



The Child Development Supplement
The Institute for Social Research
426 Thompson Street
Ann Arbor, MI 48106
Tel: 734-763-5166
Email: PSIDHelp@umich.edu
<http://psidonline.org/CDS>

The Panel Study of Income Dynamics Child Development Supplement User Guide for CDS-II

July 30, 2010

[This page intentionally left blank]

Acknowledgements

The Child Development Supplement (CDS) to the Panel Study of Income Dynamics was made possible by the generous funding of the National Institute of Child Health and Human Development, the National Science Foundation, and the Center on Philanthropy at Indiana University. CDS-I received additional support from the William T. Grant Foundation, the Annie E. Casey Foundation, the U.S. Department of Agriculture, and the U.S. Department of Education. The CDS project is based at the Institute for Social Research at the University of Michigan. Frank Stafford is the Principal Investigator of the CDS. Jacquelynne Eccles, Robert Schoeni, Wei-Jean Yeung, and Katherine McGonagle are Co-Principal Investigators of the CDS. Tina Mainieri was responsible for the day-to-day management of the CDS-II as Project Manager and wrote much of this user guide. Zoanne Blackburn had primary responsibility for all aspects of data collection of the CDS-II. Special thanks go to the CDS Advisory Board and Sandra Simpkins Chaput and Elizabeth Vandewater who provided guidance and invaluable input on CDS-II instrument development, and Mary Dascola, Brian Haggerty, Malgorzata Grodzky, and Brian Madden who made important contributions to the implementation of this project.

List of CDS-II User Guide Revisions/Additions

7/30/10 – Added Media Guide Documentation

Table of Contents

CHAPTER 1 – AN INTRODUCTION TO THE CDS.....	1
CHAPTER 2 – OVERVIEW OF THE CDSII MODULES	3
PRIMARY CAREGIVER CHILD INTERVIEW	5
PRIMARY CAREGIVER HOUSEHOLD INTERVIEW	6
CHILD INTERVIEW	6
ASSESSMENTS.....	6
TIME DIARY	6
INTERVIEWER OBSERVATIONS	7
OTHER CAREGIVER CHILD AND HOUSEHOLD INTERVIEWS.....	7
SCHOOL DATA.....	7
LOOKING FORWARD: FOLLOWING THE FULL CDS SAMPLE INTO ADOLESCENCE AND YOUNG ADULTHOOD	7
CHAPTER 3 – THE CDS SAMPLE.....	10
THE INITIAL 1997 SAMPLE: CDS-I.....	10
RE-INTERVIEW SAMPLE IN 2002: CDS-II.....	12
DATA COLLECTION OUTCOMES FOR THE INDIVIDUAL MODULES	14
SUMMARY: RESPONSE RATE ACROSS WAVES	15
CHAPTER 4 – THE CDSII SAMPLE WEIGHT.....	17
PART 2: CHANGES IN THE CDS SAMPLE BETWEEN 1997 AND 2002/2003.....	17
CHAPTER 5 – DATA COLLECTION PROCEDURES.....	23
CHAPTER 6 – PRIMARY CAREGIVER INTERVIEWS: TOPICAL GUIDE.....	25
HEALTH.....	25
EDUCATION.....	26
BEHAVIOR PROBLEMS INDEX.....	27
THE POSITIVE BEHAVIORS SCALE.....	28
HOUSEHOLD TASKS	29
PARENTAL MONITORING	29
CHILD CARE	31
ABSENT PARENTS	31
SAMPLE CHILD’S EXTRA CURRICULAR ACTIVITIES	31
CHILD-SPECIFIC EXPENDITURES AND SAVINGS	31
HOME-SF	32
NEIGHBORHOOD MEASURES	33
ROSENBERG SELF-ESTEEM FOR PRIMARY CAREGIVERS.....	33
PEARLIN SELF-EFFICACY SCALE FOR PRIMARY CAREGIVERS	34
MEASURE OF NON-SPECIFIC PSYCHOLOGICAL DISTRESS FOR PRIMARY CAREGIVERS.....	35
PARENTING ATTITUDES, STYLES, AND BELIEFS	35
CHAPTER 7 – CHILD INTERVIEW: TOPICAL GUIDE	37
INTERVIEW CONTENT FOR ALL PARTICIPATING CHILDREN 8 YEARS AND OLDER.....	37
INTERVIEW CONTENT FOR PARTICIPATING YOUTH 10 YEARS AND OLDER	39
INTERVIEW CONTENT FOR PARTICIPATING YOUTH 12 YEARS AND OLDER	40
CHAPTER 8 ASSESSMENTS	42
HEIGHT AND WEIGHT MEASURES.....	42
WOODCOCK JOHNSON TESTS OF ACHIEVEMENT	44
WISC DIGIT SPAN TEST FOR SHORT TERM MEMORY	46

CHAPTER 9 – TIME DIARIES	47
DATA COLLECTION.....	47
CODING.....	48
DATA FILES	49
CHAPTER 10 – CDSII SCHOOL DATA.....	54
CHAPTER 11 CODED VARIABLES IN THE PCG AND CHILD INTERVIEWS	55
OCCUPATION & INDUSTRY FOR YOUTH’S JOBS.....	55
CODE FRAMES FOR OPEN-ENDED QUESTIONS	55
CHILD INTERVIEW	56
PCG CHILD INTERVIEW.....	56
CHAPTER 12 – PSID DATA RESOURCES	58
THE PSID SAMPLE.....	58
PSID CONTENT	59
FAMILY AND INDIVIDUAL LEVEL VARIABLES IN THE PSID	62
THE BUILDING BLOCKS TO USING PSID DATA ON FAMILIES & INDIVIDUALS.....	63
FAMILY COMPOSITION CHANGE.....	66
PSID DATA CENTER FILES.....	67
PSID SUPPLEMENTAL FILES	68
CHAPTER 13– DATA FILE STRUCTURE WITHIN CDS FILES.....	70
CHAPTER 14 – USING THE DATA CENTER AND MERGING FILES.....	72
USING THE DATA CENTER: THE BASICS	72
LINKING DATA FILES.....	80
CHAPTER 15 – A FEW TIPS ON USING THE PSIDCDS DATA RESOURCES	90
APPENDICES	1
APPENDIX A: MEASUREMENT RESOURCE TABLE FOR CDS MODULES	2
APPENDIX B: HOME SCALE MAPPING TO NLSY AND PHDCN	17
APPENDIX C: HOME SCALE RECODING INSTRUCTIONS	25
APPENDIX D: CDSII DATA GROUPINGS	31
APPENDIX E: DOCUMENTATION FOR CDS MEDIA FILES.....	33

Chapter 1 – An Introduction to the CDS

The Child Development Supplement (CDS) is one research component of the Panel Study of Income Dynamics (PSID), a longitudinal study of a representative sample of U.S. individuals and the families in which they reside. Since 1968, the PSID has collected data on family composition changes, housing and food expenditures, marriage and fertility histories, employment, income, time spent in housework, health, consumption, wealth, and more.

While the PSID has always collected some information about children (see the PSID bibliography for research papers on child development) in 1997, PSID supplemented its main data collection with additional data on 0-12 year-old children and their parents. The goal of this new data collection was to improve our understanding of the socio-demographic, psychological, and economic aspects of childhood from a nationally representative longitudinal perspective. The CDS-I successfully completed interviews with 2,394 families (88%), providing information on 3,563 children. In 2002-2003, CDS re-contacted families in CDS-I who remained active in the PSID panel as of 2001. CDS-II successfully re-interviewed 2,019 families (91%) who provided data on 2,907 children and adolescents aged 5-18 years. In 2005, the CDS piloted a specialized interview for cohort members transitioning into adulthood, aged 18 years and older.

The CDS gathers a broad array of measures on developmental outcomes across the domains of health, psychological well-being, social relationships, cognitive development, achievement motivation, and education as well as a number of measures of the family, neighborhood, and school environments in which the sample members live and learn. The breadth and depth of measurement offers a substantially rich resource to study development of children and teens alike from infancy/early childhood through middle childhood and adolescence.

In addition to the survey measures described above, the CDS collects time diaries from the sample children 3-18 years of age. These diaries provide a basic foundation for understanding how children across ages, ethnic groups, and socioeconomic status engage in a range of activities and social circumstances. They additionally offer an excellent opportunity to investigate research questions that examine relationships among time spent in various activities, aspects of the family environment, and outcomes related to achievement, social and behavioral development, and health. The time diaries provide

By nature of the CDS being a supplement to the PSID, the study takes advantage of an extensive amount of family demographic and economic data about the sample children's family—not only parents, but also grandparents, siblings, cousins, and other relatives—providing more extensive family data than any other nationally-representative longitudinal survey of children and youth in the U.S. This rich data structure allows analysts a unique opportunity to fully link information on children, their parents, their grandparents, and other relatives to take advantage of the intergenerational and long-panel dimensions of the data.

Through a separate biennial interview, data are further collected on sample members during their early years of transitioning into adulthood, and in the main PSID when they leave home and become heads of their own households. The extensive information collected on experiences during childhood and adolescence can then be used to help explain patterns of adult life experiences and development. Truly, the multi-level, multidisciplinary, and longitudinal nature of the combined PSID-CDS research design make the CDS an important and very unique resource to the research community.

¹ URL: <http://psidonline.isr.umich.edu/Publications/Bibliography>

The goal of this User Guide is to provide you with detailed information about the overall CDS study design, specific measures, data structure and relationships with the “parent study” – the PSID. More specifically, the next chapter, Chapter 2, introduces the CDS-II data collection modules and their relationship to the CDS-I modules. Chapter 3 describes the CDS-II sample and data collection response rates. Chapter 4 follows with documentation on the sample weight and related issues. Chapter 5 provides highlights of the data collection efforts, and Chapters 6-10 detail the goals and measures in each of the modules – Primary Caregiver, Child, Assessments, Time Diaries, and External School modules, respectively. Chapter 11 briefly lists the variables that were coded in the CDS-II. Chapter 12 is an important chapter for CDS data users to review. This chapter provides information about the rich data resources in the PSID main and supplemental files, the data structure within the PSID, and fundamental aspects of its data structure as they relate to the CDS respondents. In Chapters 13-15, we continue with more information on data structure and linking within the various data groups in the PSID-CDS Data Center.

Chapter 2 – Overview of the CDSII Modules

Both waves of the CDS collect measures of developmental outcomes in the areas of health, achievement, and psychosocial wellbeing, and measures of inputs into development within the context of the child's family, neighborhood, and school. Due to the panel nature of the study, most measures in the second wave duplicated those assessed in the baseline to permit analysis of change over time. The CDS-I User Guide (Hofferth et al., 1999²) provides an overview of the initial wave and its measures. Changes in the type and content of the instruments at the second wave largely reflected measurement needs related to changes in the developmental stages of the children across the data collections. In CDS-I, almost all of the participating children were 0-12 years of age, with a good number not yet in school. The interviews focused on developmental issues appropriate to the infant, early childhood, and middle childhood stages and emphasized contacts with caregivers and preschool /elementary school teachers. In CDS-II, almost one-half of the CDS sample was in adolescence. We added substantially more youth-reported measures, new adolescent-appropriate scales, such as the Adolescent HOME Short-Form Scale, and more psychological and educational scales. Additionally, we expanded the interview with the children themselves, and incorporated many of the new question items in an Audio-Computer Assisted Self-Interview (ACASI) whereby adolescents listened to the questions through a headset and recorded their responses directly into the laptop. This method has been shown to improve data quality for reports of sensitive information such as questions about psychological well-being, sexual behaviors, and experiences with tobacco, alcohol, and drug use (Aquilino, 1994)³.

The CDS-II data are collected from multiple individuals using varied methodology, but all with a focus on the sample child and his/her social and physical environment. The CDS interviewers are, as follows: (1) Primary Caregiver Interview that focused on the child and household characteristics; (2) Child Interview for children aged 8 years and older, including an Audio-Computer Assisted Self-Interview (ACASI) component for sensitive topics asked of adolescents; (3) Standardized Educational Achievement Assessments for all children using the Woodcock Johnson and WISC Digit Span Test; (4) Time Diary for weekday and weekend accounts of use of children's time filled out by either the child alone or the child in collaboration with the mother; (5) Other Caregiver Questionnaire that focuses on the child and household characteristics; (6) Elementary/Middle School Teacher questionnaire, that focuses on the child's behavior and achievement at school, characteristics of the classroom and the school, in general, and characteristics of the teacher; (7) Course and Grade Reports of adolescents as one measure of academic choices and achievement patterns (8) External School Database to gather information about the school and school district environments. Table 2 below provides a summary of the content of the CDS modules. In this chapter, we review the modules included in the second wave. Chapters 6-10 provide more detail about the specific measures within each module, and Chapter 12 provides details on measures collected in the PSID main family interview.

² Hofferth, S., Davis-Kean, P.E., Davis, J., & Finkelstein, J. (1999). The Child Development Supplement to the Panel Study of Income Dynamics: 1997 User Guide. <http://psidonline.isr.umich.edu/CDS/usergd.html>.

³ Aquilino, W. S., (1994) Interview mode effects in surveys of drug and alcohol use: a field experiment. *Public Opinion Quarterly*, 58, 210-240.

Table 2.1 CDS-I and CDS-II Measurement Domains

Domain	Description	Module
Health Status & Behaviors*	General health status, chronic conditions, obesity, limitations, health care utilization, health-related expenditures, nutrition, exercise, sleep, smoking	PCG Child
Psychological & Social Well-Being*	Behavior problems, depression, self-esteem, worry, social well-being; risky behaviors, thrill seeking, anti-social behaviors; drug and alcohol abuse /dependence	PCG Child
Family Environment*	HOME SF cognitive & emotional stimulation; parental warmth; household tasks; involvement, closeness, conflict w/ father & mother	PCG OCG
Parental Monitoring*	Caregivers' knowledge of the child's whereabouts, activities, and associations; child disclosure of activities	PCG; OCG Child
Child Care	Type, frequency of use, and costs of arrangements for children up to Kindergarten.	PCG
Education	Parental expectations; enrollment; type of school; tuition; attendance; federal lunch & breakfast programs; attended special class/school for gifted students; classified as needing special education; repeated grade; dropped out	PCG
Achievement	Woodcock Johnson tests of achievement; course grades; WISC Digit Span short-term memory; ability self-concepts in reading and math	Child
Time Use	Stylized questions about structured and unstructured activities, activities with parents, extra curricular, part-time jobs; Time Diary measures of type, number, duration, and location of weekday and weekend activities	PCG Child
Religiosity*	Comfort, importance of religious affiliation or spirituality	Child
Future Work & Schooling Expectations*	Achieved occupational certainty and identity; job values, career orientation and expectations for future work and schooling; negative economic expectations	Child
Sibling Relationships	Type and frequency of cooperation with, kindness towards, and helping behaviors towards siblings	PCG
Caregiver Social and Psychological Resources	Rosenberg self-esteem; Pearlin self-efficacy; K-6 non-specific psychological distress; social support; parenting attitudes; aggravation in parenting; gender role beliefs; family conflict; economic strain; work schedules; community involvement	PCG; OCG

Domain	Description
Absent Parents	Frequency/types of activities absent parents are involved with their children; conflict between resident and absent parent
Spending & Savings*	Variety of expenses for household members, non-household members, and absent parents; savings mechanisms for child
School Environment	School type; racial/ethnic composition; pupil-teacher ratio; completion rates; expenditures per child; other school resources; climate; curriculum tracks; science/math courses

*New domain added for CDS-II, or expanded in CDS-II.

**CCD=Links to Common Core of Data, which replaced school administrator interview in CDS-I.

Primary Caregiver Child Interview

The Primary Caregiver Child Interview (“PCG-Child”) is the anchor to all other CDS modules and required for a sample child to be considered “response” in the study. We considered it “anchor” since it contained key developmental measures for the study, including almost all of the HOME SF scale items and other family environment measures, the Behavior Problem Index, parent-report on schooling and education activities, extra curricular activities, health, and child expenditures and savings.

As in CDS-I, by definition, the PCG must have lived with the child to be considered a primary caregiver. In most cases, the PCG was the child’s biological mother. If the biological mother was not living with the child, the appropriate respondent was the first person on the list of relationships below who was living with the child:

- Stepmother, adoptive mother, or foster mother
- Other female legal guardian (must be at least 18 years of age; oftentimes the grandmother)
- Father (biological, adoptive, step, or foster)
- Male legal guardian of the sample child
- An adult (at least 18 years of age or older) who lives with the child and takes primary responsibility for caring for him/her. This does not include someone who is paid or hired to take care of the child (i.e., not a babysitter or nanny).

The variable RELPCG02 in the data center specifies the familial relationship of the PCG to child. We are constructing a parallel variable for CDS-I – to be released in the winter of 2006.

We expanded the PCG-Child interview in CDS-II to include new questions about the sample child’s experiences with asthma and asthma-related symptoms, extended items on family rules and rule enforcement, extended HOME Scale items, questions about extracurricular involvement in school, sports, and the community; and expanded questions directly related to parents’ economic and time investments in their children.

Primary Caregiver Household Interview

While the PCG-Child interview focused on the sample child, the PCG-Household Interview (PCG-HH), in turn, focused on household characteristics (e.g., neighborhood, family rules) and the PCG's own psychological resources and social support, parenting values, stresses, and styles, and beliefs about the parental role in child rearing. This module remained largely the same across CDS data collection waves.

Child Interview

Children who were at least eight years of age and older were eligible to participate in an in-home interview themselves. For the most part, the Child Interview was a new addition to CDS-II, focusing on: self-concepts for reading and math abilities, school connectedness, peer victimization and bullying, family relationships, electronic media use, work experiences (for youth 12 years and older), and future expectations. The Eccles self-concept items for reading and math and the self-esteem measure were the only measures in the Child Interview that were also assessed in the CDS-I.

Children in the adolescent-stage additionally participated in an Audio-Computer Assisted Self-Interview (ACASI) component whereby the children listened to the questions themselves on a headset and then responded directly into the laptop. Children 10 years and older participated in Section K of the Child Interview that queried the youth about their hobbies and extracurricular activities, characteristics of friends, and health behaviors such as dietary, exercise, and sleep patterns. For youth 12 years and older, the ACASI module included items on the youth's own savings and expenditure patterns, tobacco, alcohol and drug use, delinquent behavior, educational aspirations and expectations, parental monitoring reported from the youth's perspective, and languishing and flourishing.

Assessments

The assessments in CDS-II replicated those in CDS-I: physical measurements and achievement tests. Interviewers collected height and body weight of the children, and administered standardized educational achievement tests using the Woodcock-Johnson Revised Tests of Achievement and the WISC Digit Span Test for Memory. Height and weight were gathered in CDS-I in the primary caregiver interview, and the assessments were included in the "Child Interview". The physical and cognitive assessments provide important, high-quality outcome measures for the study. More information about the assessment administrative procedures and data quality are described in Chapter 8 of this guide.

Time Diary

The time diaries are one of the unique features of the CDS design. While the PCG and Child interviews included stylized questions about the children's structured and unstructured activities, and activities with parents and absent parents, the time diaries provide detailed accounting of the type, number, duration, and location of activities during sampled 24-hour days, beginning at midnight for one randomly sampled weekday and one randomly sampled weekend day. Using the time diaries, we additionally collected information on the social context of the activity by specifying with whom the child was doing the activity and who else was present, but not engaging.

Interviewer Observations

After the in-home interview, the interviewer recorded observed interactions between the primary caregiver and the child/adolescent, and characteristics of the physical environment of the home and neighborhood. Most of these items are replicated from the HOME Scale Short Form, as used in the NLSY79 and can be summarized in terms of cognitive stimulation and emotional support. In CDS-II, we included additional observational items that were not a part of the NLSY79 version, but included in the Project on Human Development in Chicago Neighborhoods Study. Observations are fully described in the appendix under the HOME Scale Specifications.

Other Caregiver Child and Household Interviews

CDS-II collected information about the sample child and home environment from a secondary caregiver in the home (“OCG”), broadly defined as a second person living in the household who had responsibility for raising the sample child. There has been substantial interest in the involvement of fathers in the lives of their children. In many low-income households, however, the second caregiver can often be a grandmother, not a father or partner. The expanded “OCG” definition in both waves of the CDS permit examination of the contribution of a second caregiver more generally over many other studies that just focus on the father. During data collection, OCGs were identified for 2,009 CDS sample children (69%). The content of the OCG-Child and OCG-Household interviews was drawn directly from the PCG interviews and remain largely consistent across CDS data collections.

School Data

In CDS-II, we bolstered our data resources that focus on the school environment. These data were designed to allow researchers to both model school level influences on child and adolescent development and to control for school level influences when modeling family and neighborhood level influences on developmental outcomes. We replaced the school administrator interview from CDS-I with linkages to the National Center for Education Statistics Common Core of Data (NCES CCD). The CCD provides access to a broader range of variables about the schools and data at a high response rate than the CDS-I School Administrator surveys. Selected Variables from NCES CCD that contains information about the school that the CDS sample child attended at the time of the CDS-II data collection (school year 20022003) were posted to the CDS Data Group in the PSID-CDS Data Center. The variables we posted are a small sampling of data about the school and school district environment that are available through the CCD (<http://nces.ed.gov/ccd/index.asp>). Some examples of the data we have selected data are school type, racial/ethnic composition, pupil-teacher ratio, completion rates, and expenditures per child. Sensitive data contracts are available for the school and school-district linkages for the sample children.

In CDS-II, we added a module that asks middle school and high school students about school courses to obtain detailed information on educational progress and achievement. These data are planned for release in 2006.

