVID-19 pandemic on family health and finances and children's schooling, activities, and well-being.

CDS-2021 can be used as a standalone data source that captures the cross-sectional experiences of families and children in 2021. However, it can also be used in conjunction with the other COVID-19 era waves of CDS to study pandemic influences on a nationally representative sample of children at four different points in time. CDS-2019 serves as the baseline, or "pre-pandemic" observation point; CDS-2020 assesses the impact of COVID-19 in the short-term and CDS-2021 in the medium-term; and the in-production CDS-2023 will assess the pandemic's longer-term impact on children in PSID families.

Overview of the CDS-2021 User Guide

This user guide provides information about the CDS-2021 study design, questionnaire instrument and measures, fieldwork outcomes, data structure, and relationship with Core PSID and other components of PSID.

In Chapter 2, we provide a brief description of the CDS-2021 questionnaire instrument content. In Chapter 3, we provide an outline of the CDS-2021 sample. In Chapter 4 we describe the CDS-2021 data file structure and the procedures for merging files. Finally, in Chapter 5 we describe the construction and use of the CDS-2021 weights.

CHAPTER 2. THE CDS-2021 QUESTIONNAIRE, MEASURES, AND VARIABLES

In this chapter, we provide a brief overview of the CDS-2021 questionnaire. We begin by describing the general principles that guided the design and content of the CDS-2021 questionnaire. Next, we describe the questionnaire content domains and provide an overview of measures. We then describe naming conventions. Finally, we describe the CDS-2021 questionnaire and the major sections in detail.

See the CDS Cross-Wave Variable Index for a comprehensive list of questionnaire items and scales available in CDS-2021 and prior waves. This index is available in the documentation section of the PSID website and the CDS website.

General Principles for the CDS-2021 Questionnaire

- The 2021 PSID Child Development Supplement was designed to support research on children's health and well-being within family, neighborhood, and school contexts during the COVID-19 pandemic.

- Continuity with CDS-2019, CDS-2020, and CDS I–III (1997–2007). The CDS-2021 questionnaire was designed to provide directly comparable to measures to CDS-2019, CDS-2020, and earlier waves of CDS in several key content areas including food security, mental health, and behavior problems. The description of individual questionnaire modules below includes information about items or content areas that were omitted, revised, or added during CDS-2021 questionnaire development.

- Updated and new content. Although CDS-2021 closely paralleled CDS-2019 and CDS-2020, some changes to content items were implemented where necessary. Key changes and updates included:

    - A module included in CDS-2020 about school closure, attendance, and remote learning during the 2019–2020 school year was continued in CDS-2021, covering the 2020-2021 school year.

    - The COVID-19 module was updated for CDS-2021 from the version implemented in CDS-2020.

    - CDS-2021 included no in-home or mail-out/mail-back components. There was thus no collection of time diaries, reading and math skills assessments, anthropometric measurements, or saliva samples.

Questionnaire Content Domains

Table 2.1 summarizes the questionnaire content domains in CDS-2021.

Table 2.1. CDS-2021 Content Domains

Content Domain	Description of content
Health status & behaviors	Health-related limitations and chronic conditions; obesity; health care utilization; smoking; health insurance
Psychosocial & social well-being	Prosocial behavior; social integration; social identity; social anxiety; behavior problems; strengths and difficulties; depression; self-esteem; social, emotional, and psychological well-being; risky behaviors; thrill seeking; anti-social behaviors; drug and alcohol abuse /dependence; peer bullying
Family environment	HOME scale for cognitive & emotional stimulation; parental involvement, closeness, time spent and conflict with father, mother, and parent figures; household composition
Sibling relationships	Type and frequency of cooperation with, kindness towards, and helping behaviors towards siblings
Peer influences	Closeness to friends; friends' activities
Parent monitoring	Caregivers' knowledge of the child's whereabouts, activities, and associations; child disclosure of activities
Non-coresident parents	Conflict between resident and non-coresident parent
Child care	Type, frequency of use, and costs of arrangements for CDS children up to sixth grade
Caregiver social & psychological resources	Rosenberg self-esteem scale; Kessler K6 30-day psychological distress scale; social support; parenting attitudes; aggravation in parenting; gender role beliefs; family conflict; economic strain; work schedules
Spending & saving	Variety of expenditures for child; savings mechanisms
Work & wages	Employment experiences for older children; job aspirations
Education	Parental expectations; enrollment; type of school; tuition; attendance; government lunch & breakfast programs; attended special class/school for gifted students; special education; repeated grade; dropped out
Work & education expectations	Economic expectations; occupational identity; job values, career orientation and expectations for future work and schooling
Computer & media use	Access to television, computers, smartphones, and other digital devices; frequency of television, computer, and social media use
Intellectual skills & abilities	Ability self-concepts in reading and math

Variable Naming Conventions

This user guide refers to individual items by their names as they appear in the questionnaire, typically a one-letter section prefix followed by one or more digits in order of item sequence. For example, item J1 in the Primary Caregiver (PCG) Interview refers to the first item appearing in Section J: Neighborhood Measurements of the Household Interview. The questionnaires are downloadable under Documentation in the PSID Online Data Center and the CDS Online Data Center.

Variables associated with specific interview components are named using the following structure:

1. The leading character(s) refers to the study component from which the questionnaire item is drawn:
   - H = PCG Household Interview
   - P = PCG Child Interview
   - C = Child Interview (interview completed by CDS child)
   - X = Demographics
   - R = Roster

2. The following two characters in the variable refer to the calendar year that data collection began. For variable names associated with CDS-2021, these characters are always "21." 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 variable constructed from the household food security items for Household Food Security Status is H21HHFD.

The list of variables, generated scales, and source of each item is available on the CDS generated variables and scales tab in the CDS Cross-Wave Variable Index.

Description of Questionnaire Modules and their Major Sections

Primary Caregiver Household Interview (PCG-HH File)

The PCG Household (PCG-HH) Interview focuses on the characteristics of a child's family, household, and neighborhood. The interview also collects extensive information on the PCG's own psychological resources, social support, parenting stress, parenting style, and childrearing values. Unless otherwise noted, items in the PCG-HH interview were administered to all PCGs.

Topics included in the PCG-HH Interview are described below. The CDS Cross-Wave Variable Index documents the source and original author of questionnaire content where appropriate. Note that PCG-HH Interview content begins with Section J (Neighborhood Measurements). Sections A–H appear in the PCG-Child interview.

Neighborhood Measurements (Section J). Eight items assess the PCG's perception of neighborhood quality, including residential stability, residential satisfaction, neighborhood anonymity, social cohesion, and neighborhood safety. The series appears in the questionnaire as Items J1 to J8.

PCG Self Esteem (Section K). The Rosenberg Self-Esteem Scale measures global self-worth. The scale is widely used, with substantial documentation on its validity and reliability. PCGs reported on a series of ten items using a response scale ranging from 1 to 4, where 1 indicates "Strongly Disagree" and 4 indicates "Strongly Agree." The series appears in the PCG-HH Instrument as Items K1 to K10. The scale score is computed as an average of responses to these ten items and is available for respondents who have valid values on at least eight items (H21SLFEST). A new Scale score has been added for CDS-2021 to aid in interpretation (H21SLFSC).

Childrearing Values (Section M). Respondents ranked the qualities or traits they consider most important to prepare a child for life from a set of five choices. Traits include obedience, popularity, autonomy, a strong work ethic, and altruism. The series appears in items M3A to M3D. These items appeared in the Detroit Area Study and the General Social Survey.

Aggravation in Parenting (Section M). The aggravation in parenting scale (M4–M10) measures parenting stress that may result from changes in employment, income, and other factors in the lives of PCGs. Items M4 to M7 address parenting in general. Items M8 to M10 focus on the PCG's feelings about his/her children in CDS collectively. The generated variable is a mean score derived from the seven items in the scale (H21PARENT). A mean score was computed for all cases with valid values on at least five items.

Work/Life Adjustments for Children (Section M). In Items M11 to M13, PCGs reported whether they ever changed neighborhoods or employment to improve circumstances for their children.

Attitudes about Gender Roles (Section M). Items M14 to M27 measure the PCG's level of agreement with statements pertaining to gender role attitudes, including three statements drawn from the "Being a Father" Scale. These statements measure the constructs of traditional marriage values, traditional mothering values, equity, and father involvement. Each construct is represented by three variables.

PCG Psychological Distress (Section N). The Kessler 6 (K-6) Non-Specific Psychological Distress Scale (N1–N6) 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, TAS, 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 H21K6_14 is a sum score computed for all cases with valid responses to all six items in the scale. A new scale score has been added for CDS-2021 to aid in interpretation (H21K6_SC).