Looking Forward: Following the Full CDS Sample into Adolescence and Young Adulthood

Almost ten years ago, the CDS sample was selected from all PSID families with children 0 to 12 years of age. At the second wave of CDS, almost one-half of the sample had aged into adolescence. In 2007, the remaining younger half will be in adolescence. We plan to collect a third wave of CDS data at this point, using much of the same measures as used in the CDS-2.

During CDS-II, we observed a small number of individuals who had already turned 18 years. By 2005, the number of CDS youth at least 18 years and no longer in high school was close to 900. These postsecondary school years have been recognized as ones with new developmental opportunities and challenges as individuals' self identities continue to evolve, future expectations and goals further develop and begin to materialize, choices in how discretionary time is used and invested increase, and investments in tertiary education and careers take root.

In 2005, with the growing proportion of CDS youth 18 years and older, we developed a new survey module to capture data on young adult developmental pathways and outcomes, which served a scientific purpose of continued collection of detailed information about development for the sample, and data on the status of early adulthood markers (as collected in the PSID main family interview for Heads and Wives). The interview domains coordinated with the CDS adolescent measures, provided new questions critical to assessing development at this stage, and linked into the adult status markers in the PSID on employment, income, and health. In 2007, the CDS young adults interviewed in 2005 were reinterviewed, and CDS youth who had newly turned age 18 and completed or left high school were also interviewed.

The table below shows the sample distribution of the CDS youth by age in upcoming PSID data collection years. Cells shaded in green are the CDS Youth 10-17 years who will be eligible for CDS-3, and cells shaded in light purple are the CDS Youth who will be eligible for the Transition into Adulthood Interview. Note that sample size estimates account for family attrition from PSID only up through 2003, but beyond that, and other individual-level attrition are not accounted for in this table.

Table 2.2. Estimated Number of CDS Individuals by Upcoming PSID Data Collection Year for all CDS Sample Members Whose Families Were Still Active in PSID as of 2003

Birth Year	2003	2005		2007		2009		2011	
	# Kids by Age in 2003	Age (Yrs)	N						
1997	51	8	51	10	51	12	51	14	51
1996	222	9	222	11	222	13	222	15	222
1995	257	10	257	12	257	14	257	16	257
1994	243	11	243	13	243	15	243	17	243
1993	245	12	245	14	245	16	245	18	245
1992	268	13	268	15	268	17	268	19	268
1991	251	14	251	16	251	18	251	20	251
1990	215	15	215	17	215	19	215	21	215
1989	224	16	224	18	224	20	224	22	224
1988	224	17	224	19	224	21	224	23	224
1987	242	18	242	20	242	22	242	24	242
1986	236	19	236	21	236	23	236	25	236
1985	238	20	238	22	238	24	238	26	238
1984	218	21	218	23	218	25	218	27	218
1983	2	22	2	24	2	26	2	28	2
TA Sample Size per Year		936		1384		1850		2363	
CDS-III Sample Size 2007				1752					

When the CDS youth become economically independent from their parents, that is, when they

“split off” into their own households, they become eligible for participation in the main PSID family interview, the point in the life course when the PSID begins to collect its most detailed, long term information on the sample members’ adult health, family, work, wealth accumulation, and more.

Chapter 3 – The CDS Sample

Using sample weights, the CDS is designed as a nationally representative sample of children in the United States. This chapter describes the initial sampling in 1997, the re-interview sample at the second wave, and data collection response rates by module, sex, age, and race for the CDS-II. Chapter 4 provides information on the CDS-II sample weight.

The Initial 1997 Sample: CDS-I

In 1997, the sample for the CDS was drawn at the time the 1997 PSID interviews. As interviews were completed for the main PSID, family units with followable children under the age of 13 were identified for inclusion in the Child Development Supplement. Both the PSID core sample and the new PSID immigrant refresher sample generated Child Development cases. Out of the 2,705 families selected for the CDS-I, 2,394 families (88%) participated, providing information on 3,563 children. Households outside the continental U.S. were excluded from interviewing and were coded as non-sample. All other households were included in the sample.

After the data were collected and prepared for release, it was determined that 83 of the 3,563 children were non-sample and not a part of the PSID. These children were kept in the CDS-I data file for the purpose of cross-sectional analysis. At each wave of the PSID, when new information is provided about the families, it is possible that we may find that other families and/or children within families are not actually sample members. When this happens, the variables described below are updated. If you are mapping CDS-child level data with family-level data in the PSID, and using the data longitudinally, the variables described below, located on the PSID Individual File, 1997, are very useful.

Sampleness, Selection, and Participation Indicators for CDS Sample Members

(ER 33418) Whether Eligible for the 1997 Child Development Study

Code	Code Description
1	Eligible for CDS
5	Ineligible for CDS (child was under age 13 but was nonsample and was mistakenly selected for CDS)
0	Inap.: this individual was from Immigrant recontact or Latino sample (ER30001=3442-3511, 7001-9308); main family nonresponse by 1997 or mover-out nonresponse by 1996 and was not selected for recontact in 1997 (ER33401=00000); born or moved in after the 1997 interview (ER33401>00000 and ER33402=00); born or moved in after the 1997 interview (ER33401>00000 and ER33402=00); aged 13 or older (ER33404=013-999)

(ER33419) Whether Selected for the Child Development Supplement (CDS) to the PSID

Code	Code Description
1	Selected for CDS
5	Eligible but not selected for CDS
0	Inap.: this individual was from Immigrant recontact or Latino sample (ER30001=3442-3511, 7001-9308); main family nonresponse by 1997 or mover-out nonresponse by 1996 and was not selected for recontact in 1997 (ER33401=00000); born or moved in after the 1997 interview (ER33401>00000 and ER33402=00); born or moved in after the 1997 interview (ER33401>00000 and ER33402=00); aged 13 or older (ER33404=013-999)

(ER33420) Response Indicator for the CDS-I – 1997

Code	Code Description
1	Interview (IW) with sample child
2	IW with non-sample child
3	Non-response—refused
4	Non-response—lost
5	Non-response—incapacitated, permanent condition
6	Non-response—deceased
7	Non-response—could not contact primary caregiver for actual IW
8	Non-response—area too dangerous for interviewer
9	Non-response—language barrier
20	Non-response—study ended before IW could be completed
96	Office Error—interviewer reports IW taken, but not logged
97	Office Error—eligible child who should have been selected for IW
98	Office Error—outside continental U.S., should have been called
0	From Immigrant recontact or Latino sample (ER30001=3442-3511, 7001 9308); main family nonresponse by 1997 or mover-out nonresponse by 1996 (ER33401=00000); born or moved in after the 1997 interview (ER33401>00000 and ER33402=00); in institution, moved out or died in 1997 (ER33402=51-89); aged 13 or older (ER33404=013-999); ineligible and no IW taken or eligible but not selected (ER33419=5)

Re-interview Sample in 2002: CDS-II

For CDS-II, we essentially followed the sample children of CDS-I whose families were still active in the PSID 2001 main study. Specifically, eligibility criteria for inclusion in the second wave of CDS were, as follows:

- Response in CDS-I (n=3,563)
- Considered sample at the time of the main PSID data collection in 2001 (n=3,480)
- Family of the targeted CDS child still active in the PSID panel as of the 2001 main interview. In-between 1997 and 2001, 94% of the CDS children had families who remained active in PSID (n=3,271).

During data collection, there were some situations in which we did not attempt interviews with or for sample children who met the criteria above. We attempted interviews for the CDS Sample child if:

- Sample child still resided under the care of a primary caregiver (meaning, he/she had not yet established own residence);
- Sample child lived in U.S.;
- Sample child did not reside on a military base; and
- Sample child was non-institutionalized.

CDS-II data collection resulted in completed interviews for 2,907 of the 3,271 “followable” children in 2002-2003 who still met data collection eligibility, resulting in a response rate of 91%. The reasons for not completing an interview during the CDS-II data collection are specified in Table 3.4.

Table 3.1. CDS-II Data Collection Outcomes

Description of Data Collection Outcomes	Count
Interview completed about the with the PCG	2,907
Final Refusal; interview not taken	242
Lost; tracking efforts exhausted; interview not taken	15
Institutionalized for physical or mental health reasons; interview not taken	4
Language barrier with PCG; interview not taken	10
Interviewer error; interview not taken	13
Deceased (No longer eligible)	4
No longer residing with a PCG (No longer eligible)	59
Lived outside of U.S. (No longer eligible)	5
Lived on military base (No longer eligible)	7
Incarcerated or in detention home/center (No longer eligible)	5
Total	3,271

The ages of the sample children at the time of the Primary Caregiver Interview in CDS-II ranged from 5 years up to 19 years. Age and gender distributions for response children are highlighted in Table 3.2.

Table 3.2. CDS-II Response by Sample child's Age and Sex

Age in Years	Number of Males	Number of Females	Total for Age	Percent of Total Sample
05	26	25	51	1.75
06	104	96	200	6.88
07	115	135	250	8.60
08	121	115	236	8.12
09	124	95	219	7.53
10	144	108	252	8.67
11	110	121	231	7.95
12	108	111	219	7.53
13	109	99	208	7.16
14	104	109	213	7.33
15	102	122	224	7.71
16	119	111	230	7.91
17	112	95	207	7.12
18	73	91	164	5.64
19	0	3	3	0.10
Total	1,471	1,436	2,907	100.0

In the CDS-I PCG Child Interview (Question A6), the Primary Caregiver reported the sample child's race as: White non-Hispanic, Black non-Hispanic, Hispanic, Asian or Pacific Islander, American Indian or Alaskan Native. Using this variable, the distribution of CDS-II response to the Primary Caregiver Interview in 2002-2003 is, as follows:

Table 3.3. CDS-II Response by Sample child's Race

Race	Total for Age	Percent of Total Sample
White Non-Hispanic	1,365	46.96
Black Non-Hispanic	1,187	40.83
Hispanic	213	7.33
Asian Or Pacific Islander	38	1.31
American Indian Or Alaskan Native	11	.38
Other	86	2.96
Not Answered	7	.24
Total	2,907	100.0

Out of the 2,907 response to the Primary Caregiver interview in CDS-II, 63.9% of the sample children (n=1,857) had a sibling who also participated CDS-II.

Data Collection Outcomes for the Individual Modules

Each CDS sample child could have up to eight modules of data collected from three different family members and a school resource. Eligibility for some modules depended on child’s age or whether a secondary caregiver resided in the home for the child. Chapter 4 defines the modules and eligibility in more detail. The table below provides an overview of data collection response rates per CDS-II module.

Table 3.4 Response Rate per CDS-II Module

Data File	Number of Complete Interviews	Number of Eligible Participants	Data Collection Response Rate
Primary Caregiver (PCG) Child Interview	2,907	3,191	91%
Primary Caregiver (PCG) Household Int.	2891	3,191	91%
Assessment Interview	2,644	2,907	91%*
Child Interview	2,182	2,406	91%*
Time Diary	2,569	2,907	88%*
Other Caregiver (OCG) Child Interview	1,686	2,009	84%*
Other Caregiver (OCG) Household Interview	1,686	2,009	84%*
Elementary/Middle School Teacher	699	1,305	54%*
School Environment	2,203	2,538	87%*

*Contingent on participation in the Primary Caregiver Child interview.

In the CDS-II Demographic File, you will find a series of variables that identify the status of individual modules. Each variable is coded to specify whether the sample child had a completed interview for that module (“2”), an incomplete for that module (“5”), or the module was not applicable for that child (“0”). PCG module response pattern for CDS-II by age is, as follows:

Table 3.5. Total PCG Interviews and Non-Interviews by Age of CDS Child

Count Col Percent PCG Interview	Age of CDS Child					
	5 Yrs	6-9 Yrs	10-14 Yrs	15-17 Yrs	18+ Yrs	Total
Interview	51 100%	905 90%	1123 91%	661 91%	167 94%	2907
Non-Interview	0	99	111	63	11	284

Out of all CDS-II children who had a completed PCG interview, the module-level interview rate, for each module, by age of the CDS child is, as follows:

Table 3.6. Total Assessment Interviews and Non-Interviews by Age of CDS Child

Count Col Percent of Eligible Assessment	Age of CDS Child					
	5 Yrs	6-9 Yrs	10-14 Yrs	15-17 Yrs	18+ Yrs	Total
Interview	49 96%	844 93%	1013 90%	592 90%	146 87%	2644
Non-Interview	2 4%	61 7%	110 10%	69 10%	21 13%	263
Total # Children	51	905	1123	661	167	2907
Child Interview 8+ Yrs						
	5 Yrs	6-9 Yrs	10-14 Yrs	15-17 Yrs	18+ Yrs	Total
Interview	NA	425 93%	1013 90%	587 89%	144 86%	2169
Non-Interview	NA	30 7%	110 10%	74 11%	23 14%	237
Age Ineligible	51	450	0	0	0	501
Total # Children	51	905	1123	661	167	2907
Time Diary						
	5 Yrs	6-9 Yrs	10-14 Yrs	15-17 Yrs	18+ Yrs	Total
Interview	49 96%	820 91%	991 88%	570 86%	139 83%	2569
Non-Interview	2 4%	85 9%	132 12%	91 14%	28 17%	338
Total # Children	51	905	1123	661	167	2907
OCG						
	5 Yrs	6-9 Yrs	10-14 Yrs	15-17 Yrs	18+ Yrs	Total
Interview	33 87%	567 87%	630 82%	372 83%	84 83%	1686
Non-Interview	5 13%	84 13%	143 19%	74 17%	17 17%	323
Ineligible—no OCG in the home	13	254	350	215	66	898
Total # Children	51	905	1123	661	167	2907

Summary: Response Rate across Waves

3,563: CDS-I children interviewed.

3,480: CDS-I children were part of the PSID sample at time of PSID 2001 Interview (3,563 - 83 non-sample children who were interviewed for CDS-I = 3,480).

3,271: CDS-I children had families who were still active in the PSID panel as of the main PSID interview in 2001

(3,480 - 209 children whose families no longer were in PSID = 3,

271).

2,907: CDS-I children have completed data for CDS-II data collection.
CDS panel and family data 1997-2001: $2,907 / 3,480 = 84\%$.

Chapter 4 – The CDSII Sample Weight ⁴

Elena Gouskova, Ph.D., Survey Research Center, Institute for Social Research

As described in Chapter 2, in the first wave of the CDS, 2,394 families participated, providing information on a total sample of 3,563 children. The second round of data collection took place in 2002-2,003 when 2021 families were successfully re-interviewed, resulting in total of 2,907 child interviews. The reduction in the sample is due to reclassification of the eligibility status of a number of 1997 sample participants and also due to nonresponse among the remaining eligible CDS families. To account for differential probabilities of selection due to the original PSID sample design and subsequent attrition, CDS data are provided with weights. This note describes the process of the 2001 Child Supplement weight construction and discusses some related issues.

The chapter is organized as follows. Part Two provides information on the sample attrition between 1997 and 2001. The Primary Caregiver/Child weight construction procedure is described in Part Three. In Part Four, we discuss the complex structure of the CDS which consists of several interview modules. In this section we highlight the issue of missing data for special modules and suggest possible ways to deal with this missing data problem in the analysis of these data.

Part 2: Changes in the CDS Sample Between 1997 and 2002/2003

Out of the 3,563 children with complete CDS-I interviews in 1997, information was gathered for 2,907 children in CDS-II during 2002/2003. The reduction in the sample comes from two sources: the decision not to follow some families from the original 1997 CDS sample (292 children) and non-participation by eligible families in the 2002-2003 interview process (364 children). A summary of the sample size changes in the CDS panel is provided in Table 4.1.

Table 4.1. Sample Size Changes in the CDS Panel

Suspended in 2001		292
	Subtotal	292
Targeted in 2002	PCG interview: NO	364
	PCG interview: YES	2,907
	Subtotal	3,271
Total (1997 sample size)		3,563

Prior to the 2002-2003 CDS data collection, a decision was made to follow only 1997 CDS baseline families with children that had sample status⁵. Of the 3,563 children who participated in CDS-I, 83 were “non-sample” according to standard PSID rules for longitudinal panel eligibility for newborns and other new household members.

⁴ This chapter is from *PSID Technical Report: the 2002 PSID Child Development Supplement (CDS-II) Weights*, written by Elena Gouskova (August 24, 2004).

<http://psidonline.isr.umich.edu/CDS/questionnaires/cdsiiweights.pdf>

⁵ PSID respondents are distinguished as sample and non-sample members. Sample respondents are those who belong to the original 1968 family for the core PSID sample or original 1997 family for Immigrant sample. Sample members are also those who are related by blood or adoption to a PSID sample member.

The second eligibility criterion for CDS-II was that the family of the targeted CDS child needed to be active in the PSID panel as of the 2001 main interview. Between 1997 and 2001, 94% of the CDS children had families who remained active in PSID (that is participated in the 2001 PSID biennial interview). Thus, the suspended portion of the sample consisted of families with non-sample children and families that left the PSID study before 2001. The total number of suspended cases is 292, reducing the targeted sample of children from 3563 in 1997 to 3271 who were eligible to participate in 2002/2003.

Starting with this sample, the CDS-II successfully obtained interviews with 2006 PSID families, resulting in 2907 child observations. Accounting for both sources of attrition in the 1997 baseline CDS-I panel, the CDS-II re-interview rate is 81.6%. Among PSID “sample” children determined to be eligible for CDS-II re-contact, the interview completion rate for the primary care giver (PCG) interview was 88.9%.

CDS-II Weight for Primary Caregiver (PCG)/Child Analysis

CDS-II, like CDS-I before it, includes different interview modules that are designed to collect data from multiple “observers” that may be associated with each sample child. In CDS-II, these “observers” always include the child’s primary caregiver (PCG) but may also include: 1) a secondary care giver; 2) a teacher; and 3) the child herself/himself (both interview and test assessment). The primary caregiver interview is the minimum data required to qualify as a CDS-II case. Table 1 summarizes the eligibility and response outcome for the PCG interview. Eligibility and response dispositions for the remaining CDS observation modules are summarized in Table 4.2.

Table 4.2. Eligibility and Response Outcome for the CDS-II PCG Interview

Questionnaire Module	Status	N	Response Rate*
Child Interview	Not eligible Yes interview	501 2,182	
	No interview	480	
	Total	2,907	91.0%
Child Assessments	Not eligible Yes interview	0 2,644	
	No interview	263	
	Total	2,907	91.0%
Secondary Caregiver	Not eligible Yes interview	898 1,686	
	No interview	323	
	Total	2,907	83.9%
Elementary School Teacher Interview	Not eligible Yes interview	1,602 699	
	No interview	606	
	Total	2,907	53.6%

*Excluding ineligible cases

The CDS-II analysis weight described below pertains to the analysis of the PCG interview data. Special weights are not provided for the analysis of the other CDS-II modules listed above. Section 4 below provides guidance on how analysts may address the problem of missing data for other child observation modules.

Sample survey data are typically provided with weights designed to compensate for unequal probabilities of sample selection and non-response or data that is missing at random (MAR) (Little and Rubin, 2002)⁶. These weights are inversely proportional to the probability that each observation is selected and conditional on selection that they respond to the survey questions. With longitudinal data this joint probability at time t, where the study has started at t-1 or earlier, can be expressed as following

$$P(St=1)=P(St-1=1)*P(Rt=1|St-1=1)$$

where St is an indicator of participation in a study at time t and Rt is an indicator of response at time t. The probability of being a participant at time t is a product of probability of participating in previous period and conditional probability of being a response in the current period. As the first term on the right hand side of (1) is proportional to a reciprocal of weight in the previous period, the weight in the current period is a product of the weight in previous period and the inverse of probability of response (the second term on the right hand side of (1)). We will refer to $1/ P(Rt=1|St-1=1)$ as attrition adjustment factor.

Accordingly, the CDS-II weight is a product of CDS-I weight and the attrition adjustment factor. The individual level Primary Caregiver/ Child weight (CH97PRWT) was used as the previous year weight⁷. To obtain the attrition adjustment factor, the probability that a sample person was successfully re-interviewed in the CDS-II was modeled with the linear logistic model. As we discussed above there were two different reasons for attrition. We modeled them together, i.e. observations where a person was not interviewed included both suspended and non-response cases⁸. The estimates of the logistic model are given in Table 4.3.

⁶ Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis With Missing Data, 2nd Edition. John Wiley & Sons, New York.

⁷ For the description of the 1997 CDS weights construction see <http://psidonline.isr.umich.edu/CDS/weightsdoc.html>

⁸ We experimented with separate modeling of the two types of attrition. The results turned out to be very similar.

Table 4.3. Estimates of the Logistic Model to Obtain the Attrition Adjustment Factor

Variable	Estimate	Err	Std	Wald ChiSq	Prob ChiSq
Intercept	-0.2847	0.3664		0.60	0.4372
		**			
D Sample	-0.9745 *	0.1204		65.51	<.0001
		**			
D SRC sample	-0.6414 *	0.228		7.91	0.0049
D SEO sample	-0.5121 **	0.246		4.33	0.0374
		**			
D Head and Wife in FU	-0.9601 *	0.2131		20.29	<.0001
D Child is male	0.0804	0.0889		0.82	0.3655
Age of child at 1997 PSID core interview	0.0336 **	0.0133		6.39	0.0115
D White (child race)	-0.1442	0.2028		0.51	0.477
D Black (child race)	-0.447 **	0.2134		4.39	0.0362
D age<=30 (head)	0.1447	0.1603		0.81	0.3667
D 30< age<=45 (head)	-0.0411	0.1374		0.09	0.7648
		**			
D Head is male	0.6654 *	0.2226		8.94	0.0028
D No High School Degree (head)	0.2276	0.1596		2.03	0.1538
D High School Degree (head)	0.0872	0.1452		0.36	0.5482
D Some College (head)	0.1417	0.1499		0.89	0.3446
D Head is employed	-0.2221 *	0.1279		3.02	0.0824
D Inc in 1st quartile	-0.3431 *	0.1807		3.61	0.0576
D Inc in 2nd quartile	-0.1062	0.1506		0.50	0.4807
D Inc in 3rd quartile	-0.088	0.137		0.41	0.5206
		**			
D Northeast	0.5613 *	0.1723		10.61	0.0011
		**			
D North Central	0.4618 *	0.1581		8.53	0.0035
		**			
D South	0.603 *	0.1493		16.32	<.0001
D MSA 1990	0.0365	0.1104		0.11	0.7407
Summary statistics					
N:	3563				

As the final step in the weight development, the newly constructed CDS-II weights for Primary Caregiver/child observations were censored to reduce the influence of extreme weights on the variances of sample estimates of population statistics. One percent of weights at the top and bottom of the distribution were assigned values of the 99th and 1st percentiles respectively.