The scale includes three follow-up items about persistence and impairment associated with symptoms of nonspecific distress (N7–N9). 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.

Perceived Social Support (Section N). Six items (N10–N15) describe the PCG's perceived practical and emotional support received from their spouse or partner, other family, and friends.

Family Pets (Section P). Section P includes a Pet Attachment scale based on seven questions about the number and types of pets in families and the PCG's interaction with and attachment to their pets (H21PCGPET). The source of the items is the Center for the Study of Human-Animal Relationships and Environments Pet Attachment Scale.

Disagreement in Parenting and Joint Goals (Section Q). The Parental Disagreement Scale measures the extent of agreement on daily activities between a PCG and his or her spouse or partner (Q1–Q5). The items were administered only to PCGs who had a spouse or cohabiting partner in the household. The generated variable H21DISAGR is a mean score derived from the five items in the scale for all cases with valid values on at least four items. Three items measure the extent to which the PCG and his or her spouse or partner have joint goals for the future (Q6–Q8). Five items measure methods of conflict resolution among family members (Q9–Q13).

Negative effects on children (Section Q). PCGs were asked whether there was anyone living in the household whose alcohol or drug use, and mental or physical health had a negative effect on the children in the household (Q14–Q19).

Food Security (Section R). The PCG-HH interview included an 18-item version of the US Household Food Security Survey Module developed by the Economic Research Service at the US Department of Agriculture (R1–R15). 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 information about household (R1–R8) and child (R9–R15) food security separately. These data allow the food security status of CDS-2021 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 (H21HHFOODR) and separately for adults (H21ADFOODR) and children (H21CHFOODR). A parallel set of items describes the food insecurity status of the household overall (H21HHFOOD) and of adults (H21ADFOOD) and children (H21CHFOOD) in the household. A raw sum score (H21FOOD6R) and a food insecurity status indicator (H21FOOD6) based on a six-item subset of questionnaire items (R2–R6) are also available.

Home Environment (Section S). The Home Environment section collects information about children's access to learning resources and technology in the home, the PCG's involvement in her or his children's school and learning at home, and the PCG's own school enrollment, employment circumstances, and religiosity.

Children's Access to Technology (Section S). Technology questions in Section S include the types and number of electronic devices in the home, including televisions, computers, tablets, cellular telephones, and smart speakers (S1–S5, S14A–S14AA); shared television viewing habits (S9–S10); and household rules about television viewing (S11–S13) and use of other electronic devices (S14G– S14L). This section also includes an adapted six-item web-use skills index originally developed by Hargittai and Hsieh to measure the PCG's familiarity with computer and internet-related terminology (S14N1–S14N6).

Home Observation for Measurement of the Environment (HOME Scale) (Section S). The HOME Scale measures characteristics of a child's home environment that are associated with cognitive development and emotional support. HOME Scale content in the PCG-HH interview includes questions about how often the family engages in specific activities together, including meals (M1), socializing (M2), and television viewing (S9–S10); the number of books in the home (including electronic books, S15–S16); and the number of books the PCG has read in the last year (S17–S18). Note that a calculated HOME Scale score is not provided.

School Involvement (Section S). Two items in the PCG-HH interview address the PCG's volunteer activities at his or her child(ren)'s school (S19–S20). Six other items addressing the PCG's school involvement are included in the PCG-Child interview.

Response to Poor Grades (Section S). Twelve items describe actions PCGs would expect to take in response to a child's poor grades (S21–S30B).

Own Schooling (Section S). PCGs report whether they are currently attending school, and if so, the number of hours they attend school each week and travel time (S31–S31B).

Employment Characteristics (Section S). PCGs report whether they are currently working, and if so, report on characteristics of their employment such as number of jobs, hours worked weekly, nonstandard work schedules, and commuting time (S32–S39).

Religiosity (Section S). PCGs reported how often they attended religious services in the past year (S39– S40) and on the importance of religion and spirituality in their lives (S41A–S42A).

Primary Caregiver and COVID-19 (Section S). The CDS-2020 COVID-19 Health module was dropped for CDS-2021 as the items it covered were added to Section S. Nine items capture the household and PCG's experience with COVID-19, including vaccination status (S55-S59).

Interview Observations (Section OB). Interviewers provided structured and open-ended observations on the interviews they conducted, on the respondents, and on the respondents' household environment.

Primary Caregiver Child Items (PCG-Child File)

In the rest of the PCG instrument, questions were asked based on the age and/or school grade range of the items for up to three children living in each CDS household. See the CDS-2021 questionnaire for age and/or school grade ranges for each item and for rules governing skip patterns throughout the instrument.

Child Health (Section A). Questions about the physical health of each child (A2–A19) are drawn from the National Health Interview Survey and from the National Longitudinal Survey of Youth. Topics include general health status (A2), birth weight, breastfeeding, medical care, immunization status, diagnosis of chronic conditions, asthma, and disability.

Questionnaire items about birth weight are directed to PCGs only when the child's birth weight does not appear in the birth history collected as part of the PSID Core interview (A4–A4_KG). Where birth weight was already available, this information is provided in the CDS-2021 Demographics file (X21BWTP1–X21OS3B1 [biological mother report], X21BWTP2– X21OS3B2 [biological father report], X21BWTP3–X21OS3B3 [adoptive mother report], and X21BWTP4–X21OS3B4 [biological mother report]).

Information on breastfeeding duration is collected only where a CDS child is aged 0–5 years at the time of the CDS Coverscreen Interview.

A new module was added to CDS-2021 to assess each child's experience with COVID-19, vaccination status, and masking practices in public and school (A10SC1-A10SC8).

Psychological Wellbeing, Personality, and Behavior (Section B). Modules in this section include the Strengths and Difficulties Questionnaire, prosocial behavior, and sibling interaction.

Prosocial scale (Section B). The Prosocial Scale was administered about all children aged 3–18 years. Responses for each child are recorded in items B1A, B1D, B1J, B1R, and B1U. Since all children are now eligible for the SDQ, items B43-B47 are no longer included in the instrument. A prosocial behavior scale score is available for all cases with a valid response to each of the five items in the scale (P21PROSOC).

Strengths and Difficulties Questionnaire (Section B). Twenty items from the Strengths and Difficulties Questionnaire (SDQ) were included in CDS-2021 for children aged 3–18 years. The

administration of the questionnaire included items assessing hyperactivity/inattention (B1B, B1K, B1P, B1V, and B1Z), emotional problems (B1C, B1H, B1N, B1Q, and B1Y), conduct problems (B1E, B1G, B1M, B1S1/B1S2, and B1W1/B2W2), and peer relationship problems (B1F, B1L, B1O, B1T, and B1X).

Four subscales are generated using the twenty preceding items, based on the SDQ items and scoring instruction. These subscales are each the rounded mean of non-missing responses to the five component items and are only calculated if at least three of the five items have a valid response. These include the SDQ Hyperactivity Scale (P21SDQH), the SDQ Emotional Scale (P21SDQEM), the SDQ Conduct Scale (P21SDQC), and the SDQ Peer Relationships Scale (P21SDQP).

Three summary scores are also available. The SDQ Externalizing Score (P21SDQE) is comprised of the SDQ Conduct and Hyperactivity subscale and the SDQ Internalizing Score (P21SDQN) is comprised of the Child SDQ Emotional and Peer Relationships subscale.

The SDQ Total Difficulties Score (P21SDQ) comprises all four subscales. These scores are only calculated if at least three items in each subscale are non-missing.

Sibling Interaction (Section B). Five items describe the frequency of helping and prosocial behaviors expressed toward siblings, if any in the household (B48–B52, children aged 3–11 years).

Parenting and Family Interaction (Section C). Section C includes information about family routines, parental monitoring, household rules, discipline, and parent-child discussion topics. The universe for the items in this section varies depending on child age and grade. Many of these items characterize the aspects of children's home environments that are conducive to cognitive development and emotional support.

Items pertaining to household rules were revised compared to CDS I–III in order to accommodate new response options. In all waves, respondents were asked about whether there were household rules governing a variety of activities, including where and how children spend their time, homework, and television viewing. Previously, the response options were limited to "Yes" (i.e., there are household rules) and "No" (i.e., there are no household rules). In order to better characterize how household rules are implemented, the response categories were expanded as follows: "Yes, clear rules that are enforced (1);" "Yes, general rules and they are monitored (2);" "Yes, there are rules, but child makes own choices (3);" and "No (there are no rules) (5)."

Non-Coresident Parent (Section D and Non-Coresident Parent Block). The Non-Coresident Parent modules are administered to PCGs where at least one biological or adoptive parent is not living in the child's household at the time of interview. Content in Section D includes whether the child has another adoptive parent, stepparent, or parent figure in the household; whether the non-coresident parent is still living, and if not, when the parent died; when the child and the parent last lived together, if ever; and how often the parent and child communicate and visit. Questions are asked separately for mothers and fathers.