To check the 2002/2003 CDS-II PCG weight we compared weighted estimates for some basic demographic, geographic and socio-economic variables in 1997 (weighted by the 1997 CDS-I weight) to the same estimates for the smaller CDS-II re-interview sample (weighted by the CDS-

II weight). The results of this comparison are provided in Table 3.4. The comparison of weighted estimates for the baseline CDS-I and follow-up CDS-II suggest that the CDS-II attrition adjustment factors included in the PCG/Child analysis weights are compensating for potential attrition bias in the family type and demographic composition of the CDS panel data. Note: This comparison does necessarily not rule out the possibility of spurious or more subtle forms of selection bias that may not be associated with the demographic, geographic and socio-economic characteristics of CDS respondents.

Table 4.4: Comparison of Weighted Estimates in CDS-I and CDS-II

		Weight 1997		Weight 2001	
		N	Percent	N	Percent
Total		3563	100.0	2907	100.0
Region	Northeast	483	18.0	389	17.8
	North Central	877	24.1	720	23.8
	South	1599	33.4	1283	33.0
	West	604	24.5	515	25.4
Immigrant sample	Non-IMM	3234	83.9	2659	83.8
	IMM	329	16.1	248	16.2
MSA	Non-MSA	867	29.6	710	30.0
	MSA	2696	70.4	2197	70.0
Education of Head	No High School Diploma	806	19.0	630	18.8
	High School Diploma Only	1284	32.2	1057	32.1
	Some College or More	797	23.3	650	22.9
	College or More	676	25.4	570	26.1
Age of Head	30 or younger	912	22.0	735	22.2
	31-45	2208	66.2	1820	66.1
	46 or older	443	11.8	352	11.7
Gender of Head	Female	1063	22.1	852	21.4
	Male	2500	77.9	2055	78.6
Race of Head	Non-Black	2102	84.1	1721	84.8
	Black	1461	15.9	1186	15.2

The final CDS-II weight is stored in the variable CH01PRWT. The CH01PRWT weight should be used for analyses involving child level data or data involving the relationship of the child with their PCG or with family characteristics.

The Primary caregiver/ Child weight, CH01PRWT, can be converted into a caregiver level weight by averaging the weights for each CDS-II child observation provided by a caregiver. This PCG/family weight could be used to develop an analysis that is focused on the caregiver or their family as the unit of observation/analysis.

Addressing Missing Data in Other CDS-II interview Modules

The CDS-Interview has a complex structure involving several modules. These modules are the Primary Caregiver (sections A-H, J), Child Interview, Child Assessment, Secondary Caregiver, and Elementary School Teacher Interview. The Primary interviews form the core data collection of the Supplement. Completion of these interviews determined whether an observation is included in the final data set or not.

Thus for all children in the CDS data set the Primary Caregiver interview were responses. Response /nonresponse the Primary caregiver interview was also an indicator for attrition in calculation of the CDS-II attrition adjustment parameter described in the previous section.

Other CDS-Interview modules are secondary to the Primary Caregiver module and may or may not have been successfully completed. It is possible to have information on the Primary Caregiver interview and have missing data for any of the secondary module interviews. It is important to note that all secondary interviews, except for the Child Assessment interview, had an additional eligibility condition. For example, the Child Interview was administered only to children ages 8 or older. Thus missing data on the secondary interview with eligibility condition can be due to two reasons: non-response or ineligibility.

When accounting for the eligibility restrictions, the response rate to the secondary interviews is generally lower than the response to the Primary Caregiver interview. The case counts by eligibility and response status for each secondary module are given in Table 4.2.

Researchers interested in analyzing the CDS-II data for the secondary modules should carefully examine the extent of the partially missing data problem and address it if necessary.

The issue can be approached in a number of ways. The simplest way is to assume that the data are missing completely at random (MCAR). That is, given the adjustment that has already been performed for attrition in the CDS PCG observation sample, non-response to a secondary module is completely random with respect to the dependent variables of interest. When MCAR is the appropriate assumption, analysis of the observed data for the secondary modules will not lead to biased estimation. The MCAR assumption, however, is too strong and may not always be a good approximation for an attrition process. For example, our preliminary investigation indicates that for the Child Interview, age was an important factor in participation decision. The younger children tend to be more likely to participate than older children. If the probability of nonresponse can be modeled as a function of observed variables such as age or gender, Little and Rubin (2002) describe the missing data mechanism as missing at random (MAR).

When the nonresponse or missing data mechanism is MAR, one can approach the problem by either imputing the missing values or constructing additional weighting adjustments using procedures that take into account variables (such as age in the above example) that predict the probability of a response. Analysts who are interested imputation of missing response data for the CDS secondary modules may wish to review Raghunathan et al. (2001)⁹ and access the corresponding SAS-compatible imputation software at <http://www.isr.umich.edu/src/smp/ive/>. However, since the missing data problem for CDS-II secondary modules typically involves complete missing data for all variables, analysts may choose to develop a weighting adjustment to compensate for an MAR missing data mechanism. The algorithm for constructing these weighting adjustments should be similar to the one described above in Section 3 for the PCG/Child. Here, the logistic regression model of response propensity would be estimated based only on characteristics of eligible respondents and nonrespondents, excluding cases that were not eligible for the given module (see Table 4.2). The new attrition adjusted weight would be constructed as the product of the individual weight (CH01PRWT) and reciprocal of the estimated probability of response in the secondary interview.

⁹ Raghunathan, T.E., Lepkowski, J.M., VanHoewyk, J., and Solenberger, P., (2001). "A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models". *Survey Methodology*: Vol. 27, No.1. pp. 85-95. Statistics Canada.

Chapter 5 – Data Collection Procedures

The data collection for CDS-II is generally described in this chapter. More specific details on administration of assessments, anthropometric measures, and time diaries are described in each of their individual chapters that follow.

Data Collection for the CDS-II began in October of 2002 and extended through May of 2003. We employed 119 interviewers on the study who were extensively trained in all aspects of general interviewing techniques and study-specific data collection procedures, including unique protocols to administering the Woodcock Johnson tests of achievements, measuring height and weight, conducting the HOME Scale home and neighborhood observations, administering time diaries, and working with children of all ages. Approximately two-thirds of the interviewers had already worked at least one field study at the Survey Research Center at the University of Michigan prior to working on CDS.

CDS employed a mix-mode interviewing approach to increase flexibility in scheduling interviews. As mentioned in Chapter 2, each CDS sample child could have up to eight modules of data collected from three different family members (primary and secondary caregivers; sample child) and a school resource (teacher and/or school administrative data). The PCG interview was designated as the priority interview, and could be administered in-person or over the telephone, at the convenience of the respondent. The other interviews could be juggled depending on the respondent’s schedule (OCG, Child, PCG household) and could be conducted in different modes. Time diaries and other caregiver interviews could be administered in-person or over the telephone but had the added option of self-administration. Child interview components (assessments, height and weight measurements, and child CAPI/ACASI interviews) were conducted in-person.

Table 5.1. The CDS-II Survey Design

Respondent	Module	Mode
PCG ¹⁰	Child-Level Interview	In Person or Telephone
	Family-Level Interview	
OCG	Child-Level Interview	In Person or Telephone
	Family-Level Interview	
Child (8+ Yrs)	Child Interview	In Person
Child (All Ages)	Time Diary	Mail Ahead; In Person
	Child Assessments	In Person
	NCES CCD/PSS Data	Secondary Data Links
Teacher (Elementary School)	Elementary School Interview	Telephone Interview

¹⁰ See Chapter 2 for information about the selection protocol for the PCG and OCG.

Participating families received between \$95 and \$115 for their contributions to CDS, depending on the number of CDS sample children in the family. Primary Caregivers received \$75 for their interviews; each child in the household received \$20 for participation. Elementary school teachers for whom the primary caregiver gave permission to contact were interviewed by telephone and received \$20 for their participation.

CDS employed a number of flexible contact approaches to making the survey request. Emphasis was placed on strategizing how/when the interviewers attempted contact with the respondent over setting a pre-defined number of maximum contacts. Interviewers tailored their survey introductions to the respondent's concerns and individual situations, rather than using a pre-defined script. During field interviewer training, interviewers were given guidelines and hands-on practice. This information and experience helped the interviewers address issues that arose on a family-to-family basis, and helped make certain that the family felt comfortable with the wide array of questions and instruments administered to both the caregivers and children. Interviewers were given specialized training and a handbook to facilitate response to reluctance to participate and also were given the option of working with the research staff on tailored persuasion letters for resistant situations. The refusal conversion rate for CDS-II was 44%, resulting in less than 7% of the families in a final refusal.

The biggest source of support in reaching response rate goals, however, was the close relationship to the main PSID study. During the five-year gap in-between CDS-I and CDS-II, PSID interviewed the families of the CDS children (in 1999 and 2001) and also kept in touch with them through respondent postcard and report mailings. The PSID achieved a 96% response rate in the prior wave (2001), retaining a high proportion of the sample, which in turn, helped to maintain the CDS sample. CDS was able to capitalize on the contact information and efforts of that wave of the main study.

During CDS-II field period, however, some tracking was required. A specialized expert field tracking team carried out the more involved or complex tracking efforts for the hard-to-locate families. Just over 10% of the sample required tracking at either the interviewer or tracking-team level. At the end of the data collection, only 1% of the overall sample was unable to be located.

Team-based interviewing was used for 7% of the completed households for situations such as: traveling interviewers into areas, interviewing in dangerous neighborhoods, and special family situations whereby the interviewer thought a team-approach would help with gaining cooperation as it reduced the amount of time spent in the respondent's home.

Chapter 6 – Primary Caregiver Interviews: Topical Guide

As described earlier, the PCG-Child interview focused on the primary caregiver’s report of the sample child’s health, schooling, family relationships, and extra-curricular activities, while the Household interview focused on the Primary Caregiver characteristics and behaviors and the household environment. This chapter provides more detailed information about the key measures in the Primary Caregiver interviews. Appendix A summarizes these measures and provides the question number references for both CDS-I and CDS-II. Cross-wave references for individual items are available in the Cross-Year Index selection within the Data Center. Questionnaires can also be obtained on the web at: <http://psidonline.isr.umich.edu/CDS/quescodetable.html>.

Health

In the first section of the CDS-II PCG-Child interview, we ask the PCG a series of questions assessing the sample child’s health status, use of health care services, and costs associated with health care. The section starts out with a query about breastfeeding practices and age stopped breastfeeding the child. These items are the same as asked in CDS-I. Then the interview launches into a series of questions about health care utilization items, also included in CDS-I and modeled from the NLSY79. Items A3-A3b, A5A8b, and A10 query the PCG about: frequency of overnight hospital stays, date of most recent stay, reason for this hospital stay, month and year of last doctor visits for illness, injury, and mental health and reason for each visit, and status of shots and immunizations. New to CDS in the second wave, we coded responses to the open-ended questions related to reasons for doctor visits. These data are available in the Data Center.

There are several groupings of questions within the health section that assess health status. These measures included a global/general health rating, indication of limitations in activities that result from health condition, and diagnosis of chronic conditions. More specifically, at A11, the PCG rated the sample child’s health on a scale of ‘1’, meaning “Excellent” to ‘5’, meaning “Poor”. New to CDS-II was a parallel youth-reported item. Children who were 10 years and older reported their own health using this same item in the Audio Computer-Assisted Self-Interview. At A9, we gathered information from the PCG about activity limitations of the sample child, measured as three “YES/NO” items that queried the PCG if the child had any physical/mental conditions that limited him/her in play, school attendance, or school work. These items are also included in the child self-report interview for children 10 years and older. At A4a-p, the PCG reported whether a doctor ever diagnosed the child with any of 16 chronic illnesses. CDS-II used the same list as CDS-I, except “developmental delay” and “learning disability” were combined into one item, and allergies were added. Individual items for CDS-II were:

- An epileptic fit or convulsion
- Asthma
- Diabetes
- More than 3 ear infections in a year
- Speech impairment or delay
- Serious hearing difficulty or deafness
- Serious difficulty seeing or blindness
- Mental retardation

- A serious emotional disturbance
- Anemia or iron deficiency
- Elevated levels of lead in the blood
- Orthopedic impairment
- Developmental problems, such as developmental delay or learning disability
- Autism
- Hyperactivity, ADHD, or ADD
- Allergies
- Other condition

CDS-II added a series of questions (A12-A16) that focused on the PCG's report of the sample child's experience with asthma-related symptoms in the 12 months prior to the interview:

- Wheezing with exercise or playing hard
- Frequency of wheezing
- Hospital visits for asthma or wheezing
- Doctor visits for asthma or wheezing
- Missed school because of asthma or wheezing in the past 12 months

Finally, some information about health care costs to augment similar data collected in the main PSID family interview was collected. At A17-A24, the PCG reported on the amount paid for dental and medical insurance and for out-of-pocket dental and medical expenditures in the prior 12 months to the interview and who paid these expenses. This series was expanded from CDS-I.

Education

Information about the child's schooling and school environment was collected in a number of places in the CDS-II, including the teacher interview for the sample children in elementary school, the linkages to the U.S. Department of Education's Common Core of Data, and in the PCG Child Interview. In Section B of the PCG interview, the PCG reported on her expectations for the sample child's educational attainment (B1-B3), enrollment in school and special/gifted programs (B4-B6), the number of days in the past 12 months that the sample child missed school because of illness and injury (B7), enrollment in federal breakfast/ lunch program (B8-B9), number of times the CDS sample child changed school in current school year (B10), private and public school attendance, school characteristics, fees, and tuition (B11-B22); age at which the sample child began kindergarten, and if the sample child was ever suspended or ever drop out of school (B28).

PCGs (and OCGs) are asked about their participation in the child's education— specifically, their interactions with teachers, counselors, administrators and volunteerism at the school; the frequency and nature of communication between parent and child about school. These questions were modeled after questions in the NELS: 88 parent baseline survey.

Behavior Problems Index

Both waves of the CDS included the Behavior Problems Index (BPI) to assess the incidence and severity of behavioral problems among children.¹¹ The Peterson and Zill (1986) measure was developed from the more extensive Achenbach Behavior Problems Checklist.¹² Beginning at question B29a in the PCG-Child Interview, the PCG reported whether each behavior was “often”, “sometimes”, or “never” true of the sample child. Note that in CDS-I, the Secondary Caregiver, Absent Father, and Teacher were asked the same questions about the sample children. Behaviors were divided into two subscales, a measure of externalizing or aggressive behavior and a measure of internalizing, withdrawn or sad behavior. B29aa-dd were added by the NLSY79 staff and included in the CDS to provide additional measurement for the withdrawn behavior scale. Table 6.1 shows how the items map onto the external and internal subscales in the CDS-II data.

We performed a confirmatory factor analysis on our two expected subscales. The results showed that the items grouped into these two factors quite well, with one variable overlapping on both subscales, as did in CDS-I, and two variables not loading at all. We constructed an overall or total BPI score, using all 30 items that loaded, as well as separate scores for each of the two subscales, internal or withdrawn and external or aggressive behaviors. Before scoring, the individual items were recoded such that a score of “2” became “0” and a score of “2” or “3” became a “2”. Scores for the total BPI and Externalizing and Internalizing are sum scores. Higher scores on these measures imply a greater level of behavior problems. Cases were included if they had valid data on all of the variables contributing to the BPI Index: 2872 cases met this criterion for the BPI total scale; 2,893 met this criterion for the BPI externalizing subscale and 2880 met this criterion for the internalizing subscale. Please note that in CDS-I, the items were simply averaged for a BPI score.

Table 6.1. Survey Items in the Behavior Problems Index

B29. For the next set of statements, decide whether they are not true, sometimes true, or often true, of (CHILD)’s behavior. He/She ...	External	Internal	Total
a. ... has sudden changes in mood or feeling.	<input type="checkbox"/>		<input type="checkbox"/>
b. ... feels or complains that no one loves him/her.		<input type="checkbox"/>	<input type="checkbox"/>
c. ... is rather high strung, tense and nervous.		<input type="checkbox"/>	<input type="checkbox"/>
d. ... cheats or tells lies.	<input type="checkbox"/>		<input type="checkbox"/>
e. ... is too fearful or anxious.		<input type="checkbox"/>	<input type="checkbox"/>
f. ... argues too much.	<input type="checkbox"/>		<input type="checkbox"/>
g. ... has difficulty concentrating, cannot pay attention for long.	<input type="checkbox"/>		<input type="checkbox"/>
h. ... is easily confused, seems to be in a fog.		<input type="checkbox"/>	<input type="checkbox"/>
i. ... bullies or is cruel or mean to others.	<input type="checkbox"/>		<input type="checkbox"/>
j. ... is disobedient.	<input type="checkbox"/>		<input type="checkbox"/>

¹¹ Peterson, J.L., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavioral problems in children. *Journal of Marriage and the Family*, 48(2), 295-307.

¹² Achenbach, T., & Edelbrock, C. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs of the Society for Research in Child Development*, 46(1), No. 188. pp. 82.

B29. For the next set of statements, decide whether they are not true, sometimes true, or often true, of (CHILD)'s behavior. He/She ...	External	Internal	Total
k. ... does not seem to feel sorry after misbehaves.	<input type="checkbox"/>		<input type="checkbox"/>
l. ... has trouble getting along with other people (his/her) age.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
m. ... is impulsive, or acts without thinking.	<input type="checkbox"/>		<input type="checkbox"/>
n. ... feels worthless or inferior.		<input type="checkbox"/>	<input type="checkbox"/>
o. ... is not liked by other people (his/her) age		<input type="checkbox"/>	<input type="checkbox"/>
p. ... has a lot of difficulty getting (his/her) mind off certain thoughts.		<input type="checkbox"/>	<input type="checkbox"/>
q. ... is restless or overly active, cannot sit still.	<input type="checkbox"/>		<input type="checkbox"/>
r. ... is stubborn, sullen, or irritable.	<input type="checkbox"/>		<input type="checkbox"/>
s. ... has a very strong temper and loses it easily.	<input type="checkbox"/>		<input type="checkbox"/>
t. ... is unhappy, sad or depressed.		<input type="checkbox"/>	<input type="checkbox"/>
u. ... is withdrawn, does not get involved with others.		<input type="checkbox"/>	<input type="checkbox"/>
v. ... breaks things on purpose or deliberately destroys (his/her) own or another's things.	<input type="checkbox"/>		<input type="checkbox"/>
w. ... clings to adults.			
x. ... cries too much.		<input type="checkbox"/>	<input type="checkbox"/>
y. ... demands a lot of attention.	<input type="checkbox"/>		<input type="checkbox"/>
z. ... is too dependent on others.		<input type="checkbox"/>	<input type="checkbox"/>
aa. ... feels others are out to get (him/her).		<input type="checkbox"/>	<input type="checkbox"/>
bb. ... hangs around with kids who get into trouble.	<input type="checkbox"/>		<input type="checkbox"/>
cc. ... is secretive, keeps things to (himself/herself).			
dd. ... worries too much.		<input type="checkbox"/>	<input type="checkbox"/>
ee. ... is disobedient at school.	<input type="checkbox"/>		<input type="checkbox"/>
ff. ... has trouble getting along with teachers.	<input type="checkbox"/>		<input type="checkbox"/>
Number of Items	17	14	30
Cronbach's alpha	.86	.83	.90
Unweighted N	2,893	2,880	2,872

The Positive Behaviors Scale

In addition to measures of problematic child behaviors, as described above, both waves of the CDS included measures of more favorable development, including the Positive Behaviors Scale. The Positive Behavior Scale (B30), originally developed for the New Chance Evaluation (Polit, 1998¹³), measures the positive aspects of children's lives, including self-esteem, social competence, self-control, obedience/compliance, and persistence. The original scale consisted of 25 items evaluated on a 10-point scale, from "Not At All Like My Child" to "Very Much Like My Child". The scale used in the CDS consisted of 10 items that were selected by Child Trends, Inc., for use in the JOBS evaluation study. The question reads: "Please tell me how much each statement applies to (CHILD) on a scale from 1-5, where one means "not at all like your child," and five means "totally like your child," and two, three and four are somewhere in between."

¹³ Polit, D. (1998). The Positive Behavior Scale. Saratoga Springs, NY: Humanalysis.

We performed a confirmatory factor analysis on the set of 10 items to verify that the items mapped together. Factor loadings are provided in Table 6.2 below; reliability coefficient was .82. The scale is a mean score of all 10 items. Cases were included if they had valid data on approximately 75% of the variables contributing to the Positive Behavior Scale. Two cases did not meet this inclusion criterion and were dropped from the analysis, leaving a total of 2,905 cases. The (unweighted) mean score on the Positive Behavior Scale (Variable “POSBEH02” in the Data Center) was 4.13, with a range of 1 to 5, and a standard deviation of .60. Cronbach’s alpha = .82.