The Non-Coresident Parent Block collects information from the PCG about the nonresident parent of each CDS child. That is, when two children have different non-coresident biological parents, the Non-Coresident Parent Block collects information on each parent separately. Content includes the parent's residential proximity, whether she or he has other children and/or

is married; whether the parent is currently in jail or prison; and the PCG's frequency of contact and conflict with the parent.

Home Environment (Section E). Section E includes information about children's access to learning resources and technology at home and about children's learning and social activities in the community. PCGs also report who paid for children's fee-based activities such as arts instruction, athletics, and tutoring.

Response options to spending items differ from CDS-2014 and CDS I–III due to backcoding on some open-ended responses and to keep response categories consistent with those used with similar items elsewhere in the instrument (e.g., item E14, "Who paid [CHILD NAME]'s tutoring programs? Include contributions from family members or friends living elsewhere.").

The Home Environment section also includes content on children's use of technology at home, including whether the child has their own electronic device or devices such as a computer, tablet, or cellular telephone or smartphone (E47A–E49B); frequency of activities such as homework and social interaction using electronic devices (E51–E55C); and recent help-seeking and help-giving associated with computer use at home (E57–E58).

Child Education (Section F and Schools Block). Section F collected information on the PCG's educational aspirations and expectations for the CDS child (F2–F3) and the CDS child's educational history and current status. Content includes whether the child attended an early intervention preschool program such as Head Start (F4–F8); age at kindergarten entry (F9–F12); attendance at public and private schools (F14–F19); attendance in classes for gifted students (F20); classification as requiring special education (F21–F22); suspensions and expulsions (F23); grade retention (F24–F24A); school dropout (F25–F26); home schooling (F33); participation in subsidized meal programs at school F27–F32); PCG involvement at the child's school (F34–F38); and PCG involvement with the child's education at home (F39–F41).

The universe for the items in this section varies depending on child age and grade. Question wording is identical to CDS-2019, CDS-2014 and CDS I–III. Response options to some items differ due to backcoding on some open-ended responses and to keep response categories consistent with those used with similar items elsewhere in the instrument (e.g., item F15C, "Who paid [CHILD's] private school expenses?").

Expenditures and Savings (Section G). Seven items measure frequency, amount, and conditions of children's receipt of an allowance (G1–G6). Eleven items measure family members' savings and investments on behalf of children, including savings for college (G7–G13, G20–G20A). Six items describe expectations about college expenses (G14–G19).

Parallel Blocks

School Attended. For children aged 5–18 years and Grade PreK–12, information is collected on the total number of schools ever attended by each child, including school where currently enrolled and the name and location of each school. This information is matched to the Common Core of Data and Private Schools Survey databases maintained by the National Center for Education Statistics of the US Department of Education. Numeric school identifiers are available to qualified researchers under a restricted-use data agreement. Visit the PSID web site for more information on restricted-use agreements.

Child School Closure & Attendance. For school year 2020–2021, this section collected information on school attendance, 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 (SCHCV1, SCHCV2, SCHCV3), as well as student attendance (SCHCV4, SCHCV5), parental involvement in school activities (SCHCV6), and time spent on learning activities (SCHCV7). Finally, PCGs reported on the learning quality for children during this period (SCHCV8). Two other items inquired about the beginning of the 2021–2022 school year, as well (SCHV9 and SCHCV10).

Child Care. For children in sixth grade and younger, PCGs describe arrangements in the past four weeks for all child care regularly provided by someone other than the PCG and his/her spouse or partner. This includes information on the type of arrangement (e.g., relative-based in-home care or childcare center), the number of days and hours a child is in care each week, and the cost of care.

Child Interview (Child File)

Race and Ethnicity (Section A). Adolescents aged 12 years and older self-reported their racial identity and ethnic origins or background (A1_1–A2D_1). The only change in the CDS-2021 Child file instrument is the addition of a fourth mention that children can select to describe their racial identity.

For confidentiality purposes, racial and ethnic origin/background categories endorsed by ten or fewer respondents are not included in public release data. Original coded responses are available to qualified researchers under a restricted-use data agreement. Visit the PSID website for more information on restricted-use agreements.

Ability Self-Concepts in Math and Reading (Section B). The ability self-concepts items (B2– B9) reflect two scales to self-assess ability in the domains of math (C21MATH) and reading (C21READ).

Academic performance (Section B). Adolescents reported grades earned in the most recent completed semester in mathematics and English (B10–B11), current cumulative grade point average (B12), and grades earned in eighth grade (B14).

Future plans (Section B). Adolescents (11th grade and higher, including high school leavers or graduates) described their aspirations and plans for college attendance and information about college provided by their high school; and plans to serve in the armed forces (B15–B32A).

Health (Section C). This section covers questions on general health status, depression, and physical development.

General health (Section C). Adolescents reported on their general health status (C1); perceived weight status (C2); recent efforts to change or maintain weight through diet or exercise (C3–C5); and general emotional health (C6–C7).

Depression (Section C). Adolescents completed the Children's Depression Inventory (CDI) Short Form (C8–C17). The CDI (C21CDI) is an assessment that rates the severity of symptoms related to depression or dysthymic disorder in children and adolescents.

To protect respondent privacy, interviewers directed adolescents to read the response options for each item to themselves in their response booklets and to provide the numeric code corresponding to the statement that best describes their feelings during the last two weeks. The interviewer presented the response options aloud only where the respondent did not have his or her response booklet available.

Social Relationships (Section D). Adolescents described how close they feel to parents, stepparents, friends, siblings, teachers, and other adults (D1A–D1F, D1I–D2); help to siblings (D1G–D1H); friends' positive and negative behaviors (D3A–D3M); and characteristics of and attachments to family pets (D4–D12). (C21CHPET).

Personality and Behavior (Section E). This section includes modules on children's self-esteem, perseverance, and peer problems.

Self-Esteem (Section E). Adolescents responded to a five-item version of the Rosenberg Self-Esteem scale (C21RSBGSE) that excluded negatively worded items (E1–E5). Four of the items appeared in the original Rosenberg scale. A fifth item, "I feel good about myself," is a positively worded version of the statement "At times I think I am no good at all."

Perseverance (Section E). Adolescents responded to a five-item scale measuring perseverance (E6–E10).

Employment (Section F). Adolescents described current and past summer employment, including occupation, industry, tenure, work hours, wages, and job satisfaction (F1A–F21PER) and job aspirations (F21_1–F25).

Computers and Electronic Media Use (Section G). For adolescents, the questions collected information on their own electronic devices (G1–G3), internet access (G6–G7), computer/electronic device use for schoolwork, information-seeking, social interaction, and entertainment (G8–G21).

A six-item web-use skills index (C21CWBCM) adapted from Hargittai and Hsieh measured the adolescent's familiarity with computer and internet-related terminology (G23A–G23F). One item assesses confidence in understanding new terms and words related to computers and the internet (G22).

Two items measure the exchange of assistance with computers or other electronic devices between the adolescent and their PCG in the past 30 days (G24–G25).

Financial Behavior (Section H). Three items collect information about the frequency, amount, and conditions of an adolescent's allowance (H1–H3). Six items address the amount and intended purpose of an adolescent's own financial savings (H4–H9). Adolescents report on their own and their PCG's past-year charitable donations (H10, H11).

Sensitive Topics (Section J). Questions on sensitive topics were administered to CDS adolescents using interactive voice response (IVR) technology in order to ensure respondent privacy and minimize response bias during the telephone interview.

Some content from Section J of the Child Interview is available only through a special data application process. This content is denoted as sensitive data below. Please visit the PSID website for more information on how to obtain access to sensitive data.

The introduction to the IVR interview included items on right- or left-handedness, self-reported age, and availability of the response booklet (J0A–J0C).

Peer Victimization and Bullying. Four items (J1A–J1D) address peer victimization and bullying. These items were drawn from work by Kochenderfer and Ladd. A bullying scale (C21BULLY) is available.

Dating. Four items pertain to adolescents' experience with dating (J2–J4). These items were drawn from the National Longitudinal Study of Adolescent to Adult Health.

Physical development. Adolescents reported on the onset of puberty, including physical appearance relative to age peers; breast development and age at menarche for girls; and facial hair growth and voice changing for boys. These items were drawn from Pubertal Development Scale (PDS). Adolescents responded to these items as part of the IVR interview. The variables in the public-use data file are JC18–JC25.

Sexual Health and Activity (Sensitive Data). Adolescents reported on age at first sexual intercourse (J5–J7B), recent frequency (J8–J9), and lifetime number of partners (J10). All sexually active respondents report on frequency of condom use (J11); female respondents also report on use of birth control pills (J12–J13C). All respondents report on whether they have ever been tested for or diagnosed with a sexually transmitted infection (J14–J16A). Male and female adolescents responded to items about pregnancy experience (becoming pregnant or impregnating someone else), frequency, and outcomes (J16–J21). Adolescents also reported their sex at birth (C18A), sexual orientation (J21A), and gender identity (J21B).

Risky Behavior (Sensitive Data). Adolescents reported on the frequency of behaviors including staying out past curfew, physically harming others, damaging property, bringing a weapon or drugs or alcohol to school, and truancy (J22–J34). The series also included questions about contact with law enforcement, including being stopped and questioned or arrested (J28–J29).

Tobacco, Drug, and Alcohol Use (Sensitive Data). Adolescents reported lifetime and past 30-day use of tobacco products, electronic cigarettes, alcohol, marijuana, inhalants, hallucinogens, prescription drugs taken without a prescription from a doctor, amphetamines, and tranquilizers (J35–J64A). In addition, respondents reported on frequency of heavy drinking, type of alcohol most often consumed, and frequency of driving while intoxicated or riding with an intoxicated driver. Items were originally drawn from the National Longitudinal Study of Adolescent to Adult Health. Content updates were made in consultation with investigators from Monitoring the Future, an annual study of middle and high school students designed to track trends in adolescent substance use.

IVR Interview Experience. Adolescents responded to three items about the accuracy of their responses and the ease or difficulty of completing the IVR interview (J65–J67).

CHAPTER 3. THE CDS-2021 SAMPLE

The CDS-2021 sample—and the PSID sample more generally—was designed to be representative of the corresponding US population of children and families. By design, PSID and CDS-2021 have certain gaps in coverage. For CDS-2021, a new eligibility criterion was participation in the previous full wave of CDS in 2019.

The CDS-2021 sample eligibility criteria were defined as follows:

- Family participated in the CDS-2019 survey.
- Family participated in the 2021 Core PSID survey.
- Child's reported birth year was 2004–2018.
- Child was classified as belonging to the PSID sample (i.e., has the "PSID gene").
- Child was not classified as a reference person or spouse/partner of a family unit.
- Not living outside of US in both Core 2019 and Core 2021.

In addition, if there were more than three eligible children in a family unit (based on sharing the same PCG), only three children were selected randomly for CDS-2021.

A total of 4,629 children participate in CDS-2019. Of the 4,629 children, 4,154 were projected to be age-eligible for CDS-2021. The final eligible CDS-2021 sample comprised 3,637 children. A total of 2,590 children participated in CDS-2021. Table 3.1 summarizes the final contact and interview dispositions in CDS-2021.

Table 3.1 CDS-2021 Sample Disposition from projected age eligible children

CDS-2021 outcome	Count
Completed CDS-2021	2,590
Non-response	
Non-response in PSID Core-2021	494
Non-response in CDS-2021	975
Child is not followable	1
Non-sample	
Move out on own	2
Deceased	1
Living outside US in both Core 2019 and Core 2021	19
Sub-selected out because PCG has more than 3 children	72
Total	**4,154**

The overall, unconditional response rate at the child level for CDS-2021 was (2590/(2590+494+975+1)) = 63.8 percent.

Indicators for CDS-2021 fieldwork outcomes are available in the PSID Online Data Center. The individual file variable ER35060 classifies interview outcomes for 3,643 cases. This variable captures the full range of fieldwork outcomes for why an otherwise eligible child did not participate in CDS-2021.

Children in CDS-2021 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-2021 fieldwork (children were selected based on their age in 2021 to be age 2 to 17 years, but interviews occurred in 2022 when some of the oldest children reached age 18 years). Table 3.2 also shows that the CDS-2021 sample was divided approximately evenly between males and females.

Table 3.2. Age and Sex of Children in CDS-2021

Birth year	Males	Females	Total	Percent
2004	75	93	168	6.49
2005	86	86	172	6.64
2006	90	86	176	6.80
2007	76	84	160	6.18
2008	67	89	156	6.02
2009	99	99	198	7.64
2010	93	76	169	6.53
2011	86	80	166	6.41
2012	84	106	190	7.34
2013	80	80	160	6.18
2014	89	82	171	6.60
2015	98	87	185	7.14
2016	95	82	177	6.83
2017	88	90	178	6.87
2018	83	81	164	6.33
Total	1,289	1,301	2,590	100.00

CHAPTER 4. THE CDS-2021 DATA FILE STRUCTURE

The CDS-2021 data package includes the following files:

1. 2021 Demographic File (one record per child, N=2,590)
2. 2021 Primary Caregiver Child Interview File (one record per PCG child interview, N=2,590)
3. 2021 Primary Caregiver Household Interview File (one record per interviewed primary caregiver, N=1,576)
4. 2021 Household Roster File (one record per household member listed in CDS-2021 roster, N=6,201)
5. Cumulative CDS ID Map File 1997–2021 (one unique record across all waves per CDS-selected child, primary caregiver, or other caregiver, N=19,686)

Table 4.1 summarizes these files according to the CDS-2021 individual for whom data are available and lists the number of records in each component/file.

Table 4.1. CDS-2021 Study Component Completion by Individual Sample Member Type

Individual	CDS-2021 file				
	DEMOG	PCG-CHILD	CHILD	PCG-HH	HHROSTER
Child					
Age 0–2 years	X	X			X
Age 3–4 years	X	X			X
Age 5–7 years	X	X			X
Age 8–11 years	X	X			X
Age 12–18 years	X	X	X		X
PCG				X	X
Other HH members					X
Num. of records					
Total	2,590	2,590	749	1,576	6,201
With associated PCG/child record		2,590		1,572	

Primary Caregiver Household Interview (PCGHH2021)

The Primary Caregiver Household Interview is provided at the primary caregiver (PCG) level, with one record per interviewed PCG (N=1,576). Note that there are some CDS-2021 households that completed the Primary Caregiver Household Interview but did not complete the Primary Caregiver Child Interview for any child (N=4 primary caregivers). As a result, merged file content between the PCGHH2021 interview file and PCGCHILD2021 will not yield a file with complete data for all fields for all individuals.

Primary Caregiver Child Interview (PCGCHILD2021)

The Primary Caregiver Child Interview is provided at the child level (N=2,590). The data file includes records for all children aged 2–18 years for whom a primary caregiver provided a Primary Caregiver Child interview.

Household Roster (HHROSTER2021)

The Household Roster File includes one record for each person residing in a CDS-2021 household at the time of the completed coverscreen interview (N=6,201).

The roster file includes three sets of unique identifiers for each person: a CDS-2021 family household identifier (R21CDHID) and roster position (R21INST); a PSID 2021 Core interview family unit identifier (R21YRID) and sequence number (R21CYPSN); and a time-invariant family lineage identifier (R21ID68) and person number (R21PN). These unique identifiers may be used to merge together individual-level content files within CDS-2021 or between CDS-2021 and CDS-2019 or other components of the PSID suite of studies, using the merging instructions provided below.

Demographic Data (DEMOG2021)

The Demographic Data File is provided at the child level (one record per child, N=2,590). Records are included for all children who have a record on either the Primary Caregiver Child Interview (PCGCHILD2021), the Child Interview (CHILD2021), or both. This file may be used to link children's records from CDS-2021 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-2021,
- Children's unique identifiers in the 2021 PSID Core interview,
- For each CDS-2021 data file, an indicator of whether a record for the child is included,
- CDS-2021 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).

Cumulative ID Map (CDSIND2021)

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=19,686 as of CDS-2021). 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 CDSIND2021 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**

Values on the caregiver identifier variables are set to “0” on caregivers’ own records.

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

- DEMOG_**, PCGCH_**, CHILD_**), PCGHH_** (primary caregiver-level).

CDSIND2021 also includes an indicator of which type of record the individual contributed at each CDS wave (CDSTYPE**). CDS-2021 includes the following records types (CDSTYPE21):

0. Not a selected CDS child or designated primary caregiver in CDS-2021 (N=14,473)
1. Participating child (N=2,590)
2. Participating primary caregiver (N=1.576)
4. Non-participating selected child (N=975)
7. CDS eligible child in eligible household with greater than 3 eligible children, not selected (N=72)

The variable PCGHHNO21 indicates whether a CDS-2021 child's caregiver was the first or second designated primary caregiver residing in a household.

File Merging

Users may wish to combine information from multiple components of CDS-2021 into a single data file or to incorporate information about children and their families from the PSID Core (main) interview, earlier waves of CDS, or other PSID studies. The PSID Online Data Center and CDS Online Data Center 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-2021 packaged data will need to merge records across files using the following guidance.