Table 6. 2 Factor Loadings for Positive Behavior Scale

B30. Please tell me how much each statement applies to (CHILD) on a scale from 1-5, where one means “not at all like your child,” and five means “totally like your child,” and two, three and four are somewhere in between.	Factor Loading
a. Is cheerful, happy.	.596
b. Waits (his/her) turn in games and other activities.	.552
c. Does neat, careful work.	.465
d. Is curious and exploring, likes new experiences.	.381
e. Thinks before (he/she) acts, is not impulsive.	.529
f. Gets along well with other people (his/her) age.	.759
g. Usually does what you tell (him/her) to do.	.593
h. Can get over being upset quickly.	.508
i. Is admired and well-liked by other people (his/her) age.	.774
j. Tries to do things for (himself/herself), is self-reliant.	.479
Number of Items	10
Cronbach’s alpha	.82
Unweighted N	2905

Household Tasks

The questions concerning household tasks that the primary caregiver did with the sample child (B31) were intended to be one measure of cognitive stimulation in the family environment. These items were modeled after similar items in the National Survey of Families and Households and the National Longitudinal Survey of Youth. CDS-II used the same set of items as CDS-I, but added “do yard work or gardening (together)”. The item “look at books with (CHILD)” was modified in CDS-II to additionally include “...or talk with (CHILD) about books (he/she) has read”.

Parental Monitoring

Stattin & Kerr (2000)¹⁴ define parental monitoring as “a set of correlated parenting behaviors involving attention to and tracking of the child’s whereabouts, activities, and adaptations”. This definition emphasized both parents’ knowledge/awareness of their children’s activities as well as their active monitoring of them. Based on these two components, CDS-II included a number of items on parental knowledge and active monitoring at B33, B34, and B35a-B35r1 in the PCG-Child and, from the child’s perspective, L23-L28 in Child interviews.

¹⁴ Stattin, H., & Kerr, M. (2000). Parental monitoring: A reinterpretation. *Child Development*, 71, 1072-1085.

From the caregiver’s perspective, measures of the caregivers’ knowledge of the child’s whereabouts, activities, and associations can be found at B33-B34:

PCG-Child Interview	B33.	How many of (CHILD)’s close friends do you know by sight and by first and last name?
	B33a.	How many of (CHILD)’s close friends’ parents do you know by sight and by first and last name?
	B33c.	About how often did (CHILD) do things with (his/her) friends outside of school in the last month?
	B34.	About how often do you know who (CHILD) is with when (he/she) is not at home?

From the child’s perspective, measures of the caregiver’s knowledge of the child’s activities can be found in the Child Interview, at L23-L25:

Child Interview	L23.	Do your parents know what you do during your free time?
	L24.	Do your parents know which friends you hang out with during your free time?
	L25.	Do your parents know what you spend your money on?

One source of information about children’s activities is the children themselves. They may choose to disclose their activities or whereabouts without being first asked. Child Disclosure items can be found in the child interview, at L26-L28.

Child Interview	L26.	Do you keep a lot of secrets from your parents about what you do during your free time?
	L27.	Do you hide a lot from your parents about what you do during nights and weekends?
	L28.	If you are out at night, when you get home, do you tell your parents what you did that evening?

Parents or caregivers may set rules about where, how, and with whom the child spends time, and, in practice, they may enforce those rules to varying degrees, if at all. The PCGs reported on household rules and enforcement practices of those rules at B35 in the PCG Child Interview. We asked if the caregiver had rules about television watching, bedtime, snacking, homework, friends, dating, curfews, and work hours. We then asked if the caregiver discussed and enforced each of these rules with their children.

Child Care

The PSID main interview (see Section F, 2001 interview) contained some questions about the family's use of child care and its associated costs. In Section C of the CDS-II primary caregiver interview, we asked the primary caregiver about all child care arrangements used during weekday, weekend, and summer. We asked the type, frequency of use, and costs of arrangements for children up to and after Kindergarten. When CDS-I and CDS-II data are put together, there should be a fairly complete picture of formal and informal child care arrangements used from birth up to the age of the child at the time of the CDS-II interview. Information on arrangements used during weekday, weekend, and summer were obtained separately. Types of arrangements queried were expanded in CDS-II to capture full range of options used through adolescence: care by relatives under and over 13 years of age; care by non-relatives in or out of the home; Head Start Program, Pre-kindergarten program, nursery school, preschool, or child care center; before or after-school program, extra-curricular activities, and part-time jobs

Absent Parents

Questions on absent parent interactions obtained information about the frequency and types of activities in which absent parents are involved with their children (D15a-d; D31a-d in PCG-Child Interview) and any conflict between the resident parent and absent parent (D14a-h; D30a-h). These items were drawn from the National Survey of Families and Households, supplemented by additional items from other scales. It measures the degree of conflict between the parent with whom the child resides and the absent parent. CDS-II included the same items as CDS-I except two items were dropped: "how the absent parent spends money" and "disciplining (CHILD)".

Sample Child's Extra Curricular Activities

In the primary caregiver interview, we ask about the sample child's involvement in a number of school-based and community-based activities. The questions are located in sections F (children 3-5 years of age), G (children 6-9 years), and H (youth 10 years and older). We ask about participation in lessons in dance, arts, hobbies, or academics; school and summer sports, community groups such as scouts, library story hour, YMCA, or neighborhood groups, and religious worship and social activities. For each activity, the PCG provides information about costs associated with each activity.

Additional parallel measures are found in the time diary and child interview modules. The time diaries have a detailed set of codes for sports, after-school, and community group participation on the randomly selected weekday and weekend day. The child reports for him/herself in the Child Interview – Section K.

Child-Specific Expenditures and Savings

In CDS-II, we expanded the measures on expenditures and savings for each sample child over the content included in the CDS-I. In Section H, beginning at H23, the primary caregiver reported on amount of money spent by household and non-household members, including absent parents on: gifts, vacations, school supplies, clothing, car insurance, car payments, car maintenance, and food. The primary caregiver also reported on amount of money in savings and assets that are in the child's name.

HOME-SF

The Home Observation for Measurement of the Environment-Short Form measured cognitive stimulation and emotional support that parents provide to their children. This short form scale is based on the original Caldwell and Bradley HOME Inventory (Caldwell & Bradley, 1984¹⁵) but modified for the NLSY79. The particular items used in the CDS closely reflect (but do not exactly replicate) the measures in the NLSY79. The HOME-SF items included both parent/caregiver-reported items and interviewer observations of the home and neighborhood environment. The measures can be used as independent variables to help explain developmental outcomes for children and youth, or can be used as an outcome measures themselves, to better understand relationships among the quality of environment and earlier familial and maternal characteristics.

In CDS-II, we included the same age-appropriate items used in CDS-I (CDS-I also included infant module), and supplemented this set with additional interviewer observations and items from the Project on Human Development in Chicago Neighborhoods (PHDCN). In CDS-II, items from the Early Adolescent Scale were administered for sample children 10 years and older. The appendix to this guide shows the mapping of CDS items onto the NLSY79 and PHDCN.

All HOME-SF items are located in the primary caregiver-child and primary caregiver-household interviews, as well as in interviewer observations of caregiver-child interactions and the neighborhood environment. HOME-SF questions, which vary by child age, were asked among children of three age groups:

- Early Childhood (EC) - 3 to 5 year olds
- Middle Childhood (MC) -6 to 9 year olds
- Early Adolescent (EA) - 10 year olds and older

For each age group we constructed the total score and a subscore for cognitive stimulation. In order to construct scores, we first recoded the individual items into dichotomous variables, with values of zero and one. The recoding was guided by specifics from the NLSY79¹⁶ and detailed in the appendix. Items with scores of zero and one were then summed, by age group. All HOME-SF variables have one assigned decimal place.

Missing data on raw items were imputed prior to recoding of zero or one. In the imputation procedures, all missing data were replaced by mean, median, or mode, depending on the variable type and its distribution. Most of the variables had low levels of missing data; however, on a few input variables, the number of missing data was so significant that we decided to exclude them when constructing the HOME-SF scores. This procedure affected the following variables: Q21E4 and Q21E4A (27% of the cases missing, mainly due to children residing without a father) and all input variables from the interviewer observations of mother-child interactions (18% of the cases missing) and neighborhood environment (14% of the cases missing). Missing data on the interviewer observations are largely attributed to a high geographic distribution of families in the PSID-CDS sample, making some families out of range of our interviewers. If a family was not close enough to at least two other eligible families to justify the high cost of traveling, we did not make an in-person visit to the home, and relied on telephone and mail methods to obtain the other parts of the interview.

¹⁵ Caldwell, B., & Bradley, R. Home Observation for Measurement of the Environment. Little Rock, AR: University of Arkansas at Little Rock.

¹⁶ <http://www.bls.gov/nls/97guide/nls97usg.htm>

We provide proration flag variables on the PCG-Child file that represent the number of imputations conducted for each scale (HT1IMP02, CGIMP02, HT2IMP02, CG2MP02, HT3IMP02, and CG3MP02). A flag variable with a score of zero indicates that all input items were non-missing. Note, however, that sometimes the number of imputations is higher than the number of dichotomous variables used to construct the scale. This is caused by the fact that some dichotomous variables were created from more than one input item. Note that NLSY79 provides scores for the Emotional Support and Cognitive Stimulation subscales only for cases where no more than one item is missing. In CDS-II, we constructed the subscales regardless of the number of items requiring imputation – the flag variables allow you to set your own inclusion criteria based on the number imputations.

In order to properly construct the HOME-SF scales, some of the input items regarding caregiver reaction to child's swearing/hitting – discipline items-that had originally been coded as “other” were reviewed, and all verbatim comments recoded. The discipline item was generally scored as 0 if it was classified as a “harsh” response, such as a physical/severe response or extensive/excessive deprivation. Non-harsh discipline items were coded as 1. There were four cases in age group 10+ that were originally coded as “Other” in Release 2. Those cases led to the creation of three new variables in Release 3 (Q21H17P – “Eat soap”, Q21H17Q – “Counseling”, Q21H17R – “Make him my shadow/make him pay”) among which one – “eat soap” -was classified as a “harsh” response. There were no responses classified as “other” in the youngest age group, while in the middle age group (6-9 year olds) only one response was coded as “other,” though it was a non-harsh response. In CD-II, the following responses were categorized as harsh and scored as 0: “hit”, “spank”, “kick out of the house”, and “eat soap”.

While the CDS-II scores were constructed based on NLSY79 guidance, there were two questions that are part of NLSY79 scales that were not administrated in CDS-II: The question of how often the caregiver reads to the child was not asked in the 6-9 year old age group in CDS-II and the interviewer observation about the safety in the child's neighborhood (SLY CSOB-10C) was also not administrated in CD-II. As mentioned above, some variables were excluded in the CDS-II due to high rates of missing data.

Neighborhood Measures

The Primary Caregiver-Household module gathered some information about the sample child's neighborhood. At the beginning of Section J, the primary caregivers reported on length of residence in the neighborhood, neighborhood quality, cohesiveness, and safety.

After the visit to the home, the interviewer recorded characteristics of the physical environment of the home and neighborhood (K27-K39). These observations focused on the cleanliness and appearance of the home and street, amount of living space, noise, signs of drug and alcohol use, and safety. Interviewers were extensively trained in method for these observations.

Rosenberg Self-Esteem for Primary Caregivers

The Rosenberg Self-Esteem scale describes the degree of approval or disapproval toward oneself. The scale is widely used, with substantial documentation on its validity and reliability. Primary Caregivers reported on a series of 10 items at J9a-j, using response scale of 1-4, where 1 indicates “Strongly Disagree” and 4 indicates “Strongly Agree”. Prior to scoring, items J9c, J9e, J9h, J9i, J9j were reverse scored. A higher score designates higher self-esteem. Cases were included if they had valid response on approximately 75% of the items. There were 2,889 cases with a self-

esteem score.

The results of a confirmatory factor analysis showed that the items readily grouped together. Factor loadings are provided in Table 6.3. The scale is a mean score of all 10 items. The mean score on the Rosenberg Self-Esteem Scale (Variable “SLFEST02” in the Data Center) was 3.43, with a range of 1 to 4, and a standard deviation of .44. Cronbach’s alpha = .86.

Table 6.3. Factor Loadings for the Rosenberg Self-Esteem Scale

Question Item	Factor Loading
Feel Like a Failure 02	.675
Good Qualities 02	-.657
Positive Attitude 02	-.643
Feel Useless At Times 02	.641
Think I’m No Good 02	.639
Person of Worth 02	-.625
Satisfied With Self 02	-.607
Not Much to be Proud Of 02	.570
Want More Respect 02	.560
Do Things Well 02	-.547

Number of Items	10
Mean Score	3.43
Cronbach’s Alpha	.85
Unweighted N	2889

Pearlin Self-Efficacy Scale for Primary Caregivers

The Pearlin Self-Efficacy Scale assesses the extent to which people see themselves as having control over aspects of their lives. This measure consisted of four of the original seven Pearlin items (PCG-HH J10a-d) which were answered on a 4-point response scale (“Strongly Agree”, “Agree”, “Disagree”, “Strongly Disagree”). The original scale has been shown to exhibit reasonable internal reliability and good construct validity (see Pearlin et al, 1981¹⁷). The four items used for CDS-II were selected based on factor analysis of CDS-I data where all seven items were used. Items are averaged to produce an overall score.

Factor loadings from the confirmatory factor analysis are provided in Table 6.4. The scale is a mean score of all four items. The mean score on the Pearlin Self-Efficacy Scale (Variable “PEARLN02” in the Data Center) was 3.11, with a range of 1 to 4. Cronbach’s alpha = .79.

Table 6. 4. Factor Loadings for the Pearlin Self-Efficacy Scale

Item Label	Factor Loading
Feel Helpless 02	.774
Little Control 02	.707
Pushed Around 02	.683
Can’t Solve Probs 02	.630

Number of Items	4
Mean Score	3.11
Cronbach’s Alpha	.79
Unweighted N	2885

¹⁷ Pearlin, L., Lieberman, M., Menaghan, E., and Mullan, J. (1981). The stress process. *Journal of Health and Social Behavior*, 22, 227-356.

Measure of Non-Specific Psychological Distress for Primary Caregivers

The K-6 Non-Specific Psychological Distress Scale was developed by Ronald Kessler, Professor of Healthcare Policy at Harvard Medical School, with support from the U.S. National Center for Health Statistics. As described in more detail in Kessler et al. (2003¹⁸), the K-6 Non-Specific Psychological Distress Scale was designed to discriminate cases of serious mental illness from non-cases in a general population survey. The K-6 is now included in the core of the NHIS as well as in the National Household Survey on Drug Abuse.

The K-6 includes six items that queried the primary caregivers about how they have been feeling during the prior four weeks. Response items are based on a scale of 1-5, where 1 indicates “All of the Time” and 5 indicates “None of the Time”. The items are rescored, following the coding in the table on the next page, and summed. A score of 13 or higher indicates a potential for nonspecific distress. In the CDS-II data, K6_02 had a mean of 4.12, standard deviation of 3.71, range of 0-24, and total number of cases scored equal to 2,889.

Table 6.5. Recoding of the K-6 Scale

In the last 4 weeks, about how often did you feel ...	All of the Time	Most of the Time	Some of the Time	A Little of the Time	None of the Time
a. so sad that nothing could cheer you up?	4	3	2	1	0
b. nervous?	4	3	2	1	0
c. restless or fidgety?	4	3	2	1	0
d. hopeless?	4	3	2	1	0
e. everything was an effort?	4	3	2	1	0
f. worthless?	4	3	2	1	0

The K-6 included not only the six Likert scale questions in the scales, but additional questions about persistence and impairment that result from the symptoms. These additional questions are not required to score the K-6.

Parenting Attitudes, Styles, and Beliefs

The PCG-Child and PCG-Household interviews have a number of measures on parenting attitudes and experiences. Briefly, these variables are:

Parenting attitudes: types of compromises or life style adjustments to make life better for their children.

Gender roles: PCG perceived responsibilities of mother and father roles.

Aggravation in Parenting Scale (E8a-c) and (J15a-d) measures parenting stress that may result from changes in employment, income, and other factors in their lives.

Response to Poor Grades (G16a-j; H15a-j): These items query PCG on actions taken in response to poor grades. CDS-II used the same items as CDS-I.

¹⁸ Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S-L. T., Manderscheid, R. W., Walters, E. E., Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60(2), 184-189.

Encouragement of Reading (G18-G20): These items assess PCG encouragement of reading and availability of dictionary and encyclopedia in the home.

Family conflict: methods of conflict resolution among family members.

Economic strain: experiences of economic or financial stress and strain and practical responses to such pressures.

Work schedule: PCG work status; work schedule, and travel time one-way to the job site for the respondent's job or main job.

Chapter 7 – Child Interview: Topical Guide

Children who were at least eight years of age and older were eligible to participate in an in-home interview themselves. Almost all of the Child Interview was new to CDS-II, as more children were old enough to report on their feelings and behaviors themselves. The interview content was sub-divided into sections for children of different ages – the measurement resource table in Appendix A gives details on various measures throughout the child interview.

Interview Content for All Participating Children 8 Years and Older

For all children participating in an interview, we asked about their own self-concepts in math and reading, expectations for the future, connectedness to school, use of electronic media, and peer bullying. We assessed the frequency in which the child conversed with mother, father and friends about school, future plans, and friendships. Questions H5-6, in turn, queried the child about his/her closeness with family members and other adults. We asked the children to describe the frequency in which they experienced a range of health symptoms such as: feeling really sick, waking up tired, having skin problems, feeling dizzy, experiencing chest pains, headaches, muscle soreness, and stomachaches.

We generated three summary variables for the interview content that targeted all participating children: ability self concepts for reading and math, and peer bullying. Methodology for constructing these measures is described below.

Ability Self-Concepts

The Ability Self-Concepts items at E1-E20 reflect two scales developed and validated by Jacquelynne Eccles for the domains of math and reading¹⁹. The same items were used both in CDS-I and CDS-II, and are asked of children 8 years and older. We created summary variables for reading and math by first reverse scoring E6 and E16 so that the scores ranged from low to high, as other items. A confirmatory FA forced the analysis to two factors and generated a clear delineation: all of the reading items loaded on the Reading Self-Concept factor (Factor 1 on the table below, which accounted for 20% of the variance), while all of the math items loaded on the Math Self-Concept factor (Factor 2, which accounted for 19% of the variance). We then constructed an average score for each scale, and included a case if it had valid data on about 75% of the variables. Three cases did not meet the scaling criteria for both Reading and Math scales and were dropped, leaving a total of 2179 included cases. The scores ranged from 1.6 to 7 on the Reading Self-Concept Scale, with a mean of 5.29 and a standard deviation of 1.0. On Math Self-Concept Scale, the scores ranged from 1.30 to 7, with a mean of 4.92 and standard deviation of 1.01.

¹⁹ Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P., (1993). Age and gender difference in children's self-and task perceptions during elementary school. *Child Development*, 64, 830-847.

Table 7.1. Reading and Math Self Concept Factor Loadings

Variable	Self-Concept	
	Reading	Math
Reading Skill General Rating 02	.765	
Reading Compared to Other Skills 02	.743	
Achieve in Reading this Year 02	.717	
Reading Skill in Context Peers 02	.714	
Learning Something New in Reading 02	.646	
How Much like Reading 02	.586	
Interest in Reading 02	.572	
How Hard is Reading 02	.511	
Importance of Reading 02	.494	
How Useful is Reading 02	.430	
Math Skill General Rating 02		.763
Math Compared to Other Skills 02		.746
How Much like Math 02		.692
Math Skill in Context Peers 02		.670
Achieve in Math this Year 02		.614
How Hard is Math 02		.594
Interest in Math 02		.565
Learning Something New in Math 02		.527
Importance of Math 02		.359
How Useful is Math 02		.304
N	2179	2179
Mean Score	5.29	4.92
Alpha	.82	

Peer Bullying

We assessed frequency of peer victimization with four items at the beginning of Section H²⁰ (H1a-d). These four items were rolled into a mean score for all cases having valid data on at least 75% of the items.

A confirmatory factor analysis produced one factor that accounted for 31.24% of the variance and had strong correlations to each item. We then constructed an average score, and included a case if it had valid data on 75% of the variables. Seven cases did not meet the scaling criteria and were dropped, leaving a total of 2175 included cases. The scores ranged from 1.0 to 5.75, with a mean of 1.45 and standard deviation of .71. The alpha coefficient was .62. Factor loadings are shown on the next page.

²⁰ Kochenderfer, B. J. & Ladd, G. W. (1996). Peer victimization: Cause or consequence of school maladjustment? *Child Development*, 67, 1305-1317.

Table 7.2. Factor Loadings on the Peer Victimization Scale

Variable	Factor Loading
Picked on You 02	.657
Left Out of Things 02	.575
Hit You 02	.559
Taken Your Things 02	.419
N	2175
Mean score	1.45
Alpha	.62

Interview Content for Participating Youth 10 Years and Older

Children in the adolescent-stage additionally participated in an Audio-Computer Assisted Self-Interview (ACASI) component whereby the children listened to the questions themselves on a headset and then responded directly into the laptop. Children 10 years and older could participate in Section K that included: a parallel set of questions to the PCG that queried the child about time spent in after school and community activities, involvement of friends, hobbies, sports participation, and pro-social behaviors. We gathered some additional information about how the youth themselves rate and evaluate their health. We included the self-reported health question, a single item that asked the child to rate his/her own health on a scale of 1 (Excellent) -5 (Poor). This item has a parallel measure in the Primary Caregiver interview, where we asked the caregiver to describe the child’s general health status. We asked the youth about their practices for maintaining their weight, their dietary habits, exercise at and outside of school, any limitations in activities that they experience because of their health (parallel item to the ones in the Primary Caregiver Interview), and sleeping patterns.