Table 4.2 Unique Identifiers in CDS-2021
(Note: When Two variables are Listed in a Cell, the Variables are Jointly Unique)

		2021 PSID Core interview (Family ID, person ID)		CDS-2021 (Family ID, Person ID)			Fixed (time-invariant) (Family ID, Person ID)	
	Record unit (Ego)	**Ego**	**Caregiver to Ego**	**Ego**	**Caregiver to Ego**	**Caregiver number**	**Ego**	**Caregiver to Ego**
DEMOG2021	Child	X21YRID, X21CYPSN						
PCGCH2021	Child	P21YRID, P21CYPSN						
CHILD2021	Child	C21YRID C21CYPSN						
PCGHH2021	Primary caregiver	H21YRID H21CYPSN		H21CDSHID, H21INST		H21PCGHH		
HHROSTER2021	Household member	R21YRID R21CYPSN		R21CDSHID, R21INST	R21CDSHID, R21CDSHPIN	R21PCGHH	R21ID68, R21PN	
CDSIND2021	CDS sample (selected children & designated caregivers)	CRFID21 CRSN21	CRPCGFID2 CRPCGSN21	CDS_HID21, CDS_SN21	CDS_HID21 CDSPCGSN21	PCGHHNO21	CDSCUMID68 CDSCUMPN	ID68PCG21 PNPCG21

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-2021.

See Table 4.3 for the expected number of matched records for selected file merges.

Table 4.3. Number of Matched Records between Merged Files (N=Appears on File 1 and File 2 / N=Appears on File 1 only / N=Appears on File 2 only)

File1/File 2	PCGHH2021	PCGCH2021	CHILD2021
CHILD2021	749 / 0 / 1,845	749 / 0 / 1,841	-
PCGCH2021	2,590 / 0 / 4	-	749 / 1,841 / 0
DEMOG2021	2,590 / 0 / 4	2,590 / 0 / 0	749 / 1,841 / 0

*The first value in each cell describes the number of records on File 1 affiliated with a primary caregiver who appears on PCGHH2021. These records are matched using the CDS household identifier and primary caregiver number as shown in Table 4.2.

Merging Data Files within CDS-2021

Merging Child-Level Records

In the PSID Online Data Center and CDS Online Data Center, a user may include variables from multiple CDS-2021 child-level files in a data cart. The Data Centers will deliver a downloadable data file with all child-level records already merged across files.

Alternatively, a user may wish to use CDS-2021 packaged data, which includes separate data files for each study component. Each of these data files includes a pair of identifier variables that, in combination, uniquely identify each CDS child. The identifiers are drawn from the 2021 PSID Core interview and refer to the child's family interview identifier (ER34901) and sequence number (roster position in the family listing—ER34902)) from that interview. In each file, the family interview identifier includes the root YRID in the variable name and the variable name for the sequence number includes the root CYPSN. On each file, the prefix (three leading characters) is unique to the file and survey interview year. Refer to the column labeled 2021 PSID Core interview/Ego in Table 4.2 for a complete set of variable names.

To merge child-level records together across files, change variable names as needed so that they match between the two files. Then merge records together, being sure to use both the family interview ID and sequence number in order to uniquely identify children.

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

The guidance below is directed to users working with CDS-2021 packaged data or who use the PSID Online Data Center to create a data extract containing variables from the CDS-2021 PCG-Household and PCG-Child Interviews. Users who create a data extract containing variables from both files in the CDS Online Data Center 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.

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 (CDS2021IND) as a bridge between files. The Cumulative ID Map can be downloaded from the PSID packaged data page, and variables from the Cumulative ID Map can be added to data carts via the PSID Online Data Center and CDS Online Data Center.

Use the following steps:

1. Conduct a one-to-one merge between the child-level file and CDSIND2021 using the two child ID variables (CRFID21 and CRSN21 in CDSIND2021; [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 so that they will match as needed. This will merge the CDS household interview ID (CDS_HID21), the PCG household number (PCGHHNO21) and PCG PSID family identifiers (CRPCGFID21 and CRPCGSN21) to the child-level file.

2. Conduct a one-to-many merge between the Primary Caregiver Household Interview file (PCGHH2021) and CDSIND2021 using the CDS household interview number and PCG household number (CDS_HID21 and PCGHHNO21 in the CDS2021IND file and H21CDSHID and H21PCGHH in the PCGHH2021 file) 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 CRFID21 and CRSN21 to merge the files created in steps 1 and 2. Users may wish to remove records which are in either the Primary Caregiver Household Interview file or child-level data file but not in the other (e.g., PCGHH2021 records that do not have a corresponding child record or vice versa).

Merging CDS-2021 Records to other PSID Studies

Merging Individuals' Records

Users may merge a CDS-2021 content file directly to records from the PSID 2021 Core interview by using the Core 2021 family interview number and sequence number included on all CDS-2021 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 ER34901 (family interview number, equivalent to [z]YRID in CDS-2021) and ER34902 (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 (HHROSTER2021, variables R21ID68 and R21PN) and the Cumulative ID Map (CDSIND2021, variables CDSCUMID68 and CDSCUMPN). The Household Roster includes records for all CDS-2021 household members. The Cumulative ID Map includes records for children who were selected for CDS-2021 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 (ID68PCG21 and PNPCG21) on the Cumulative ID Map (CDSIND2021) 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 four primary caregivers who were present in CDS-2021 but did not complete the PCG-Child interview. Other records will appear only on the content file. This includes all individuals who were not the primary caregiver to a child in CDS-2021. 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. In the data cart, include at least one child-level variable from CDS-2021. 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.

CHAPTER 5. THE CDS-2021 WEIGHTS

CDS-2021 includes cross-sectional and longitudinal weights. This chapter describes the construction and use of the CDS-2021 weights. We recommend that researchers use an appropriate CDS-2021 weight with all of their analyses.

Cross-Sectional Weights

The CDS-2021 Child Cross-Sectional Weight (X21CHWGT) allows researchers to generalize their statistical results to the US national population of children aged 2–17 years in 2021. Note that the CDS-2021 sample was conditioned on participation in CDS-2019—and the weights are designed to project backwards to the full CDS-2019 sample who are still age-eligible in 2021.

The Child Cross-Sectional 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-2021 Child Cross-Sectional Weight was derived from the CDS-2019 Child Cross-Sectional Weight, which incorporates differential patterns of non-response and corrects potential under-coverage of some demographic subgroups using post-stratification. The CDS-2021 Child Cross-Sectional Weight also accounts for differential non-response in 2021 Core PSID and in CDS-2021.

The CDS-2021 Child Cross-Sectional Weight (X21CHWGT) is provided for 2,590 children. X21CHWGT should be used when analyzing CDS-2021 outcomes and when comparing survey outcomes collected in CDS across both the 2019 and 2021 waves.

The CDS-2021 Child Cross-Sectional Weight should be used for all child-level analyses based on the interview data from CDS-2021 that are undertaken with one observation for each child in the sample.

The PCG Weight (H21PCGWGT) was directly derived from the Child Cross-Sectional Weight. It has one value for each of the 1,576 PCGs in the CDS-2021 sample, including 4 PCGs who completed the CDS-2021 PCG Household Interview but for whom no corresponding child-level data was collected in CDS-2021.

Table 5.1 summarizes for CDS-2021 the Child Cross-Sectional Weight and the PCG Weight.

Table 5.1 Use of CDS-2021 Weights for Analyzing Interview Data

Analysis sample	Recommended weight	Cases
Child-level interview data	X21CHWGT	2,590
PCG or household interview data	H21PCGWGT	1,576

Method to Construct the Child Cross-Sectional Weight

The CDS-2021 Child Cross-Sectional Weight (X21CHWGT) was based on the CDS-2019 Child Cross-Sectional Weight, the construction of which is described in the CDS-2019 User Guide.[3] The CDS-2021 Child Cross-Sectional Weight incorporated an adjustment for non-response

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

based on a regression model that predicted which cases from CDS-2019 completed the 2021 Core PSID and CDS-2021 interview using a comprehensive set of covariates. The non-response adjusted weight was then post-stratified to 2021 population totals.

Step 1. Within PCG Child Selection Adjustment

When a PCG had more than three children, three of these children them were randomly selected for CDS-2021 and a within-PCG child selection adjustment factor, the inverse of the selection probability, was applied to the selected children's weights. The children who were not selected were coded as non-sample and excluded from data collection and from the weight construction. Children whose PCG has three or fewer children were assigned a value of "1" for the within-PCG child selection adjustment factor.

Step 2. 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-2021 child cases (n=4,060) 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. 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-2021 Child Cross-Sectional 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-2021 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.

Research by Little and Vartivarian[4] recommends adjusting for design variables in non-response weights. The product of each child's CDS-2019 weight and within-PCG child selection adjustment factor incorporates unequal probability of selection in the sample design and was added to the CDS-2021 response propensity model to adjust for design variables as part of the CDS-2021 non-response weighting.

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-2021 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-2021 was higher among children in families with an older reference person, from the SEO sample, and in the North Central region. The probability of response was lower for children in families with a less-

[4] Little, R.J. and Vartivarian, S. (2003). On weighting the rates in non-response weights. Statistics in Medicine, 22, 1589–1599.

educated reference person and with more children and for children with no behavior problems and with some distress. Although a number of variables in the model are not statistically significant predictors of CDS-2021 response, all variables were retained in the model used to derive the predicted probabilities of response. Overall, the Hosmer-Lemeshow test of goodness of fit test (χ^2=18.01, 8 df, p=0.021) suggests lack-of-fit. Although we would prefer to have a better model fit, because the objective of this propensity model is to provide predictions of the response propensities rather than explaining nonresponse, this lack-of-fit suggested by the Hosmer-Lemeshow test is not a particularly concerning for our nonresponse weighting purposes.

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-2021 child response case was assigned a non-response adjustment factor equal to the inverse of the median predicted probability of successful completion of the CDS-2021 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-2021 Child Cross-Sectional Weight

Response propensity decile	Median response propensity	Adjustment factor
1	0.465	2.15
2	0.542	1.846
3	0.580	1.724
4	0.611	1.637
5	0.637	1.569
6	0.660	1.515
7	0.682	1.466
8	0.705	1.418
9	0.731	1.368
10	0.770	1.298

The probability of selection weight for each CDS-2021 child response case was then multiplied by the non-response adjustment factor and the within PCG child selection adjustment factor, described in Step 1, to produce an interim weight that adjusts for the non-response to the CDS-2021 interview.

Step 3. Non-US Cases

There were 10 children in CDS-2021 with interview data that resided outside the US 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 US resident population. At this step, for the non-US cases, the Child Cross-Sectional Weight is designated to be complete.

Step 4. 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 Cross-Sectional Weight from Step 2, 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 5. Post-Stratification to Population Control Totals

We next post-stratified the trimmed, attrition-adjusted weights from Step 4 to population control totals from the 2021 American Community Survey (ACS). Post-stratification cells were formed based on the following respondent characteristics:

- Child sex (male/female),
- Birth year of child (2004–2018),
- Child race/ethnicity (Hispanic, non-Hispanic Black, non-Hispanic White, or other), and
- 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 10–20 individuals in each cell. The post-stratification adjustment factors were computed as the ratio of the ACS control totals to the CDS-2021 weighted population estimate (using the interim weight from Step 4). Appendix Table A.2 shows the CDS-2021 sample count, CDS-2021 weighted estimates, the ACS population estimates, and the post-stratification adjustment factors for each of the 91 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 6. Combining the US and Non-US Cases

The final step in creating the Child Cross-Sectional Weight was to combine the weights from Step 3 for non-US cases with the weights from Step 5 for cases in the US.

Method to Construct the PCG Weight

The CDS-2021 PCG Weight (H21PCGWGT) was derived entirely from the CDS-2021 Child Cross-Sectional Weight. In particular, PCGs were assigned the average CDS-2021 Child Cross-Sectional Weight over all children for whom they were the responsible primary caregiver. For PCGs with no corresponding children in the sample (because no child interview components were completed and hence no Child Cross-Sectional Weight was constructed), a PCG weight was calculated based on imputed values for the missing child cross-sectional weights.

Child Longitudinal Weight

All children in CDS-2021 participated in CDS-2019. The CDS-2021 child sample also includes children who participated in CDS-2014. These sample children were all born between 2004 and 2013, and hence were aged 1–10 years in CDS-2014, aged 6–15 years in CDS-2019, and aged 8–17 years in CDS-2021. To support longitudinal analysis of CDS children who participated in CDS-2014 and in both CDS-2019 and CDS-2021, we provide a longitudinal child weight. This weight accounts for differential probabilities of selection due to the original PSID sample design and subsequent attrition.

The CDS-2021 Child Longitudinal Weight (X21LONGWGT) is provided for 1,296 children and is designed for analyses of outcomes for children who participated in CDS-2014, CDS-2019, and CDS-2021. The construction of this CDS longitudinal weight is described in this section.

Sample Transition from CDS-2014 to CDS-2021

Of the 4,333 children who participated in CDS-2014, 2,779 were projected to be eligible for participation in CDS-2021. Table 5.3 summarizes the CDS-2021 fieldwork outcomes for these 2,779 children. In CDS-2021, data were collected on a total of 1,296 of these age-eligible children, representing an unweighted response rate of (1,296 / (2,779 – 51)) = 47.5%. The projected eligible sample excludes a total of 51 children who died (n=2), were reclassified as non-sample (n=3), were institutionalized (n=2), resided oversea in both 2019 and 2021 (n=7), moved out on their own (n=2), or were not selected for PCGs with more than three children. Children were classified as non-response (n=1,432) because their family ended participation in PSID prior to the start of CDS-2019 (n=233), because they did not participate in CDS-2019 for any reason (n=567), because they participated in CDS-2019 but not in PSID-2021 (n=219), or because they participated in CDS-2019 but not in CDS-2021 (n=413).

Table 5.3. CDS-2021 Fieldwork Outcomes for Age-Eligible Children from CDS-2014

CDS-2021 outcome	Count
Child data collected in CDS-2021	1,296
Non-response of family for Core PSID by or in 2019	233
Non-response for child in CDS-2019	567
Non-response for child in PSID-2021 but response in CDS-2019	219
Non-response for child in CDS-2021 but response in CDS-2019	413
Total non-response = 1,432	
Child reclassified as non-sample	3
Child deceased	2
Child institutionalized	2
Living oversea both 2019 and 2021	7
Move out on their own	2
Not selected from the PCG with >3 children	35
Total non-sample =51	
Total	2,779

Construction of CDS-2021 Child Longitudinal Weight

The CDS-2021 Child Longitudinal Weight is the product of the CDS-2014 Child Weight (X14CHWGT), CDS-2021 within PCG child selection adjustment factor (described in Step 1 of constructing cross-sectional weight), and an attrition adjustment factor. Because the completion of CDS-2021 is conditional on response in CDS-2019, we can decompose the attrition adjustment factor into the product of two components: the 2014 to 2019 attrition adjustment factor and the 2019 to 2021 attrition adjustment factor. The CDS-2021 Child Longitudinal Weight can thus be constructed by applying the product of CDS-2021 within PCG child selection adjustment factor and these two adjustment factors to the CDS-2014 Child Weight:

CDS-2021 Child Longitudinal Weight = CDS-2014 Child Weight x CDS-2014 to CDS-2019 Attrition Adjustment Factor x CDS-2019 to CDS-2021 Attrition Adjustment Factor x CDS-2021 within PCG Child Selection Adjustment Factor
(1)

Because the CDS-2019 Child Longitudinal Weight is the product of the CDS-2014 Child Weight and the attrition adjustment factor between CDS-2014 and CDS-2019, Equation (1) is equivalent to:

CDS-2021 Child Longitudinal Weight = CDS-2019 Child Longitudinal Weight x CDS-2019 to CDS-2021 Attrition Adjustment Factor x CDS-2021 within PCG Child Selection Adjustment Factor (2)

The CDS-2021 Child Longitudinal Weight was thus constructed by applying the CDS-2019 to CDS-2021 attrition adjustment factor and CDS-2021 within PCG child selection adjustment factor to the CDS-2019 Child Longitudinal Weight (X19LONGWGT). The attrition adjustment factor that accounts for non-response between CDS-2019 and CDS-2021 was produced for the CDS-2021 Child Cross-Sectional Weight and is described in Table 5.2.

To examine the properties of the CDS-2021 Child Longitudinal Weight, we compared weighted estimates for selected demographic, geographic, and socioeconomic variables in the CDS-2014 data using two approaches. The first was based on the CDS-2014 sub-sample that remained eligible for CDS-2021 and used the CDS-2014 Child Weight. The second approach was based on CDS-2021 panel response cases and used the CDS-2021 Child Longitudinal Weight. The results are presented in Table 5.4, and show that the distributions of the selected characteristics are similar across the two approaches. This suggests that the attrition adjustment for the CDS-2019 Child Longitudinal Weight compensates for potential attrition bias—at least for the variables included in this comparison. Note, however, that this comparison does not necessarily rule out the possibility of selection bias associated with other characteristics of the respondents.