Global Self Concept Scale

We created a summary variable for the Global Self Concept Scale by taking the mean of items K27-K32 if the case had valid data on a minimum of five out of the six items. Confirmatory factor analysis was conducted; factor loadings are shown in Table 7.3 below. The scale included 1735 cases – 13 did not meet inclusion criterion and were dropped. The scale scores ranged from 1.0 to 5.0, with an overall scale mean was 3.30 and standard deviation of 1.78. The scale alpha was .82.

Table 7.3. Factor Loadings for the Global Self Concept Scale, Children 10+ Years

Items	Factor Loading
Good Things about Me 02	.767
Good as Others 02	.715
Do Things As Well as Others 02	.654
Proud 02	.622
Do Things Well 02	.592
Others Think I Am Good 02	.588
N	1735
Mean Score	3.30
Alpha	.82

Interview Content for Participating Youth 12 Years and Older

For youth 12 years and older, we included additional questions administered by the interviewer as well as through the ACASI module. The interviewer-administered section (Section J) included questions about the youth's self-identified ethnicity, religious affiliation, importance of religion and time spent in worship and religious groups, and importance of spirituality. We asked a series of questions about current and summer employment, including type of job, which was then coded with census occupation and industry codes, characteristics of the job, including level of responsibility, autonomy, opportunity to learn new skills; satisfaction with the job, plans to continue with current / summer job, wage rate, and work hours. We additionally asked a series of questions about job values and expectations for future careers. Next, we asked a series of questions about expectations for the future – mainly, marriage, divorce, parenting, schooling, and economic security. These questions are carried forward into the Transition into Adulthood interview.

The final set of questions in Section J focuses courses that the youth took in junior high / middle school and high school. We asked for a complete list of courses for two terms prior to the interview, exact titles of the courses, and grades.

The ACASI module for youth 12 years and older started with a series of items about savings and expenditure patterns. Youth are asked about allowances, savings accounts and goals for savings, particularly future schooling, and spending on a variety of items such as contributions to the main family for household expenses, car expenses, and entertainment.

The next long series of questions is at L11, and focuses on delinquent behaviors. Adolescents indicated the frequency of engagement in 10 different behaviors in the past six months that range in severity from staying out past curfew to harming others badly enough to require medical attention.

At L12-L15, we collected information about experiences with tobacco, alcohol, and drugs. For tobacco, alcohol, and drugs each, adolescents reported if they ever tried, age at which they first tried, frequency of use, and peer use. These questions were adapted from the ADD Health study.

The set of items L17a-b, L17d, L17g, L17i, L17j; L18a-b, L18d, L18g, L18i, L18j were adapted from work by B. Barber (1996) intends to measure psychological control, as reported by the child²¹. Six out of the original eight items were retained for CDS-II.

Children's Depression Inventory Short Form (CDI:S)

The Children's Depression Inventory (CDI)²² was developed by Maria Kovacs. The measure is designed as a diagnostic instrument that quantifies depressive symptomatology of children 7 to 17 years of age. It is an established measure, copyrighted by the Multi-Health Systems Incorporated and has been validated with normative populations of children 7-17 years of age. At L16, we included the short form of the Children's Depression Inventory (CDI), where youth reported on symptoms of depression during the two weeks prior to the interview.

²¹ For more details, see: Barber, B. K. (1996). Parental psychological control: Revisiting a neglected construct. *Child Development*, 67, 3296-3319.

²² Kovacs, M. (1992). *The Children's Depression Inventory (CDI) Technical Manual*. North Tonawanda, NY: MHS.

Social Initiative

The items at L22a-e were adapted from Bachman et al (1993) and B. Barber (1996) and are designed to measure social initiative. Five out of the original 13 items were retained for CDS-II based on Barber's recommendations. This scale was given to only the children in the sample who were at least 12 years of age. We constructed a mean score for the Social Initiative Scale and included the case if it had valid data on a minimum of four out of the five items. There were a total of 1267 cases included in this analysis. Confirmatory factor analysis had an overall of 3.61, with a standard deviation of 1.27 and range of scores from 1 to 6. The alpha coefficient for the scale was .81. Factor loadings are shown in Table 7.4 below.

Table 7.4. Factor Loadings for the Social Initiative Scale

Variable	Factor Loading
Joke with Teachers 02	.720
Converse with Adults 02	.712
Class Discussions 02	.688
Talk to Teachers 02	.665
Ask Questions 02	.608
N	1267
Mean Score	3.61
Scale Alpha Coefficient	.81

Subjective Well-Being

Questions L37a-c; L37d-h; L37k-n were adapted from MacArthur MIDUS Youth²³ and were designed to measure several dimensions of Subjective Well-Being among adolescents in CDS-II: Emotional Well-Being (L37a-c); Social Well Being (L37d-h); and Psychological Well Being (L37k-n). These questions were self-administered for youth 12 years and older. We constructed the emotional, social, and psychological well-being measures from the L37 items as mean scores if the respondent provided response on 75% of the items within each subscale. The continuous measure of flourishing is the sum of three subjective well-being measures. Descriptive statistics are provided in Table 7.5 below.

Table 7.5. Descriptive Statistics for the Subjective Well-Being Scales

Var Name	Variable Label	N	Mean	St Dev	Score Range
EWB02	Emotional Well-Being 02	1291	4.70	1.04	1-6
SWB02	Social Well-Being 02	1266	3.35	1.22	1-6
PWB02	Psychological Well-Being 02	1283	4.52	1.07	1-6
CONFL02	Continuous Flourishing 02	1257	12.58	2.78	3.2-18

²³ For more details, see: <http://midmac.med.harvard.edu/>

Chapter 8 Assessments

The CDS-II assessment module included measures of the child's height and weight, using tape measures and scales, and measures of the child's cognitive abilities, using the Woodcock-Johnson Revised Test of Achievement and WISC-III Digit Span. This chapter gives an overview of the procedures used for collecting these data.

Height and Weight Measures

In CDS-I, we gathered reported weight of the child from the Primary Caregiver and interviewer-measured height of the child, using a rafter square and tape measure (see Section A of the CDS-I Primary Caregiver-Child interview). In CDS-II, we attempted to measure and weigh all of the children and youth for whom we were able to get an in-person interview. In the assessment questionnaire for CDS-II you will find question items for (a) in-person measurement / weighing by the interviewer; or (b) self-report height and weight. Body Mass Index (BMI) and percentile rank were calculated for all children with height and weight data.

Administration Procedures

The procedure for measuring height and weight was as follows: If the child was under 6 years, interviewers asked the PCG or a family member for assistance in taking off the child's shoes, helping the child onto the scale, and encouraging the child to stand still. Directions were given directly to both the child and family member assisting the interviewer. If the child was 6 years or older, the interviewer instructed the child / youth directly and asked a family member for assistance only if necessary. If the child or youth refused to be measured, the interviewer asked them or the primary caregiver (depending on age) to report height and weight, or best estimate of, and the date the child was last measured.

To measure the child or youth's height, the interviewer asked them to take off their shoes and stand against a wall or door. The interviewer placed a Post-it note on the wall right above the child's height, and then positioned a rafter's square against the wall directly over the child's head. A mark was placed on the Post-it note at the child's height. When the child moved away from the wall, the interviewer positioned the tape measure under a door jam or her foot and measured from the floor to the mark on the Post-it.

To weigh the child or youth, the interviewer asked them to empty their pockets and remove their shoes, activated the scale and set it to zero. Once the scale showed "0.0", the child stepped onto the scale and stood in the center of the scale for 5-10 seconds. The same brand/model of scale was used for all children:

"Strain-Gauge Lithium Bath Scales" "THINNER MS-7" model
Measurement Specialities, Inc.
80 Little Falls Road Fairfield, NJ 07004
<http://www.msiusa.com/default/index.asp>

Scale Features included:

- High Contrast Digital Display
- Four-digit readout in .5LB increments
- Automatic Zero Set
- Non-slip Platform
- Snap-on carpet feet
- Lifetime lithium battery

BMI-for-Age

While Body Mass Index (BMI) for adults is based simply on height and weight, BMI for children and teens, also referred to as “BMI-for-Age”, additionally takes into account age and gender. The reason is that children's body fatness changes over time as they grow, and boys and girls differ in their body fatness as they develop. BMI-for-age is plotted on gender specific growth charts for children and teens 2 – 20 years of age. You can obtain more information on the CDC Growth Charts at: <http://www.cdc.gov/growthcharts/>.

Each of the CDC BMI-for-age gender specific charts contains a series of curved lines indicating specific percentiles. The percentile cutoff points identify underweight and overweight in children:

Underweight BMI-for-age < 5th percentile

At risk of overweight BMI-for-age 85th percentile to < 95th percentile

Overweight BMI-for-age > 95th percentile

In the CDS-II data, we have included BMI score and percentile score for each child/youth. BMI was calculated using pounds and inches with this equation:

$$\text{BMI} = \frac{\text{Weight in Pounds}}{(\text{Height in inches}) \times (\text{Height in inches})} \times 703$$

Percentile categories were calculated based on the CDC growth charts referenced above. When the CDSII data are used with the sample weight, the distribution of girls and boys in the weight percentile categories closely parallels the finding from the NHANES. See Table 1 for distribution of CDS-II and NHANES data, for children and youth 6-18 years.

Table 8. 1 Weighted Percentile by Child’s Sex

Sex	Underwt	Normal	Overwt	Overwt	Risk of Total
Male					
CDS-II	3.8	58.89	15.68	21.62	49.41
NHANES	3.6	63.91	14.45	18.04	50.67
Female					
CDS-II	2.55	66.49	15.61	15.35	50.59
NHANES	3.01	66.23	15.86	14.90	49.33
Total	3.17	62.74	15.65	18.45	100.0

Woodcock Johnson Tests of Achievement

Description of the Test

Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R) is a well-established and respected measure that provides researchers with information on several dimensions of intellectual ability, including current developmental status, degree of mastery in reading and mathematics, and group standing (either age or grade group). In the CDS-I and CDS-II, we selected three subtests as a measure of reading and math achievement: the Letter-Word, the Passage Comprehension, and the Applied Problems tests (the Calculation test was additionally administered in CDS-I). These scales can be used individually, or in the case of the four subscales, combined to create scores for Broad Reading and Broad Math. When applicable, the Spanish version of the WJ-R (Batería-R, Form A), was used for children whose primary language was Spanish.

Administration

The Woodcock-Johnson Revised (WJ-R) Tests of Achievement have standardized administrative protocols. For respondents under 6 years, the interviewer administered two subtests: Letter-Word Identification and Applied Problems. For respondents 6 years and older, the interviewer additionally administered the Passage Comprehension subtest. The paragraphs below provide information on the main procedures for administering these tests; the user is encouraged to obtain Woodcock-Johnson materials and training (see the Riverside Publishing website for details that help to fully understand the complete requirements for the standardized administration).

The Woodcock-Johnson (WJ-R) Test of Achievement is an ‘easel’ test, or a test with a response book that sits in front of the respondent. The interviewers placed the easel at an angle so that they and the respondents could both see the stimuli (pictures) simultaneously. Each test required that the interviewer administer it exactly as described in training and in their interviewer manual. Any deviation from these procedures invalidated the results.

Since the WJ-R can be used for respondents from ages 2 to 90 years, items in the WJ-R were arranged by difficulty for all persons between those ages. The easiest questions were presented first and the items became increasingly difficult as the respondent proceeded through the test. The interviewer started testing at the appropriate starting point based on education level of the child or youth as the general guideline. At the beginning of every subtest, usually on the first page, there was a chart organized by grade in school that informed the interviewer at what item they should start administering the test. To administer the test, the interviewer turned to the page where the starting item is located.

Raw scores were calculated for the WJ-R using basal and ceiling. The basal and ceiling criteria were created to limit the amount of time any one person spends on each subtest. When the respondent got six or more consecutive items correct, then they established their basal. The interviewer continued testing until the respondent established ceiling, which was six or more consecutive items incorrect and the end of the testing page has been reached.

Scoring

The Woodcock-Johnson Revised (WJ-R) Tests of Achievement have standardized scoring protocols. The tests are designed to provide a normative score that shows the CDS target child's reading and math abilities in comparison to national average for the child's age. The normed scores are constructed based on the target child's raw score on the test (essentially the number of correct items completed) and the child's age to the nearest month. Raw scores are charted on normative tables based on the child's age and what percentile the child falls into.

The first step in creating standardized test scores for the Woodcock-Johnson Revised (WJ-R) subtests was to enter all of the hardcopy scoring tables into electronic spreadsheet files so the scoring could be done by computer. The second step was to calculate the subjects' age at the time of testing. This was accomplished by comparing their birth date with the testing date. The syntax also incorporated commands to properly round days and months per WJ-R scoring rules. This variable was properly formatted to match the values given in the WJ-R scoring tables.

The next step was to create a programming loop to choose the correct version of the testing form used, either Form A, Form B, or Form Spanish. Then the raw score for each subtest could be matched to an intermediate number known as the W Score – this intermediate step allows standardization across the different forms of the WJ-R. The W Score is then compared against the Reference W Score, which is the average W Score for a child that age.

The Reference W has to be compared to child's W Score to determine the difference and whether that difference is positive or negative. One of the scoring tables provides the number of a lookup column, and which column is used depends on the child's age, the magnitude of the difference score, and whether the difference score was positive or negative. Furthermore, the value of the difference between the child's W Score and the Reference W also determines which lookup row will be used. Since negative difference scores would have been problematic in the look-up tables, their absolute values were taken and then added to 400 to make them easily distinguishable from positive scores. Likewise, a difference score of zero was converted to 900. The standardized score is found at the intersection of the proper row and column, and this information was then transferred back to the main data file.

For more information about standardized scoring and interpretation, please see: Woodcock, R.W., & Mather, N. (1989, 1990). WJ-R Tests of Achievement: Examiner's Manual. In R. W. Woodcock & M.B. Johnson, Woodcock-Johnson Psycho-Educational Battery – Revised. Allen, TX: DLM Teaching Resources.

WISC Digit Span Test for Short Term Memory

The WISC Digit Span Test for Short Term Memory²⁴ is a test of our ability to process information as it directly measures auditory short-term working memory. Digit span is the measure of how many sequential digits can be taken in, stored, processed, and recalled in the correct order. In CDS-I and II, the interviewer read a series of digits and the respondents then repeated the series. The series begin with two digits and keep increasing in length until eight digits, with two trials at each length. The test was repeated with a different set of digit series, but the second time, the respondent repeated the numbers backwards.

Administration Procedures

Interviewers were given the following instructions:

- (1) Read each digit span only once at an even rate of one digit per second.
- (2) Read Part A of a question; pause for response; then score.
- (3) Read Part B of a question; pause for response; then score.
- (4) If child does not respond after reading the question, do not encourage further.
- (5) Stop when child misses Part A and Part B of any one question.
- (6) To be scored as “correct”, no digits may be omitted or be in reversed order.
- (7) READ TO CHILD: I can only read these instructions to you one time. I cannot repeat them. I am going to say some numbers. Listen carefully, and when I am through, say them right after me.

SAY: “Ready” BEFORE EACH QUESTION.

Scoring

Items were scored by summing the number of sequences answered correctly. The data files include a digit-forward, digit-backward, and total score.

²⁴ Wechsler Intelligence Scale for children-Revised. Copyright © 1974 by the Psychological Corporation.

Chapter 9 – Time Diaries

The time diaries are one of the unique features of the CDS design and provide data of central importance to scientists in many disciplines. While the primary caregiver (PCG) and Child interviews include stylized questions about the children's structured and unstructured activities, and activities with parents and absent parents, the time diaries provide detailed accounting of the type, number, duration, and location of activities during sampled 24-hour days, beginning at midnight for one randomly sampled weekday and one randomly sampled weekend day. The time diaries additionally collect information on the social context of the activity by specifying with whom the child was doing the activity and who else was present, but not engaging.

Response rate for the time diaries was 88% in 2002/2003 (82% in 1997). The PCG completed the diary for the very young children; older child and adolescents were expected to complete the time diaries themselves or with the assistance of the PCG as needed.

With child-based weights, the time diaries give a representative national sample of children's activities and can also be converted into traditional time use aggregates, which are known to provide less biased estimates of time amounts than do self-reports of time in particular activities (Juster, Ono and Stafford, 2003²⁵). By nature of the CDS design, the rich time diary data can be used with the detailed information in the other CDS modules as a way to examine influences on key developmental outcomes and achievement patterns, and can also be used with PSID data to parcel out the ways in which family characteristics tie into children's activity choices and developmental outcomes.

Data Collection

We invited CDS children and youth to complete a set of time diaries as part of their involvement in the overall CDS data collection effort. Each participating child completed two time diaries, a weekday (Monday-Friday) and weekend day (Saturday or Sunday). These days were randomly selected when the interviewer completed the initial contact for the household. There was no substitution of diary days once they were assigned to the CDS child/youth.

In both waves, the CDS time diary was a paper diary that was mailed ahead before the scheduled interview, with instructions to complete prior to the interview. During the in-house CDS interview or by telephone, the interviewer reviewed and edited the diary with the child and primary caregiver or, in the situations whereby the diary was not completed in advance, the interviewer administered the diary as an interview. On average, the review/ interview time was just under 17 minutes per diary, per child. A summary of the data collection design is provided in Table 9.1 on the next page.

²⁵ Juster, F. Thomas; Ono, Hiromi, and Stafford, Frank (2003). An Assessment of Alternative Approaches to the Measures of Time Use. *Sociological Methodology*, 3319-54.

Design Features	Description
Sample	CDS sample aged 0-12 in 1997 and 5-18 years in 2002
Number of interviews	2,904 children in CDS-I 2,569 children in CDS-II
Response rate	82% in CDS-I 88% in CDS-II
Diary day assignment	One randomly selected WEEKDAY and WEEKEND day, beginning at midnight.
Substitution of diary day?	No substitution of days.
Contact Effort	Targeted diary day is fixed; day of the interview is scheduled at respondent's convenience.
Mode	Diary mailed ahead with interviewer review in person or by tele.
Avg. number of activities	20 activities per diary day, on average.
Interview length	16.9 minutes per diary. Including questions A1-A9 (in CDS-II)

Coding

We utilized the same interviewing procedures and coding protocols for the time use module in both waves of the CDS. With the older age-range of the children in the CDS-II (ages ranged from 5 to 18 years), the activity codes were somewhat expanded to capture more detail for participation in community, school, and active leisure activities. Codes were added to the “organizational activities” group to include participation in specific types of helping organizations and detailed extracurricular school clubs and organizations. CDS-II added details to the “sports and active leisure” group to capture specific sport or recreational activity.

In order to accommodate these new detailed codes but also allow for easy mapping back to the CDS-I data, CDS-II created a four-digit activity code whereby the first three digits reflect the 1997 code. If the code was expanded to capture more specific detail, a sequential number beginning with “2” was added as the fourth digit. The fourth (last) digit is a “0” for any code that remained identical to CDS-I code, or in the case of an expanded code category for CDS-II, a “0” in the fourth digit indicates a catch-all / unspecified code for that particular activity. The activity codebook is on the CDS website at URL:

<http://psidonline.isr.umich.edu/CDS/Codebooks/2003codebook.pdf>.

To help illustrate the point, here is an example. In CDS-I, we coded “homework” as one type of activity. In CDS-II, we expanded the homework category to capture a distinction among non-computer related homework, studying time, library time, and homework review with the parents or caregivers. In the CDSII activity codebook the specific homework codes 5491-5494 are indented under 5490, which visually identifies that they are new codes to 2002, with 5490 as a general “catch all” code that was used in 1997. An excerpt from the codebook is on the next page:

Table 9. 2 Example: Homework Codes in CDS-II Time Diaries

Code	Code Description
5490.	General category for homework/studying-specific activity NA. Note: in 1997 code “549” read: “Homework, non-computer related, Studying, research, reading, related to classes or profession, except for current job (code 0790) and “Went to the library”. These activities are coded separately in 2002 (see below).
5491.	Homework, non-computer related. Note. Homework using computer, code as 5040.
5492.	Studying, research, reading, related to classes or profession, except for current job (code 0790), college applications, working on school project.
5493.	“Went to the library”
5494.	Reviewing homework with parent/caregiver. Note. This is new item.

In addition to these variables, the CDS Media Files were made public in 2008. These include data for both CDS-I and CDS-II. The Center for Research on Interactive Technology, Television and Children (CRITC) at the University of Texas at Austin coded the television programs that were reported in the 1997 and 2002 CDS Time Diaries. There are eight variables generated by the media coding project:

- (1) Format (TD97M09, TD02M09);
- (2) Intended Audience (TD97M10, TD02M10);
- (3) Character Age (TD97M11, TD02M11);
- (4) Genre (TD97M12, TD02M12);
- (5) Comedy (TD97M13, TD02M13);
- (6) Science Fiction, Fantasy or Supernatural/Paranormal (SF/F/SP) (TD97M14, TD02M14);
- (7) Curriculum (TD97M15, TD02M15); and
- (8) Violence (TD97M16, TD02M16) of each television program.

The television coding manual and the data files are now available. This document describes the data structure and coding procedures. For more information on the Media Files please see Appendix E.

Data Files

The CDS time diary data files are posted on the PSID-CDS Data Center (<http://simba.isr.umich.edu/>), freely accessible to researchers worldwide who are interested in using the data. The data are available at the activity level for both CDS-I (131,060 total activities, for an average of 22 activities per diary) and CDS-II (99,467 total activities, for an average of 20 activities per diary) and at child level, with the activities grouped into broader or “aggregated” variables. The illustration below shows you where you first need to go to access the time diary data – either the activity level (select “Time Diaries”) or aggregate level (select “CDS incl. Time Diary Aggregates”):

Time Diary Activity-Level Files

The Time Diary Activity data file is structured at the activity level, meaning one data record per activity. There are three important identifiers in this file: (1) the ID68 (ER30001)-PN (ER30002) combination to identify the child for whom the activity belongs, (2) the “Diary_02” to identify if the activity belongs to a weekday or a weekend diary (we reviewed earlier that each child had up to two diaries taken – one for a randomly selected weekday and one for a randomly selected weekend. “T1_02” specifies the actual day of the week), and (3) the start time of the activity, variable “COLB_02”. Start and end time are represented in seconds past midnight. The duration of the activities are recorded as amount of time in seconds that the activity took place.

The other variables in the data file are few in number - each one representing a column in the time diary grid.. See <http://psidonline.isr.umich.edu/CDS/questionnaires/cds-ii/english/cdsiitdday.pdf> for an example of the instrument. Since Columns G and H are “select all that apply” questions in the diary, there is a variable for each possible option:

Participating (Column G in Time Grid)	There, but Not Participating (Column H in Time Grid)
COLGA_02: NO ONE ELSE	COLHA_02: NO ONE ELSE
COLGB_02: MOTHER	COLHB_02: MOTHER
COLGC_02: FATHER	COLHC_02: FATHER
COLGD_02: SIBLING	COLHD_02: SIBLING
COLGE_02: STEP-MOTHER	COLHE_02: STEP-MOTHER
COLGF_02: STEP-FATHER	COLHF_02: STEP-FATHER
COLGG_02: STEP-SIBLING	COLHG_02: STEPSIBLING
COLGH_02: CHILD FRIEND	COLHH_02: CHILD'S FRIEND
COLGI_02: GRANDPARENT	COLHI_02: GRANDPARENT

COLGJ_02: OTHER RELATIVE	COLHJ_02: OTHER RELATIVE
COLGK_02: OTHER NON-RELATIVE	COLHK_02: OTHER NON-RELATIVE

You can print out just a codebook of the file through the Data Center. To do this, simply select the variables through the Data Center, but at the very last step, select “codebook only”. Since the activity codes for columns A and J are so extensive, they are provided in a separate document: <http://psidonline.isr.umich.edu/CDS/Codebooks/2003codebook.pdf>.

To give a better understanding of how the time diary activity file is structured, take a look at the example data file below. The first 32 rows are activities that span two diaries T1_02=7 (Sunday) and T1_02=4 (Wednesday) for child “0004”-“039”. COLA_02 specifies the activity and COLB_02 and COLC_02 specify the start and end time of the activity, respectively. There are 17 activity records for the Sunday diary and 15 activity records for the Wednesday diary. The first data record, for example, shows that child “0004”-“039” was sleeping (activity code 4590) from midnight (start time in seconds past midnight=0) to 9:00 AM (end time in seconds past midnight =32400 seconds). The child was not watching television so the variable COLD_02, activity type, is coded as INAP (0). The child was at home during this activity (COLF_02=10). Variable COLGA_02 is coded as INAP (0) since we did not collect information about with whom the child was engaged in the activity for activities under the personal care category. There are more variables in the data file, but this illustration should get you on the way to understanding the file structure of the Activity file.

Example SAS File for the Time Diary Activity Level Data

	ER30001	ER30002	T1_02	COLA_02	COLB_02	COLC_02	COLD_0	COLF_02	COLGA_02
1	4	39	7	4590	0	32400	0	10	0
2	4	39	7	4070	32400	33000	0	10	0
3	4	39	7	4090	33000	33300	0	10	0
4	4	39	7	4390	33300	35100	0	10	1
5	4	39	7	9190	35100	37800	2	10	1
6	4	39	7	9190	37800	40500	2	10	1
7	4	39	7	9190	40500	43200	2	10	1
8	4	39	7	8620	43200	54000	0	10	0
9	4	39	7	4080	54000	54600	0	10	0
10	4	39	7	4390	54600	55500	0	10	0
11	4	39	7	8011	55500	61200	0	10	1
12	4	39	7	8620	61200	72000	0	10	0
13	4	39	7	9190	72000	75600	2	10	0
14	4	39	7	9190	75600	77400	2	10	0
15	4	39	7	9190	77400	79200	2	10	0
16	4	39	7	4090	79200	79800	0	10	0
17	4	39	7	4590	79800	86400	0	10	0
18	4	39	4	4590	0	31800	0	10	0
19	4	39	4	4090	31800	32400	0	10	0
20	4	39	4	4390	32400	34200	0	10	1
21	4	39	4	9190	34200	36900	2	10	1
22	4	39	4	9190	36900	39600	2	10	1
23	4	39	4	4080	39600	42600	0	10	0
24	4	39	4	4090	42600	43200	0	10	0
25	4	39	4	4390	43200	45000	0	10	0
26	4	39	4	1290	45000	61200	0	10	0
27	4	39	4	4390	61200	63000	0	10	0
28	4	39	4	9190	63000	68400	1	10	0
29	4	39	4	9190	68400	72000	2	10	0
30	4	39	4	9190	72000	79200	2	10	0
31	4	39	4	4090	79200	79800	0	10	0
32	4	39	4	4590	79800	86400	0	10	0

Time Diary Aggregate Files

In addition to the activity-level files, there are child-level time diary files which “roll-up” individual activities in several ways: at the four- and three-digit code levels, which take all of the time spent in each of the activities and sum them across the diary day, and at the two-digit level, which aggregate time per diary day by 39 activity categories. These aggregate files are posted at the child-level, facilitating their use with developmental data from the other CDS modules, including (but not limited to) achievement measures, physical and emotional health assessments, and family environment measures, or with PSID family economic and demographic data.

In the aggregate file, there is one record per child and a variable that will tell you if that child has weekday diary data and another variable that will tell you if the child has weekend diary data. The labeling convention is, as follows:

Variables that begin with “WD” indicate that it is a “weekday” activity; variables that begin with

“WE” indicate that it is a “weekend” activity. For CDS-II, the next two components of the variable name is “02”, meaning, data collection start year 2002.

As mentioned above, there are four-digit aggregate variables, three-digit aggregate variables, and two digit aggregate variables. You can ascertain the level of the aggregation by the naming convention. For the variables that roll up the activities at the four-digit level, the next four components of the variable name are the four digit code for the activity. For the variables that roll up the activities at the three-digit level, the next four components of the variable name are an underscore (“_”) and the three digit equivalent (or, CDS-I equivalent) for the activity. For the variables that roll up the activities at the two-digit level, the next four components of the variable name are “39”, representing 39 categories and the two-digit code for the activity.

For example, let’s take a look at our previous example about homework time.

The four-digit aggregate variables for homework on weekdays are: WD024590, WD024591, WD024592, WD024593, and WD024594.

The three-digit aggregate is the sum of all of the specialized homework codes and for weekday, is represented as “WD02_549”: WD02_549 is equal to SUM (of WD025490, WD025491, WD025492, WD025493, WD025494).

The two-digit aggregate that includes homework time is “EDUCATIONAL AND PROFESSIONAL TRAINING”, code WD023913.

As in the Time Diary Activity-Level file, the duration of the activities are given seconds.

Chapter 10 – CDSII School Data

The CDS provides information about the schools in which the target children and youth attend in several ways. As described earlier, in the primary caregiver interview, the primary caregivers reported on the target child's current school enrollment, type of school (public or private; home-schooled), tuition and other school-related costs; grade in school, age began kindergarten, program enrollment for gifted youth or for special education, and if child ever repeated a grade or ever dropped out of school. They additionally reported out on their involvement in their children's schools and own family processes that encourage academic interests and achievements.

In the Elementary School Teacher interview, the target child's teachers provided some information about the school environment for the elementary-school aged children, including teacher-child relationships, social processes in classrooms, and target children's classroom behaviors.

In 1997, CDS-I included surveys of school administrators to gather more information about the schools that the CDS children attend. In these surveys, we asked school administrators to provide data about school enrollment, staffing, special programs, general climate, revenues and expenditures. Unfortunately, only 32% of elementary school administrators responded to the survey. Instead of gathering reports from school administrators in CDS-II, we drew upon school characteristic data collected by the U.S. Department of Education National Center for Education Statistics (NCES) Common Core of Data (CCD). This approach potentially provides more comprehensive data on the school environment and the school district, helping to minimize missing data points resulting from non-response on the administrator survey, and improve overall quality of these data.

What is the "CCD"? The CCD is a national database of all public elementary and secondary schools and school districts in the United States. Data are available on the U.S. Department of Education's website. The CCD survey annually collects data about all public elementary and secondary schools, all local education agencies, and all state education agencies throughout the United States. CCD contains three categories of information: general descriptive information on schools and school districts; data on students and staff; and fiscal data. We have an "ID Map" under sensitive data contract whereby users can extract additional CCD information for the CDS sample. This map provides the CDS identifiers along with the NCES CCD school and school district identifiers for the school that the CDS child or youth attended during the CDS-I and CDS-II interview.

Chapter 11 Coded Variables in the PCG and Child Interviews

Coding was performed by the Institute for Social Research's Survey Research Operations (SRO) group. Approximately 10 percent of the cases were coded twice, once by the coder and a second time (check coding) by the SRO Production Manager assigned to the project. Check coding is an item-by-item check where a coder's work is evaluated and verified.

Occupation & Industry for Youth's Jobs

The U. S. Census' 2000 Occupation & Industry Code Lists (U.S. Census Bureau, Demographic Surveys Division), consisting of approximately 980 occupation and 980 industry codes, were used to classify responses to several job related open-ended questions in the Child interview:

- Q23J6. Think about the job in which you work the most hours. What is your occupation? What kind of work do you do on this job? What are your most important activities or duties?
- Q23J7. What kind of business or industry is that in? What do they make or do at this workplace?
- Q23J20. Think about the job in which you work the most hours. What was your occupation? What kind of work did you do? What are your most important activities or duties?
- Q23J21. What kind of business or industry is that in? What do they make or do at this workplace?
- Q23J31. What are the three kinds of jobs you would like to have when you are done with school?
- Q23J32. Which job do you want most?
- Q23J33. We can't always get the jobs we want. Which job do you think you will actually get to do?

Code Frames for Open-Ended Questions

When developing the framework for coding open-ended responses, the goal was to avoid subjective judgments as much as possible and link coding to objective decision rules. To this end, open-ended responses were first sorted into obvious groupings based on just the phrases or words that the respondent used. All mentions, not just the most common ones, were included, and counts assigned. This constituted the universe of items to be coded.

"Lowest common denominator" categories were then grouped (if possible) into larger categories. Internet searches and background research often provided guidance in this step. If a lowest common denominator category was substantially large, it was given its own category. These "lowest common denominator" categories allowed obvious groupings to be spotted, while at the same time implicating other potential groupings that could be developed after some investigation. Every effort was made to maintain similar code frames for similar questions, such

as codes 30-97 for questions about family activities, hobbies, lessons, sports and community programs in which the child may be involved.

Child Interview

At the beginning of Section J, for the CDS youth aged 12 years and older, we asked about racial/ethnic identify and religious affiliation. Questions were formatted as open-ended items whereby the youth reported on these identifications in their own words.

In the work items in the Child Interview, we ask if the youth plan to continue their current jobs into the future – and the reasons why.

PCG Child Interview

In Section A, we asked the PCG about the target child’s reasons for visits to the hospital and to the doctor – either for illness or for injury. The open-ended text was categorized by what the respondent said – we did not attempt to roll the data into any type of classification scheme.

In Section B, we asked the PCG several open-ended questions about their child’s current education and expectations for future education. We categorized the data based on the responses themselves.

Q21B3. What things might keep this child from getting as much education as (he/she) wants? (This item is also in the Other Caregiver Interview)

Q21B16b. Please specify the reason why (CHILD) is currently in a special education class or program.

Q21B20a. Why did you wait to enroll (CHILD) in kindergarten?

In Sections E, G, and H, we ask several open-ended questions about family activities, hobbies, and extra curricular activities that were then categorized into groupings based on the data:

Q21E15a. In what family activities is (CHILD) included?

Q21G2a. & Q21H2a. What are (CHILD)’s hobbies?

Q21G5b. & Q21H5b. (Was (CHILD) in a tutoring program in the last 12 months?) What was the tutoring for?

Q21G6d. & Q21H9d. Which religious activities was (CHILD) involved in?

Q21G7a. & Q21H6a. Did (CHILD) take lessons, such as music, dance, or drama in the last 12 months? Which lessons was (CHILD) involved in?

Q21G8a, Q21H7a, Q21G8g, & Q21H7e. Which school athletic or sports teams is/was (CHILD) involved in during this/last school year?

Q21G8d & Q21H7c Which organized sports programs was (CHILD) involved in during last summer?

Q21G10a. & Q21H8a. Which community groups or programs was (CHILD) involved in?

Chapter 12 – PSID Data Resources

As mentioned in the first chapter, the Panel Study of Income Dynamics (PSID) is a longitudinal survey of a nationally representative sample of U.S. families. The PSID data archive, spanning almost four decades of data collection, contains over 48,000 variables on a diverse set of topics, including the dynamics of family composition change, marital and birth histories, housing, income, wealth, and welfare participation; health status of family members, expenditures, philanthropy, and more. Over 60,000 people have ever participated in the panel, which includes up to three generations within a family. The PSID is the longest running panel on family and individual (including child) dynamics and has consistently achieved response rates of 96-98 percent. More information about the study and its instruments can be obtained from its website: <http://psidonline.isr.umich.edu/Guide/Overview.html>. The PSID bibliography provides references to studies using the data for many research topics (key word index provided): <http://psidonline.isr.umich.edu/Publications/Bibliography/>.

As a CDS data user, the rich family data resource in the main PSID provides an incredible opportunity to explore family environmental factors in your analyses of child outcomes. In this chapter, our goal is to provide you with some basic (although not comprehensive) information about the PSID sample, interview content, public and sensitive data contract file, and data structure. Let's start first with an overview of the sample origins and design.

The PSID Sample

The original 1968 PSID sample was drawn from two independent samples: a nationally representative sample of roughly 3,000 households designed by the Survey Research Center at the University of Michigan (the "SRC sample") and an over-sample of roughly 2,000 low income families from the Survey of Economic Opportunity (the "SEO sample"). From 1968 to 1996, the PSID interviewed and re-interviewed individuals from families in these two samples every year, whether or not they were living in the same dwelling or with the same people. Since 1997, interviewing has taken place every other year. Also in 1997, the PSID added an immigrant refresher sample that was intended to bring the sample up to representativeness of immigration into the U.S. since 1968.

PSID follows adults through the full life course. Children born to an original-sample member are classified as sample members and are eligible for tracking as separate family units when they set up their own households. Ex-spouses and other adult sample members who move out of the PSID family units are tracked to their new family units. This procedure replicates the population's family-building activity and produces a dynamic sample of families each year. New PSID families form when children grow up and establish separate households or when marriage partners go separate ways. The unique self-replacing design implies that for many PSID families, the data include self-reported information on three (and occasionally now, four) generations within the same family at various points in their life course. Through multiple waves collected over long time periods on a national population, the PSID is the only data set ever collected on life course and multigenerational economic conditions, well-being, and health in a long-term panel representative of the full U.S. population.

PSID Content

As touched on above, the PSID contains a large number of variables that have been measured wave after wave in much the same, if not the identical, manner. The table below gives a general idea of the data collected but is by no means comprehensive. Years are provided when data were collected for only some, but not all 34 waves of the PSID. Keep in mind that this table is by no means exhaustive of the PSID content. Please see <http://psidonline.isr.umich.edu> documentation for more details on the PSID measures.

Table 12.1 Highlights of the PSID Interview Content

Topic	Detail
Family Composition	Status and change over time Extended family identifiers Marriage, divorce, and widowhood Fertility history
Employment Head/Wife	Current Employment Status Time Away From Work Not Working <ul style="list-style-type: none"> • Unemployed and not looking for work • Unemployed and looking for work Employed <ul style="list-style-type: none"> • Start and end dates • Employer name • Occupation and industry • Number of employees • Salary or wage • Hours per week • Overtime hours and pay • Number of weeks worked overall • Looking for work or another job in past four weeks • Military service Unemployed <ul style="list-style-type: none"> • Ever worked for money • Looking for work in past four weeks • Military service
Income Detail Head/Wife	Earned Income <ul style="list-style-type: none"> • Business Income • Wage and Salary • Bonus, Overtime, Tips and Commissions Asset Income <ul style="list-style-type: none"> • Rent • Dividends; Interest • Trust Funds & Royalties Transfer Income <ul style="list-style-type: none"> • Unemployment Compensation • Worker's Compensation
Income Detail	<ul style="list-style-type: none"> • Help from Relatives and Others not Living in FU

Topic	Detail
	<ul style="list-style-type: none"> • Supplemental Security Income • TANF • Child Support
Program Participation	Food Stamps WIC Welfare Programs Home Heating Assistance Public Housing Help or Support from: church, community group, family/friends
Wealth & Active Savings 1984, 1989, 1994, 1999, 2000 to 2005	Real Estate Personal vehicles: cars, trucks, a motor home, trailer, boat Stockes, mutual funds, investment trusts Checking or savings accounts Other assets Gifts and inheritances Debt Carryover credit card and store balances Student loans Medical or legal bills Loans from relatives
Philanthropic Giving & Volunteering Head/Wife 2001, 2003, 2005	Donations <ul style="list-style-type: none"> • Amt • Type of organization (arts, religious, helping, educational, youth/family services, improving neighborhoods, environmental, international aid, research) • How decision to support made in marital pair If volunteer- <ul style="list-style-type: none"> • Volunteer activity (religious, children/youth, senior, helping for people in poor health or in need, social change, Tsunami) • Hours spent in volunteer activities in previous year
Education 1985 – 2005	Month, year of HS graduation (H, W ²⁶) Whether GED (H, W) <ul style="list-style-type: none"> • If GED, number of grades finished (H, W) • If GED, month/year of last grade attended (H, W) If no HS or GED (H, W) <ul style="list-style-type: none"> • Number of grades finished • Month/year of last grade attended Whether college degree (H, W) <ul style="list-style-type: none"> • Month/year of last college attended • Highest college year finished • Month/year of college degree • Name of college • City, state of college Highest grade or year of school completed (OFUM) Whether other degree/vocational certificate/training (H, W) <ul style="list-style-type: none"> • School or apprenticeship: up to 2 mentions

²⁶ H: “Household Head”; W: “Wife of Head”

Topic	Detail
	<ul style="list-style-type: none"> • Type of degree/certificate • Field of training • Name of institution/organization • Month/year of degree If non-US born (H, W) <ul style="list-style-type: none"> • Years of school outside US • Highest degree/certification outside US
Education Supplement 1995	For all family members (in 1995, aged 5-49) <p>Current school attendance</p> <ul style="list-style-type: none"> • Full-time or part-time status • Current grade or year of school • High School graduate or GED • Month/year of last school attendance • Full-time or part-time status • Highest grade or year of school completed <p>Private school attendance</p> <ul style="list-style-type: none"> • Grades of private school attendance • Whether private school attendance was religious <p>Gifted or advanced class attendance</p> <p>Held back or repeated a grade</p> <ul style="list-style-type: none"> • Grades held back <p>Whether needed special education</p> <p>Whether learning disabled, perceptually impaired, speech impaired</p> <p>Head Start enrolled</p> <ul style="list-style-type: none"> • Age of first Head Start attendance • Total number of months in Head Start <p>Nursery school, preschool program, day care center enrollment</p> <p>Ever expelled or suspended from school</p> <p>Ever spent time in jail, prison, youth training or reform school</p> <ul style="list-style-type: none"> • Number of times sent to a youth corrections institution • Number of times sent to an adult corrections institution • Month/year of most recent release <p>Number of times principal requested family conference for behavior problems</p> <p>Extracurricular activity participation in past 5 years</p>
Physical and Mental Health	<p>General Health Status</p> <ul style="list-style-type: none"> • Current status 1984-2005 • Status since last interview 2003-2005 • Status birth – 16 years old 1999-2005 • Height & Weight 1986, 1999-2005 • Birth weight <p>ADL-type Measures</p> <ul style="list-style-type: none"> • 7 activities, consistently asked 1992-96, 1999-2005 <p>IADL-type Measures</p> <ul style="list-style-type: none"> • 6 activities, consistently asked 1992-96, 1999-2005 <p>Specified Health Conditions</p>

Topic	Detail
	<ul style="list-style-type: none"> • 12 conditions, w/onset, duration 1999-2005 <p>Death</p> <p>Activity Limitations</p> <ul style="list-style-type: none"> • Health limits work 1968-2005 • Employment status – disabled 1976-2005 • Days missed work for sickness 1976-2005 <p>Depression/Anxiety</p> <ul style="list-style-type: none"> • 6 item, non-specific distress scale (K6) 2001-03 • 2 item, clinical depression screener (CIDI) 2003 <p>Health Behaviors</p> <ul style="list-style-type: none"> • Smoking: current, ever, # packs, when start, when stop 1986, 1999-2005 • Alcohol: ever, quantity, binge 1999-2003, 2005 • Physical activity: light/heavy 1999-2005
Housing Detail	<ul style="list-style-type: none"> • Characteristics of housing: type of dwelling, number of rooms, air conditioning, type of heating, monthly payments for utilities • Rentals – monthly payment • Home ownership – value of home, whether have mortgage, amount of remaining principle on mortgage, monthly mortgage payment amount, Interest rate on loan, year obtained loan/refinanced, remaining years to pay on the loan, second mortgage, amount of property taxes, homeowner’s insurance premium • Neighborhood characteristics: amount of property taxes, homeowner’s insurance premium
Expenditures	<ul style="list-style-type: none"> • Food: Food at home, delivered, and eaten out • Health Care: nursing, hospital, doctor, Rx, insurance • Housing: mortgage, property tax, insurance, rent, utilities • Transportation: loan, lease, insurance, repairs, gas, parking, fares • Education: tuition, supplies, room and board <p>(New for 2005)</p> <ul style="list-style-type: none"> • Home repairs & maintenance • Household furnishings & equipment • Clothing and apparel • Trips & vacations • Recreation & entertainment • Child Care

Family and Individual Level Variables in the PSID

PSID variables can be categorized as family-level and individual-level. Most of the information from any year’s data collection is categorized as **family-level variables**. The family-level variables include not only information that applies to the family unit as a whole, but also almost all information about the head and wife/ “wife”. **Individual-level variables** are wave specific data as well as time invariant information that may span many years. These variables cover basic

demographic and economic data about an individual in the FU. If the individual was head or wife / “wife”, much of the information in these variables is available among the family level variables, often in substantially greater detail.