Table 5.4. Comparison of Estimates Using: (1) the Full CDS-2014 Sample and the CDS-2014 Child Weight and (2) CDS-2014 Data for CDS-2021 Participants and their CDS-2021 Child Longitudinal Weight

Characteristic from CDS-2014 or 2013 PSID*	Value	Estimate using CDS-2014 data and CDS-2014 Child Weight		Estimate using CDS-2014 data for CDS-2021 participants and their CDS-2021 Child Longitudinal Weight		Ratio (2)/(4)
		Column 1 (N)	Column 2 (percent)	Column 3 (N)	Column 4 (percent)	Column 2/4
Region	Northeast	277	13.82	127	13.45	1.03
	North Central	732	23.92	361	24.73	0.97
	South	1206	38.09	562	38.80	0.98
	West	540	23.91	245	22.98	1.04
	Outside of US	8	0.26	1	0.03	9.23
Immigrant sample	Non-immigrant	2517	83.18	1184	83.12	1.00
	Immigrant	246	16.82	112	16.88	1.00
Metropolitan Statistical Area	MSA	2108	76.08	1042	79.94	0.95
	Non-MSA	647	23.66	253	20.04	1.18
	Outside of US	8	0.26	1	0.03	9.23
Child birth year	2004-2005	583	23.98	275	23.12	1.04
	2006-2009	1180	43.20	554	44.05	0.98
	2010-2013	1000	32.82	467	32.83	1.00
Child sex	Female	1425	48.59	677	47.56	1.02
	Male	1338	51.41	619	52.44	0.98
Race/ethnicity of child	Hispanic	363	24.22	167	24.78	0.98
	Non-Hispanic Black	1097	15.96	548	16.61	0.96
	Non-Hispanic White	1235	56.57	549	55.22	1.02
	Non-Hispanic Other	68	3.25	32	3.40	0.96
Education of reference person (RP)	Education unknown	36	1.20	9	0.79	1.53
	No high school diploma	511	16.69	243	20.01	0.83
	High school diploma only	742	25.07	318	24.50	1.02
	Some college	759	26.01	364	25.67	1.01
	College or more	715	31.03	362	29.04	1.07
Age of RP	30 or younger	1018	28.17	449	28.25	1.00
	31–45	1488	58.57	707	57.12	1.03
	46 or older	257	13.27	140	14.63	0.91
Sex of RP	Female	815	18.70	399	19.01	0.98
	Male	1948	81.30	897	80.99	1.00
Employment of RP	Unemployed	262	7.99	120	7.85	1.02
	Employed	2501	92.01	1176	92.15	1.00
Number of children in family	1	607	21.27	312	19.80	1.07
	2	1,024	38.05	491	37.84	1.01
	3	691	25.70	324	26.56	0.97
	4+	441	14.98	169	15.81	0.95
Total		2763	100.00	1296	100.00	1.00

Note: *characteristics of the reference person (RP) and household/family were collected in the 2013 Core PSID interview.

Summary of Weights

In Table 5.5 we list the CDS-2021 child weights and the PCG weight and we present case counts and summary statistics. The estimated US population of children born from 2004 to 2018 and not living in institutional group quarters for 2021 (the target year for the CDS-2021 sample) is 62,501,888. The sum of the Child Cross-Sectional Weight is slightly higher because of the inclusion of the respondents residing outside of the US. The sum of the CDS-2021 Child

Longitudinal Weight, 37,480,131, is close to the weighted total population in CDS-2014 of children born from 2004 to 2013 of 35,634,801. The weighted total population of PCGs in CDS-2021 is 37 million.

Table 5.5. Summary of CDS-2021 Weights

Weight type (variable name)	Count	1st pct.	50th pct.	99th pct.	Mean	Std. dev.	Coef. var.	Sum total
Child Cross-Sectional Weight X21CHWGT	2,590	734.05	20,221.69	100,053.78	24,217.66	22,161.75	91.51	62,723,739
Child Longi-tudinal Weight X21LONGWGT	1,296	999.74	23,481.49	118,733.37	28,919.85	27,568.38	95.33	37,480,131
PCG Weight H21PCGWGT	1,576	745.22	20,205.30	97,675.28	23,658.70	21,538.90	91.04	37,286,107

Recommendations for Using the Weights

In this section, we summarize our recommendations for using the CDS-2021 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-2021 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-2021 comprises a subset of the Core PSID sample, users may encounter instances where a cluster includes a single observation when analyzing the CDS-2021 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), 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, R, and Stata).

Child Cross-Sectional Weight (X21CHWGT)

This weight should be used for all cross-sectional analyses of child data from CDS-2021 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-2021.

Child Longitudinal Weight (X21LONGWGT)

The Child Longitudinal Weight is designed for panel analyses of child-level data between CDS waves in 2014 and 2021. For example, this weight should be used when analyzing the change

from CDS-2014 to CDS-2021 in the PCG-Child interview data. Such analyses can also incorporate data from CDS-2019 using this same Child Longitudinal Weight.

PCG Weight (H21PCGWGT)

This weight should be used for all analyses of in which the CDS-2021 PCG or household are the focus of the analysis and the unit of analysis is the PCG or the household. This is the weight to use for analyses of outcome variables from the PCG Household Interview.

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-2021 Child Response

Variable	Estimate	Std. err.	P-value	Significance
PSID sample component			0.019	*
SRC sample (ref.)	.	.	.	
SEO sample	-0.084	0.160	0.601	
1997/1999 new immigrant sample	0.317	0.169	0.060	
2017/2019 new immigrant sample	0.282	0.131	0.031	*
Child is male (0/1)	0.040	0.067	0.554	
Child birth year			0.543	
2004–2009 (ref.)	.	.	.	
2010–2014	0.089	0.081	0.274	
2015–2018	0.031	0.109	0.780	
Child race/ethnicity			0.626	
Non-Hispanic White (ref.)	.	.	.	
Hispanic	-0.131	0.136	0.334	
Non-Hispanic Black	-0.150	0.130	0.248	
Non-Hispanic Other	0.022	0.293	0.941	
Age of household reference person			0.003	**
≤30 years	-0.155	0.099	0.118	
31–45 years (ref.)	.	.	.	
≥46 years	0.295	0.106	0.005	**
Household reference person is male (0/1)	0.037	0.093	0.688	
Education of household reference person			0.001	**
≤11 years	-0.407	0.122	0.001	***
12 years	-0.460	0.105	<.0001	***
13–15 years	-0.244	0.100	0.015	*
≥16 years (ref.)	.	.	.	
Education unknown	-0.631	0.296	0.033	*
<25 years old	-0.376	0.228	0.099	
Household reference person is employed (0/1)	0.170	0.140	0.225	
Family income quartile			0.660	
1st quartile	0.069	0.133	0.602	
2nd quartile	0.136	0.113	0.228	
3rd quartile	0.057	0.101	0.572	
4th quartile (ref)	.	.	.	
Region			0.148	
South (ref.)	.	.	.	
Northeast	-0.040	0.120	0.736	
North Central	0.205	0.087	0.018	*
West	0.085	0.102	0.403	
Outside US	0.308	0.520	0.554	
Metro area (0/1)	0.183	0.097	0.058	
Number of children in the family unit			<.0001	***
1	0.089	0.098	0.362	
2 (ref.)	.	.	.	
3	-0.099	0.085	0.243	
4+	-0.561	0.103	<.0001	***

Variable	Estimate	Std. err.	P-value	Significance
The product of CDS-2019 Child Cross-				
Sectional Weight and CDS-2021 within-	<.0001	<.0001	0.705	
PCG child selection adjustment factor				
Child behavioral problems index (BPI, 0–27)			0.003	**
0	-0.367	0.117	0.002	**
1–3 (ref.)	.	.	.	
4–6	0.090	0.110	0.412	
7–10	0.045	0.111	0.688	
11–15	0.160	0.126	0.203	
16+	0.092	0.149	0.537	
Child age ≤2 years	0.095	0.144	0.509	
Safety of local neighborhood			0.179	
Completely safe	-0.095	0.076	0.213	
Fairly safe (ref.)	.	.	.	
Somewhat dangerous	-0.069	0.110	0.531	
Extremely dangerous	0.413	0.254	0.103	
Child health status			0.604	
Excellent (ref.)	.	.	.	
Very good/good/fair/poor	-0.037	0.070	0.604	
Family food security status			0.752	
High food security	.	.	.	
Marginal food security	-0.011	0.112	0.925	
Low food security	-0.095	0.118	0.424	
Very low food security	0.089	0.154	0.565	
PCG psychological distress scale (0–24)			0.247	
0 (ref.)	.	.	.	
1	-0.326	0.140	0.020	*
2	-0.161	0.133	0.225	
3	-0.315	0.132	0.017	*
4	-0.134	0.138	0.334	
5	-0.267	0.147	0.069	
6	-0.088	0.152	0.562	
7-9	-0.138	0.142	0.333	
10+	-0.069	0.161	0.669	
Hosmer and Lemeshow goodness-of-fit test	18.0109 (8 df) p=0.0211			
Note: *p≤0.05, **p≤0.01, ***p≤0.001; N=4,060 (response=2,590, nonresponse=1,470).				