The Building Blocks to Using PSID Data on Families & Individuals

Because of the complexity of the PSID data structure, and its long-panel nature, there are a few PSID concepts that are important to understand to effectively take advantage of the richness of the data archive. This section of the chapter attempts to give you some background on these concepts.

Let’s first review the PSID definition of “family”. The **Family Unit (FU)** is defined as a group of people living together as a family. “As a family” means to us that the individuals are generally related by blood, marriage, or adoption, but unrelated persons can be part of a FU if they are permanently living together and share both income and expenses. The **Household Unit (HU)**, on the other hand, is defined as the physical boundary, such as a house or apartment, where members of the PSID FU reside. Not everyone living in a HU is automatically part of the FU. The PSID survey is about FU Members only.

Families change from year to year. Each family unit in a specific data collection wave is assigned a unique **Family ID**, sometimes called “**Interview Number**”. The most critical Family ID is the one assigned in 1968 to families in the original sample. According to our following rules, we follow “splitoffs” as children and others as they set up their households. One family in 1968 can become 3-4 or more families by 2003. All of these families will have the same 1968 ID, since they originated from the same family in 1968, but will have different family ID’s in 2003, since they are separate family units in 2003.

For each family, the Family ID number will most certainly vary from year to year. For example, a 1968 family ID of 1234 will not likely be 1234 in 1969 or any other year for that matter. Yearly IDs are assigned based on receipt of the interview—the first interview in from field is numbered 1, the second, 2, and so on. In the table on the next page, the “Smith Family” in 1968 had a Family ID of “0100” and included five FU members: Jason and Andrea, and their three children Julie, Anne, and Todd:

The “Smiths” in 1968

Family Unit Member	Age	1968 Family ID
Dad (Jason)	36	0100
Mom (Andrea)	35	0100
Julie	12	0100
Anne	10	0100
Todd	8	0100

By 2003, Jason and Andrea’s children had moved out on their own, thus creating four families out of this original 1968 “Smith Family”. Notice each family keep their original 1968 ID, but also was assigned unique 2003 Family ID:

The “Smiths” in 2003

Families	68 ID	2003 Family ID “Interview Number”
Jason (Dad) & Andrea (Mom)	0100	0834
Julie, her husband and son	0100	1004
Anne	0100	0971
Todd, his wife and daughter	0100	2202

The PSID collects information about the **individuals** of each FU. In the Data Center, each individual is uniquely identified by their combined **1968 ID** (labeled as ER30001 in the Data Center) and **Person Number** (labeled as ER30002). At the time of data release for each wave of the PSID, updates to Person Numbers are made. This is important to remember since it will require updating of data files downloaded at a previous point-in-time.

The most information is collected about the Head of the household, the second most about the wife (or “wife” in the scenario of a cohabiter for one year or more), and the least amount of detail is collected about other family unit members (OFUMs), who are oftentimes the children in the FU, but can also be boy/girl friends, grandparents, etc. All of the individuals in an FU have a **Relationship to the Head** coded in the data files

As mentioned above, PSID gathers the most information about the **Head** of the FU. Within each wave of data, each FU has only one current Head. The person designated as Head may change over time as a result of other changes affecting the family. When a new Head must be chosen, the following rules apply:

The Head of the FU must be at least 16 years old and the person with the most financial responsibility for the FU. If this person is female and she has a husband in the FU, then he is designated as Head. If she has a boyfriend with whom she has been living for at least one year, then he is Head. However, if the husband or boyfriend is incapacitated and unable to fulfill the functions of Head, then the FU will have a female Head.

A new Head is selected in situations such as:

- Last year’s Head moved out of the HU (household unit), died or became incapacitated; or
- A female Head has gotten married; or
- An individual or group of individuals split-off to form a new FU (we’ll talk more about this later).

The **Wife** is the spouse of the head; “**Wife**” is the cohabiter of one year or more. Much of the information collected in the PSID about Heads is also collected about the Wives and “Wives”.

Continuing with our Smith Family example, Julie was one of the original sample members in 1968. She was the daughter of Jason (Head) back then. In 2003, she was living in her own FU, Family ID 1004, with her husband and their child. Her relationship to Head is now Wife. Her husband is Head, and their son is Child of Head.

Julie's Family in 2003

2003 Family ID	68 ID	Individuals	Relationship to Head
1004	0100	Julie (original 1968 sample member)	20 (Wife to Head)
1004	0100	Husband	10 (Head)
1004	0100	Son	30 (Child of Head)

Other Family Unit Members (OFUMs) are the other individuals living in the family unit. Not as much information is collected about OFUMs as Heads and Wives. Generally, the PSID main interview provides individual level data on the OFUMs as well as data each wave about their:

- Demographic information, such as birth month and year, sex, date of death, family unit status, wave-specific response status
- Marital and fertility histories, and cohabitation status
- Work status, number of jobs, occupation, weeks worked on job, during which months of previous year (2004), hours per week
- Other income sources – (interest, TANF, SSI, welfare, social security, unemployment compensation, worker's compensation, child support, help from relatives)
- (For OFUMs 16 years and older) Date last attended school, currently enrolled as part time or full time student, highest grade or year of school completed
- Health insurance coverage

Note that CDS "children" are OFUMs in the PSID data collections and are considered PSID individuals. Data about them in the main PSID interviews can be merged onto CDS data files.

In CDS, the **primary and secondary caregivers are in relationship to the child** (variables "RELPCG02" and RELOCG02" in 2002). If you are interested in merging data from PSID about the caregivers, it is helpful to know the caregivers' relationships to head. In some situations, the data you are interested in may have only been collected for Heads and Wives. This accounts for the majority of the PCG and OCGs.

Keeping in mind the PCG's Relationship to Head, if you are interested in constructing, for example, the PCG average work hours, you will find that PSID asks for this information and stores it in three separate variables depending on the individual's relationship to head:

HDAVG01 B79. And, on the average, how many _hours a week_ did you work on your main job(s) in 2000? (HEAD)

WFAVG01 D79. And, on the average, how many _hours a week_ did she work on her main job(s) in 2000? (WIFE) (exact same wording)

G81. During the weeks that (he/she) worked, about how many hours did (he/she) usually work per week? (OFUM)

At the beginning of this chapter, we mentioned that the PSID has a self-replacing sample design whereby children are interviewed as their own family unit after they leave their parents' households and ex-spouses and other adult sample members who move out of the family unit are

tracked to their new family units. With this design, at any wave, the PSID has main or “re-interview” families and split-off families. A **main family** is a family unit (FU) that was interviewed in the prior wave. In a divorce or separation, the main family could be the family unit of either spouse, depending on who was found first for the interview. In the case of children leaving home, the main family is almost always the parental family. A **split-off family**, on the other hand, consists of a person or group of people (at least one of whom is a “follow” person of any age) who moved out from a main family since the prior wave’s interview to form a new, economically independent family unit. Several criteria must be met for a split-off to occur. In addition to having moved out since the prior wave, and to being ‘followable’ (i.e., having an original 1968 family identification), the person or group of people have not moved to an institution such as college or prison or to another family unit within the panel study. Moreover, the person or group of people who have moved out and formed their own family unit must be economically independent (i.e., they must be paying their own living expenses) from the family unit from which they split off. Information that informs the criteria is gathered at the main family interview, during the household listing stage. The main family Respondent provides these details. It should be noted that these are general rules, and that sometimes unique situations arise that determine whether a person or group of persons becomes a split-off. For example, while moving to an institution such as college does not generally meet the criteria for becoming a split-off, if the person is working, paying their own living expenses, and paying their own educational expenses in addition to attending school, then this person would be considered a split-off. Information that can inform this type of scenario is gathered in Section G of the main family interview and in the interviewer observations. The living situation and interview data for each and every possible split-off case are first reviewed before split-off status is granted.

Family Composition Change

When looking at variables pertinent to the same person(s) over time, the variable called “Sequence Number” can be of help in understanding continuity and change in family composition wave-to-wave. The **Sequence Number**, found in the individual file, identifies an individual’s status with regard to the family unit at a given wave of data collection. The sequence number is coded as follows:

- 1-20 indicates that the individuals in the family at the time of the current wave interview
- 51-59 indicates that the individuals were in institutions at the time of the current wave interview
- 71-80 indicates that the individuals moved out of the FU or out of institutions and established their own households between the prior and current wave interview
- 81-89 indicates that the individuals died by the current interview but were living at the time of the prior wave interview.

PSID also has a variable (“**Whether Moved In/Out**”) to indicate whether the individual moved in or out of a family unit in a given wave of data. The 2003 codebook definition below helps illustrate the detail of this variable.

- 1 Moved in or born between the 2001 and 2003 interviews; split-off in 2003
- 2 Appeared; did not move in between the 2001 and 2003 interviews, but not included in FU for 2001 (i.e., listing error)
- 5 Moved out of FU or out of an institution between the 2001 and 2003 interviews and was not included in a 2003 panel FU
- 6 Moved out of FU and into an institution between the 2001 and 2003 interviews
- 7 Living in 2001 but died by the time of the 2003 interview
- Disappeared; moved out prior to the 2001 interview, but included in FU for 2001 (usually a listing error)
- 8 listing error

PSID Data Center Files

Most of the family and individual-level data on all waves of the PSID and CDS are housed in the online Data Center. There are tutorials that provide a guided overview to utilizing the data. You can create customized data sets and codebooks through the easy-to-use Data Center. More information on using the Data Center is provided in Chapter 14.

Data Groupings

In the Data Center, you will find the data grouped by PSID main study data and CDS data. Within the PSID data group, the data are clustered by individual-level files and family level files. The individual level files include both wave-specific and time-invariant data, mentioned earlier in this chapter. The family-level files include not only “raw” interview data but also complex generated variables on income, work, wealth, sample weights, and other measures. Data by wave of the study are provided as the next level within each of the data groupings. In the CDS data group, the data are clustered by child-level files, which include all of the interview modules, and by time diary activity files, which are at the activity level.

The screenshot shows the 'Data Center >> Variable Selection >> By File' interface. It includes a header, a descriptive paragraph, instructions on how to use the interface, two buttons ('Add To Cart' and 'Reset'), and a tree view of data groupings. The tree view is expanded to show 'PSID Family-level' and 'PSID Individual-level' categories, with sub-items like 'PSID Main Family Data', 'Active Savings', 'Risk Tolerance', 'Food Security', 'PSID Individual Data by Years', and 'Summary Variables (Sampling Variables, Family History Variables, and Sex)'. Other categories include 'CDS (including Time Diary Aggregates)' and 'CDS Time Diaries'.

PSID Supplemental Files

In addition to the data accessible through the on-line Data Center, PSID provides zipped files with data that focus in detail on various aspects of the main study. These data are accessible through the “PSID Packaged Supplemental Files” at <http://psidonline.isr.umich.edu/Data/zipSuppData.html>. The Family History Files may be of particular interest to CDS Data Users. These files provide information on biological and adoptive parents, childbirth and adoption history, and (3) marriage history of the PSID individuals (of interest for CDS – marriage history of the parents / caregivers). These files are updated at the time of each PSID main study data release.

Parent Identification File

The Parent Identification File (PID) summarizes information collected from various sources since the 1983 wave of the Panel PSID about parent-child relationships. This file contains relevant identifiers for the child, his or her birth and adoptive parents, and information source indicators. Linkages can be done from either the child’s standpoint – to find the identifiers of the CDS child’s parents, or from the child’s parents’ standpoint, to find the CDS child’s grandparents’ identifiers. All CDS children have information about their parents on the Parent Identification File (PID); however, note that a parent is necessarily the same person identified in the CDS as primary caregiver (PCG).

Childbirth and Adoption History File

The Childbirth and Adoption History File contains all present-year and past-year detail about the timing and circumstances of childbirths and adoptions up to and including the most current wave of the PSID. In situations where an individual has never had or adopted children, there is one record recording the fact that the individual has no such experience. Variables on this file include the identifiers for each parent and child, month and year of birth for both parent and child, birth order, birth weight and date of death for a child, year of most recent report and number of births/adoptions, etc. Data on this file are structured in a one-record-per-event format, with each record representing a specific childbirth or adoption event.

The 1999, 2001, and 2003 Family Files and the 1968-2003 Individual Files in the Data Center also contain variables for number of childbirth, birth month and year of most children, identifiers of mother and of father, whether the mother was married at the time the individual was born, and birth order of individual.

Marriage History File

The 1985-2003 Marriage History File contains detailed information about marriages of people of marriage-eligible ages living in a PSID family at the time of the interview in any wave between 1985 and 2003. Each record contains all past-year and most-recent-year detail about the timing and circumstances of a marriage for a specified individual up to and including 2003 (or whichever wave 1985-2003 when the individual was last in an interviewed family unit). Data on this file are structured in a one-record-per-marriage format, with each record containing relevant data about the spouse and timing of a specific marriage. In situations where an individual has never been married, there is one record recording the fact that the individual has no marriages. Variables on this file include the identifiers for each individual and his/her spouse, month and year of marriages/divorce/widowhood, order of the specific marriage, total number of marriages and the most recent year wave when data were collected.

The PSID Individual Files in the Data Center also include variables for number of marriages, beginning and end month and year for the first and the most recent marriages, marital status of the individual at the time of the most recent interview, and the most recent wave when marriage data were collected.

Geographic Residence and Environmental Characteristics

The PSID Geocode Match Files contain the identifiers necessary to link the main PSID data to Census data. This linkage allows the addition of data on neighborhood characteristics for the geographic areas in which panel individuals and families reside to the already rich socioeconomic variables collected in the PSID. In recent years, there has been a rapid growth of external sources that provide an increasing variety of measures of the social environment. Rather than investing our resources in duplicating this effort, we are expecting users to seek out these sources to match with the PSID files. Because the Geocode Match data are highly sensitive (usually pinpointing the census tract in which families lived), this information is available only under special contractual conditions designed to protect the anonymity of respondents.

Chapter 13– Data File Structure within CDS Files

CDS-II data are stored in 13 groupings, each file specific to a CDS-II data collection module described in Chapter 2 of this User Guide, with the addition of a background data file labeled “Demographic File”. Except for Time Diary Activity file and Media File, the data are stored as one record per child. Time Diary and Media File data are stored as one record per time use activity (each child can have more than one activity record). Once you are in the Data Center, and have selected “CDS” grouping, you will then see all of the possible data files listed, with a selection box for each wave in which the data are available. (<http://simba.isr.umich.edu/ALL/selectcats.asp>). The table in Appendix D provides an overview of the 13 CDS data groupings.

The PCG Household and OCG Household data were collected at the family level, meaning if two children in the same household were interviewed for CDS-II, the household interview was administered only once to the PCG and once to the OCG (if applicable). To keep all of the CDS files at the child-level, as done in the CDS-I, we applied the household data to each response target child household, resulting in one data record per child in the data files you will find on the PSID-CDS Data Center. The sibling indicator in the CDS-II Demographic File gives you information on whether a specific CDS response child has a response sibling in the CDS-II. Note that Family ID accompanies data downloads.

The PCG Child Interview was required for the CDS target child to be considered “response” in the CDS-II data collection. The other interview modules were completed based on eligibility and cooperation with appropriate respondent in the home or school. The data files are structured to reflect this study design.

In the 2002 Demographic file, you will find a set of indicator variables for assessment, child, time diary, OCG and teacher modules that specify (a) if a case was eligible for that module and (b) if a record exists for that case in the corresponding data file. These variables are helpful when you merge variables from multiple CDS modules. The sample weight in the Demographic file can be used to adjust for unit nonresponse at the child level. The module indicator variables, however, will inform you about item missing data across modules. It is up to you to then decide on your preferred approach for addressing item missing data that results from differential response rates across modules (for example, you may leave it as missing, impute data, etc).

To review, each response target child has a data record in the PCG child file, and every data record from each of the other data files has a corresponding record in the PCG Child interview file. You will not find a data record in any of these modules that is not in the PCG Child data file.

On the other hand, not every targeted child with a completed PCG Child interview has a completed interview for each (or any) of the other modules. There are two reasons for this: (a) the child was not eligible for a specific module, or (b) the child was eligible but did not complete the specific module. For instance, children under eight years of age were not eligible for a Child interview and thus do not have a corresponding record in the Child interview file. Children who were eight years or older and had a PCG interview completed for them, but refused to participate themselves in an interview likewise do not have a record in the Child interview file but do have one in the PCG Child data file.

Variable names were kept as similar as possible and variable labels the same or as similar as possible in CDS-II as to CDS-I for convenience in referencing variables. Variable labels in CDS-

II have “02” at the end. For example, one of the behavior problem index items Q21B30D is labeled “CURIOUS 02” in CDSII, and its companion variable in CDS-I, Q1G24D, is labeled “CURIOUS”.

Variable names in CDS-II start with “Q2x” (x=1–9). For example, variable Q21A3-“HOSPITALIZATION 02” is from CDS 2002 PCG Child interview (indicated by “Q21” in the variable name and “02” in the variable label), Question A3.

Table 13.1 Labeling References for CDS-I and CDS-II Variables

Data Center File Name	Q# Ref CDS-II	Q# Ref CDS-I
Primary Caregiver Child File	Q21	Q1
Primary Caregiver Household File	Q22	Q2
Child File	Q23	Q3
Assessment Data	Q24	<i>Not in CDS-I</i>
Other Caregiver Child File	Q25	Q5
Other Caregiver Household File	Q26	Q6
Time Diary Questionnaire	Q28	<i>No Q ref.</i>
Time Diary Aggregate Variable File	WD02	WD97
	WE02	WE97
Elementary School Teacher File	Q29	Q9
Demographic File - including weights	<i>No Q ref.</i>	<i>No Q ref.</i>
Time Diary Activity File	<i>No Q ref.</i>	<i>No Q ref.</i>
Media File	<i>No Q ref.</i>	<i>No Q ref.</i>

On the CDS website, we have posted an excel table that serves as a cross-wave index, mapping maps the individual variables across the two waves of the CDS. This file is located at: http://psidonline.isr.umich.edu/CDS/questionnaires/pcgchild_xyr.pdf. You can also download codebooks from the Data Center or review on-line when selecting variables. Codebooks give exact question wording and complete code frame.

Chapter 14 – Using the Data Center and Merging Files

With exception of sensitive data contract files, all PSID and CDS data are located on the PSID-CDS Data Center. In this chapter, we provide you with a basic step-by-step guide to using the Data Center. As mentioned in an earlier chapter, there are also tutorials on the website that are helpful to understanding basic structural concepts in the data

(see <http://psidonline.isr.umich.edu/Data/DataCenterHilites.ppt>, http://psidonline.isr.umich.edu/Guide/tutorials/tutorials_home.html). While all of the CDS data are located within the Data Center, some PSID data are located in zipped files (see <http://psidonline.isr.umich.edu/Data>). The codebook and read-me file that accompany each supplemental data file are detailed enough to illustrate the data structure and content of the files. Towards the end of this chapter, we give you some basic information on merging data across the PSID and CDS.

In addition to actual data about the CDS child and his/her family, there are a number of tools on the PSID-CDS website that facilitate use of the data, such as the Caregiver Map and the Family Identification Mapping System. This chapter gives an overview of how these tools function. Let's start with a few basic concepts of the Data Center.

Using the Data Center: The Basics

The underlying premise of the Data Center is to facilitate the selection of variables from both the CDS and the PSID panel, across multiple data collection years, and to output the data in a format that fits your specific analytic needs. More specifically, there are six main functions within the PSID-CDS Data Center that facilitate building a data file. These functions are: 1) automatic downloading and merging of multiple waves of PSID data to create PSID-CDS datasets; 2) customized codebooks in many formats; 3) subsetting option for certain types of individuals or families; 4) intergenerational dataset creation for CDS children and PSID family members; 5) data cart management; and 6) search and browse functions. The following section will describe these functions as we show you how to create customized data sets and codebooks using nearly 50,000 variables across the (current to this writing) 34 waves of PSID data (19682003), two waves of CDS data (1997 and 2002).

Data Groups

The PSID-CDS Data Center stores data for PSID and CDS into groupings of variables. In the PSID, data groups are mainly categorized as individual level data (time-invariant grouping and PSID individual grouping) and family level data, which includes data from the questionnaire and data from constructed measures (e.g., Income Plus, Work Hours and Wages, etc.). In the CDS, data groups mainly follow data the collection modules. The structure of the CDS data files is described in the previous chapter. Mainly, all files except for Time Diaries are child-level files, with one data record per participating CDS child. When you select data from the CDS grouping, your final output file should have up to 3,563 data records if you selected from CDS-I grouping, or 2,907 records if you selected from CDS-II data.