Table A.2. Post-Stratification Cells for CDS-2021 Weight

Cell	Birth year	Sex	Race/ethnicity	Region	CDS sample size	CDS weighted estimate	ACS population totals	Adjustment factor
1	2004-2005	F	NH Black	Not South	25	380,651	266,745	0.700760
2	2004-2005	F	NH Black	South	48	400,196	361,372	0.902987
3	2004-2005	F	NH White/Other	Midwest	28	707,123	645,881	0.913392
4	2004-2005	F	NH White/Other	South	33	984,502	881,752	0.895633
5	2004-2005	F	NH White/Other	West	15	487,742	526,924	1.080333
6	2004-2005	M	NH Black	Not South	17	252,273	276,482	1.095965
7	2004-2005	M	NH Black	South	46	482,342	375,557	0.778611
8	2004-2005	M	NH White/Other	Midwest	20	727,558	693,350	0.952982
9	2004-2005	M	NH White/Other	South	18	814,354	919,744	1.129416
10	2004-2006	F	Hispanic	Not West	15	469,574	957,680	2.039467
11	2004-2006	F	Hispanic	West	16	631,046	614,139	0.973208
12	2004-2006	F	NH White/Other	Northeast	11	603,311	665,416	1.102940
13	2004-2006	M	Hispanic	Not West	25	807,896	1,025,390	1.269211
14	2004-2006	M	Hispanic	West	26	841,105	647,361	0.769656
15	2004-2006	M	NH White/Other	Northeast	16	856,434	711,778	0.831095
16	2004-2006	M	NH White/Other	West	19	938,957	831,529	0.885588
17	2006-2007	F	NH Black	Not South	25	313,647	292,837	0.933652
18	2006-2007	F	NH Black	South	51	528,966	383,123	0.724286
19	2006-2007	F	NH White/Other	Midwest	19	660,717	660,282	0.999342
20	2006-2007	F	NH White/Other	South	19	833,434	866,293	1.039426
21	2006-2007	F	NH White/Other	West	25	644,014	537,940	0.835293
22	2006-2007	M	NH Black	Not South	20	260,702	300,407	1.152300
23	2006-2007	M	NH Black	South	42	482,062	401,236	0.832333
24	2006-2007	M	NH White/Other	Midwest	27	696,743	669,940	0.961531
25	2006-2007	M	NH White/Other	South	25	1,010,047	937,171	0.927849
26	2007-2009	F	Hispanic	Not West	20	925,505	1,036,069	1.119463
27	2007-2009	F	Hispanic	West	14	463,782	671,574	1.448039
28	2007-2009	F	NH White/Other	Northeast	13	539,912	665,847	1.233250
29	2007-2009	M	Hispanic	Not West	35	1,249,601	1,087,006	0.869883
30	2007-2009	M	Hispanic	West	24	660,877	705,724	1.067860
31	2007-2009	M	NH White/Other	Northeast	11	527,293	684,744	1.298603
32	2007-2009	M	NH White/Other	West	16	951,995	864,549	0.908145
33	2008-2009	F	NH Black	Not South	35	277,826	315,270	1.134773
34	2008-2009	F	NH Black	South	45	402,588	411,546	1.022250
35	2008-2009	F	NH White/Other	Midwest	35	707,064	638,693	0.903303
36	2008-2009	F	NH White/Other	South	26	1,042,929	885,853	0.849389
37	2008-2009	F	NH White/Other	West	18	623,741	541,055	0.867435
38	2008-2009	M	NH Black	Not South	21	394,226	301,547	0.764910
39	2008-2009	M	NH Black	South	37	435,680	418,822	0.961307
40	2008-2009	M	NH White/Other	Midwest	27	661,433	678,806	1.026265
41	2008-2009	M	NH White/Other	South	26	1,004,474	904,823	0.900793
42	2010-2011	F	NH Black	Not South	24	240,859	279,902	1.162101
43	2010-2011	F	NH Black	South	37	274,805	371,706	1.352620
44	2010-2011	F	NH White/Other	Midwest	19	607,147	612,846	1.009387
45	2010-2011	F	NH White/Other	South	21	795,219	843,990	1.061331
46	2010-2011	F	NH White/Other	West	10	407,330	514,857	1.263979
47	2010-2011	M	NH Black	Not South	23	299,025	301,580	1.008546

Cell	Birth year	Sex	Race/ethnicity	Region	CDS sample size	CDS weighted estimate	ACS population totals	Adjustment factor
48	2010-2011	M	NH Black	South	56	549,690	382,737	0.696277
49	2010-2011	M	NH White/Other	Midwest	24	663,563	651,064	0.981164
50	2010-2011	M	NH White/Other	South	20	722,599	903,309	1.250083
51	2010-2013	F	Hispanic	Not West	42	1,214,851	1,283,606	1.056595
52	2010-2013	F	Hispanic	West	22	739,464	802,860	1.085732
53	2010-2013	F	NH White/Other	Northeast	23	833,835	818,257	0.981317
54	2010-2013	M	Hispanic	Not West	49	1,389,486	1,323,024	0.952168
55	2010-2013	M	Hispanic	West	23	722,060	808,261	1.119382
56	2010-2013	M	NH White/Other	Northeast	16	506,576	877,912	1.733031
57	2010-2013	M	NH White/Other	West	23	945,447	1,087,292	1.150029
58	2012-2013	F	NH Black	Not South	26	239,299	268,661	1.122699
59	2012-2013	F	NH Black	South	42	273,149	349,825	1.280712
60	2012-2013	F	NH White/Other	Midwest	25	581,005	600,810	1.034087
61	2012-2013	F	NH White/Other	South	27	864,394	781,721	0.904357
62	2012-2013	F	NH White/Other	West	23	545,746	506,896	0.928813
63	2012-2013	M	NH Black	Not South	30	457,103	295,617	0.646719
64	2012-2013	M	NH Black	South	31	344,421	349,927	1.015987
65	2012-2013	M	NH White/Other	Midwest	24	569,036	641,834	1.127932
66	2012-2013	M	NH White/Other	South	17	837,246	837,763	1.000617
67	2014-2015	F	NH Black	Not South	23	245,141	273,180	1.114380
68	2014-2015	F	NH Black	South	36	340,927	352,728	1.034613
69	2014-2015	F	NH White/Other	Midwest	26	697,907	599,206	0.858576
70	2014-2015	F	NH White/Other	South	31	965,913	806,368	0.834824
71	2014-2015	F	NH White/Other	West	10	321,143	500,156	1.557426
72	2014-2015	M	NH Black	Not South	30	369,936	289,156	0.781639
73	2014-2015	M	NH Black	South	55	347,843	370,074	1.063912
74	2014-2015	M	NH White/Other	Midwest	26	755,277	628,478	0.832115
75	2014-2015	M	NH White/Other	South	21	772,022	834,733	1.081230
76	2014-2018	F	Hispanic	Not West	43	1,548,541	1,522,622	0.983262
77	2014-2018	F	Hispanic	West	37	1,111,169	909,006	0.818063
78	2014-2018	F	NH White/Other	Northeast	24	879,341	970,450	1.103610
79	2014-2018	M	Hispanic	Not West	45	1,324,743	1,614,282	1.218562
80	2014-2018	M	Hispanic	West	23	960,569	987,495	1.028031
81	2014-2018	M	NH White/Other	Northeast	39	1,197,214	1,023,158	0.854615
82	2014-2018	M	NH White/Other	West	36	1,318,766	1,288,640	0.977156
83	2016-2018	F	NH Black	Not South	27	394,514	421,557	1.068547
84	2016-2018	F	NH Black	South	60	459,164	533,309	1.161478
85	2016-2018	F	NH White/Other	Midwest	36	928,931	856,109	0.921607
86	2016-2018	F	NH White/Other	South	44	1,038,217	1,157,318	1.114717
87	2016-2018	F	NH White/Other	West	21	933,422	720,882	0.772300
88	2016-2018	M	NH Black	Not South	47	414,614	431,398	1.040480
89	2016-2018	M	NH Black	South	59	510,873	529,402	1.036269
90	2016-2018	M	NH White/Other	Midwest	42	922,138	910,457	0.987333
91	2016-2018	M	NH White/Other	South	28	929,867	1,213,210	1.304713