If you select variables from both PSID and CDS data groupings, your data file is structured as one record per CDS child, with PSID individual or family variables attached to each record. Individual level data from PSID in this scenario is about the child, and family level data is about the family unit in which the child resided in that data collection year, and the head or the wife of the child's family unit. Later in this chapter we discuss merging on PSID individual data about the child's parents and caregivers.

Welcome to the Data Center!

When arriving at the Data Center’s homepage, you will see there are four ways to add variables to your cart: 1) By File; 2) By Topical Index; 3) By Search; or 4) By Cart.

Welcome to the Data Center!

The Data Center is where you create subsets and documentation. Use the Variable Selection menu above to begin creating your customized dataset.

There are four ways to add variables to your cart:

- ◆ **By File:** The Data Center provides a drill-down interface into the PSID and CDS data files so that you can add variables of interest into your Data Cart.
- ◆ **By Topical Index:** The Data Center also provides an easy to navigate tree interface to the many topical categories of variables throughout the years. This allows you to easily view, compare, and select variables of interest for each year available.
- ◆ **Search:** Search on variable **name**, **label**, **question**, and **explanation** text that are currently in the Data Center. Other searches are of the PSID [Web Site](#), and the PSID/CDS [Bibliography](#).
- ◆ **By Cart:** Retrieve data carts created by you or others.

Use any method you want, and switch methods while you shop and the content of your cart will not be lost. To view the variables you have selected, simply click on data cart.

Selecting Variables by File

When adding by file, you will use the drill-down interface to eventually select specific variables from PSID/CDS and add them to your data cart. You must first, however, select the Data Group(s) of interest from PSID, CDS, or both. In our example, we select the “PSID Family-level” data group and the “CDS” data group:



- PSID Family-level
- PSID Individual-level
- CDS (including Time Diary Aggregates)
- CDS Time Diaries

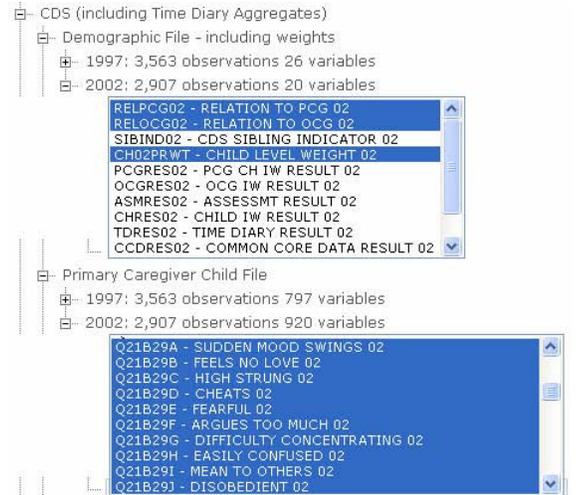
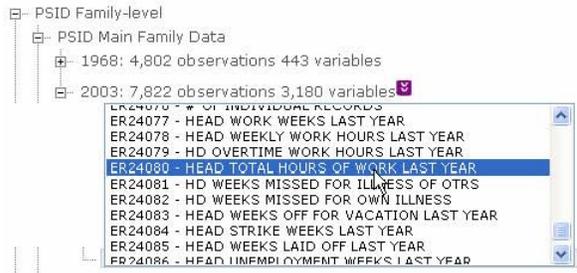
- PSID Family-level
 - PSID Main Family Data
 - Active Savings
 - Risk Tolerance
 - Food Security
- PSID Individual-level
- CDS (including Time Diary Aggregates)
 - Demographic File - including weights
 - Primary Caregiver Child File
 - Primary Caregiver Household File
 - Child File
 - Other Caregiver Child File
 - Other Caregiver Household File
 - Fathers Outside the Home Child File
 - Fathers Outside the Home Household File
 - Elementary/Middle School Teacher File
 - Preschool/Daycare Teacher File
 - Home-Based Care Provider File
 - Elementary/Middle School Administrator File
 - Preschool/Daycare Administrator File
 - Food Security Status File
 - Time Diary Questionnaire
 - Child Assessments
 - Time Diary Activity - Aggregate File
 - Selected Variables from NCES CCD
 - Selected Variables from NCES Private School Survey (PSS)
- CDS Time Diaries

Wave of Data

Expanding these selections even further allows you to view the specific waves, or years, of data that are available for each selected data group from both the PSID and CDS. You can build your data file by selecting any number of waves of data from the PSID and CDS. PSID currently has 34 waves of data in the Data Center and CDS has two waves.

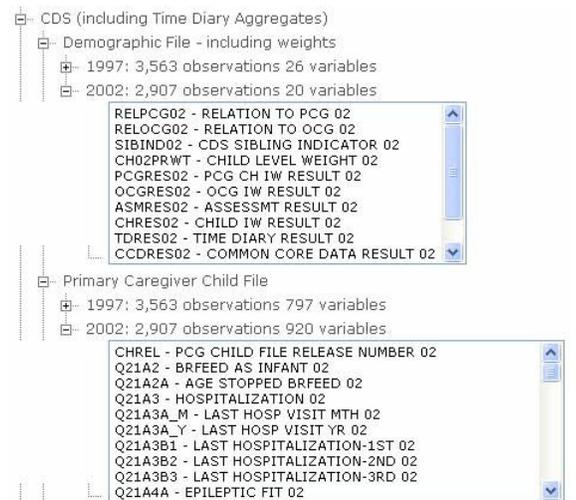
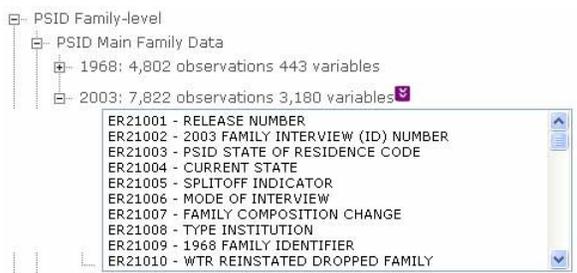
- [-] PSID Family-level
 - [-] PSID Main Family Data
 - [-] 1968: 4,802 observations 443 variables
 - [-] 1969: 4,460 observations 580 variables
 - [-] 1970: 4,645 observations 598 variables
 - [-] 1971: 4,840 observations 551 variables
 - [-] 1972: 5,060 observations 585 variables
 - [-] 1973: 5,285 observations 312 variables
 - [-] 1974: 5,517 observations 333 variables
 - [-] 1975: 5,725 observations 434 variables
 - [-] 1976: 5,862 observations 821 variables
 - [-] 1977: 6,007 observations 483 variables
 - [-] 1978: 6,154 observations 522 variables
 - [-] 1979: 6,373 observations 516 variables
 - [-] 1980: 6,533 observations 558 variables
 - [-] 1981: 6,620 observations 612 variables
 - [-] 1982: 6,742 observations 540 variables
 - [-] 1983: 6,852 observations 634 variables
 - [-] 1984: 6,918 observations 1,080 variables
 - [-] 1985: 7,032 observations 1,347 variables
 - [-] 1986: 7,018 observations 1,188 variables
 - [-] 1987: 7,061 observations 1,038 variables
 - [-] 1988: 7,114 observations 1,409 variables
 - [-] 1989: 7,114 observations 1,313 variables
 - [-] 1990: 9,371 observations 1,246 variables
 - [-] 1991: 9,363 observations 1,246 variables
 - [-] 1992: 9,829 observations 1,250 variables
 - [-] 1993: 9,977 observations 1,764 variables
 - [-] 1994: 10,765 observations 2,017 variables
 - [-] 1995: 10,401 observations 1,857 variables
 - [-] 1996: 8,511 observations 2,108 variables
 - [-] 1997: 6,748 observations 2,085 variables
 - [-] 1999: 6,997 observations 3,467 variables
 - [-] 2001: 7,406 observations 3,395 variables
 - [-] 2003: 7,822 observations 3,180 variables
- [-] CDS (including Time Diary Aggregates)
 - [-] Demographic File - including weights
 - [-] 1997: 3,563 observations 26 variables
 - [-] 2002: 2,907 observations 20 variables
 - [-] Primary Caregiver Child File
 - [-] 1997: 3,563 observations 797 variables
 - [-] 2002: 2,907 observations 920 variables

The next step is to then select data year(s) of interest from each data group. In our example on the next page, we selected 2003 data from “PSID Main Family Data” and 2002 data from CDS “Demographic File” and “Primary Caregiver Child File.”



Variable Selection

The third step to creating your PSID-CDS data file is to actually select the variables of interest and add them to your data cart. You can select as few or as many from each of the data groupings you have already chosen. In our example, we select “ER24080 – HEAD TOTAL HOURS OF WORK LAST YEAR” (the Head’s total work hours in the previous year (2002)) from the 2003 PSID Family File. From the CDS 2002 Demographic File we select “RELPCG02” & “RELOGCG02” (the relationship of the PCG and OCG to the child), and “CH02PRWT,” (the child’s sample weight in 2002). NOTE: to select more than one variable, press and hold Control. We also select all ten variables belonging to the Behavior Problem Index (BPI), “Q21B29A” through “Q21B29J,” from the CDS 2002 Primary Caregiver File. Once you have selected all desired variables by highlighting them, click the “Add To Cart” button located at the top of the page.



Review of Data Selections

The next step is to review the selections you have chosen by viewing your data cart. Clicking the blue “your cart” link will take you to a summary of the variables you have chosen. Clicking “Expand All” will provide you with a detailed list of all selected variables, and allows you to delete any that are unwanted.

The screenshot displays a data selection interface. On the left, there are buttons for "Add To Cart" and "Reset", and a link for "Variables added to your cart." Below this, a list of variable categories is shown with expand/collapse icons: PSID Family-level, PSID Individual-level, CDS (including Time Diary Aggregates), and CDS Time Diaries.

On the right, there is a detailed tree view of the selected variables. At the top, there are buttons for "Delete Checked Nodes", "Empty Cart", and "Checkout". The tree view is currently expanded, showing a hierarchy of variables for the years 1968, 2001, 2002, and 2003. The variables are organized into levels (Family-level, Individual-level) and specific data files or variables. For example, under 2002, there is a "Primary Caregiver Child File" with variables like "SUDDEN MOOD SWINGS 02" and "FEELS NO LOVE 02".

Notice that the Data Center automatically appends certain key individual and family identifier variables to the data set, such as “ER30001,” “ER33603,” and “ER33701.” These variables, which are described in depth in Chapter 12, are important to understanding relationships within the family and to the family head.

Output Options

After finalizing your variables, “Checkout” your data cart. In this step you can specify the type of format for your codebook and data file outputs, and also name your file. In our example we select a PDF codebook and a SAS data file, and name our file “Behavior Problems”.

Codebook Type

HTML
 PDF
 XML
 No thanks, I don't need a codebook
 Codebook only (no data will be output)

Data Output Type [?]

ASCII Data With SAS Statements
 Microsoft Excel Spreadsheet
 ASCII Data With SPSS Statements
 dBase Data File (DBF) / Spreadsheet
 ASCII Data With Stata Statements
 SAS V9 Transport Data File

Subsetting Criteria [?]

Compression Option [?]

Create files in compressed (.zip) format for faster downloading

Cart Options [?]

Make my cart available to the public
 Send links to my Email

Additionally, under “Cart Options,” you may elect to have the Data Center notify you via email after your data have been created and provide you with a link to your output files. This option allows saving data carts for future use.

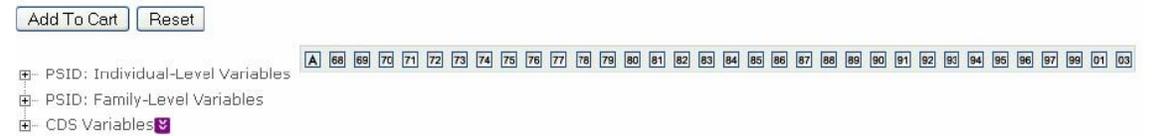
You may also decide it is necessary to limit the sample to cases where one or more variables are in specific value ranges to define a subset of observations. To subset the data, enter the appropriate command in the “Subsetting Criteria” box in the Output Options section of the data selection process.

If the email option is not elected, the Data Center will create the specified files and allow you to view, download, and save them.

File Type	File Size (KB)
Customized Codebook	64
Right-click and choose <i>Save As</i> to download.	
SAS V9 Transport File	993
Your customized data contain 8,379 observations and 22 variables, created Monday, Jul 10 at 6:11 PM.	
Output more than 7 days old will be deleted from this server	
In addition to the customized codebook, we highly recommend that you also download the complete documentation:	
<ul style="list-style-type: none"> ◆ Packaged Main Data, Documentation, and Questionnaires ◆ Packaged Supplemental Data and Documentation ◆ CDS Documentation and Questionnaires 	

Selecting Variables by Topical Index

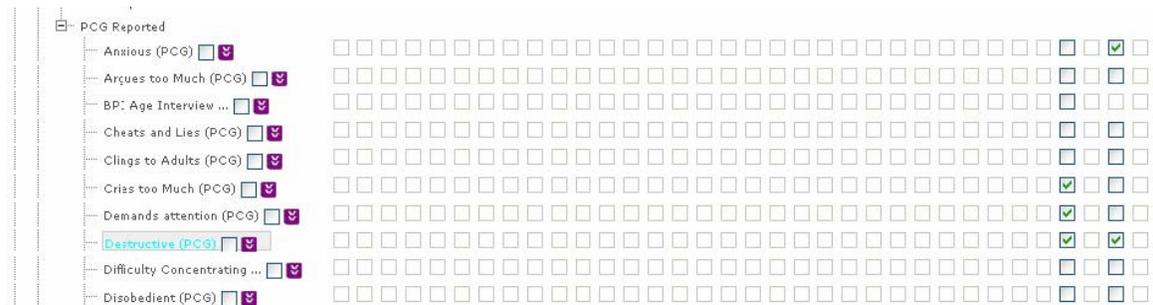
A second approach to selecting variables in the Data Center is by Topical Index. This method also employs a drill-down tree interface that allows you to browse through the many topical categories of variables and chose the desired waves of data available for those selected.



In this example, we will again attempt to select the Behavior Problem Index variables from the 2002 CDS. Expanding the “CDS Variables” data group leads us to the “Problem Behaviors” category, which contains the Behavior Problem Index.



Expanding the “PCG Reported” BPI lists the 39 possible variables in this topic, and allows you to directly select as many or as few of the variables for any combination of waves available. We can see that CDS data for these variables is available for the 1997 and 2002 waves. NOTE: If you are unsure about what wave you are selecting, float your mouse over the checkbox in question to display the year. In our example, we select the following: “Cries too Much,” “Demands attention,” & “Destructive” for 1997, and “Anxious” & “Destructive” for 2002. Once you have selected all desired variables by checking the appropriate boxes, click the “Add to Cart” button located at the top of the page.



As before, you may follow the blue “your cart” link and “Expand All” to get a detailed summary of the contents in your data cart, including the variables you selected, and the key relational identifier variables the Data Center appends automatically.

|

- 1968 (Including Summary Variables that apply to all years)**
 - PSID Individual-level
 - Summary Variables (Sampling Variables, Family History Variables, and Sex)
 - ER30001: 1968 INTERVIEW NUMBER
 - ER30002: PERSON NUMBER 68
- 1997**
 - PSID Individual-level
 - PSID Individual Data by Years
 - ER33401: 1997 INTERVIEW NUMBER
 - ER33402: SEQUENCE NUMBER 97
 - ER33403: RELATION TO HEAD 97
 - CDS (including Time Diary Aggregates)
 - Primary Caregiver Child File
 - Q1G23V: DESTROYS THINGS DELIBERATELY
 - Q1G23X: CRIES TOO MUCH
 - Q1G23Y: DEMANDS ATTENTION
- 2001**
 - PSID Individual-level
 - PSID Individual Data by Years
 - ER33601: 2001 INTERVIEW NUMBER
 - ER33602: SEQUENCE NUMBER 01
 - ER33603: RELATION TO HEAD 01
- 2002**
 - CDS (including Time Diary Aggregates)
 - Primary Caregiver Child File
 - Q21B29E: FEARFUL 02
 - Q21B29V: DESTRUCTIVE 02

Lastly, “Checkout” your cart, select your output options, and retrieve your customized data files.

Selecting Variables by Search

A third approach to selecting variables is to use the Search feature, which allows you to search codebook text for all variables in the Data Center. You may select any combination of Data File Types and Data Years, as well as specify the text field types it searches against. You may search against variable names, variable labels, question text & explanation text (notes about the question wording or variable construction), or all of the above.

In our example, we look for the phrase “Behavior Prob” in the Question/Explanation text fields for variables in the 1997 and 2002 CDS data group.

Data File Type	Data Year	Codebook Part	Search Type
<input type="checkbox"/> PSID Family-level <input type="checkbox"/> PSID Individual-level <input checked="" type="checkbox"/> CDS (including Time Diary Aggregates) <input type="checkbox"/> CDS Time Diaries	<input type="checkbox"/> 1997 <input checked="" type="checkbox"/> 2002	<input checked="" type="radio"/> Question or Explanation Text <input type="radio"/> Variable Label <input type="radio"/> Variable Name <input type="radio"/> All of the Above	<input type="radio"/> Any words (OR) <input checked="" type="radio"/> All words (AND) <input type="radio"/> Phrase
<input type="text" value="Behavior Prob"/>			<input type="button" value="Search Codebooks"/>

The engine generates a list of all records of the BPI index for 1997 and 2002, including such variables as the internalizing and externalizing scale, and total BPI score. Clicking “Check All” and “Add To Cart” will deposit the variables into the Data Cart, which we can then review and output into the desired format(s).

16 records found

Data Type	Data Category	Year	Name	Label	Select
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1A24	MENTAL PROBLEMS	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1A24M	LAST MENTAL PROBLEM	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1A24Y	LAST MENTAL PROBLEM (YEAR)	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1EXTERN	BPI EXTERNAL SCALE	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1INTERN	BPI INTERNAL SCALE	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	1997	Q1BPITOT	BPI SCALE TOTAL SCORE	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	Q21A8	DR VISIT - EMOT PROB 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	Q21A8A_M	EMOT PROBLEM MTH 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	Q21A8A_Y	EMOT PROBLEM YR 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	BPI_T02	BPI - TOTAL SCORE 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	BPI_E02	BPI - EXTERNALIZING SCORE 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Primary Caregiver Child File	2002	BPI_N02	BPI - INTERNALIZING SCORE 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Elementary/Middle School Teacher File	1997	Q9A24	BEHAVIOR/DISCIPLINE PROBLEMS	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Elementary/Middle School Teacher File	2002	Q29A15	BEHAV/DISCIPLINE PROBS 02	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Preschool/Daycare Teacher File	1997	Q10A22	BEHAVIORAL PROBLEMS SINCE 1996	<input checked="" type="checkbox"/>
CDS (including Time Diary Aggregates)	Home-Based Care Provider File	1997	Q11A11A	SUDDEN MOOD SWINGS	<input checked="" type="checkbox"/>

Linking Data Files

There are a large number of variables in the PSID that can be used along with CDS. Every individual in the PSID has both an “ID68” (1968 Family Identifier) and “PN” (Person Number) that when combined, uniquely identify that individual. Identifier variables are available through a “CDS Caregiver Map” that is created automatically whenever CDS and/or PSID individual-level data are downloaded, and through the Family Identification Mapping System. Let’s start with the Child-to-Caregiver Map file.

CDS Caregiver Map

An integral aspect of the CDS research design is to collect information about the children’s caregivers. Typically, the caregivers are the children’s biological or adoptive parents, but not

always. We define the primary caregiver as the person living with the child who has primary responsibility for caring for that child, and the secondary caregiver as the person living with the child who assumes secondary responsibility for caring for that child. Under this definition, non-biological parents may assume a central caregiver role in the child’s life.

Since the caregivers reside with the CDS child by study definition, information about them is collected in the main PSID. To make linkages possible between the rich panel data in the PSID data archive and the CDS, we provide a file with unique individual identifiers for the CDS child, primary caregiver, and secondary caregiver that is automatically appended to any Data Center download of CDS data. The map file always has the same name as the data download, with “map” in the filename. For example, if your data file name is “job39350.txt”, your map data file would be “map39350.txt”.

The file is formatted as a flat file, with one record per CDS individual. Every individual in the PSID has both a 1968 Family Identifier or “ID68” and Person Number or “PN” that when combined, uniquely identify that individual. In the Data Center, these variables are “ER30001 – 1968 INTERVIEW NUMBER” (ID68) and “ER30002 – PERSON NUMBER 68” (PN). The first two columns of the map file are ER30001 and ER30002, which are the 1968 Family Identifier and Person Number, respectively, of the child. The next two variables, PCGID97 and PCGPN97, represent the ER30001 (1968 Family ID) and ER30002 (Person Number) of the individual identified as the primary caregiver of the child in 1997. The next two variables OCGID97 and OCGPN97, represent the ER30001 (1968 Family ID) and ER30002 (Person Number) of the individual identified as the other caregiver of the child in 1997. Missing data points on the OCG variables indicates that a secondary caregiver who is in the PSID followable sample was not identified for that year.

The PCG and OCG variables are repeated in the next four columns, but here, represent the identifiers for CDS-II (2002-2003) data collection. Note that missing data points on PCG 2002 identifiers indicate that the CDS child did not participate in the second wave. Missing data points on the OCG 2002 identifiers indicate either (a) the CDS child did not participate in the second wave, or (b) a secondary caregiver who is in the PSID followable sample was not identified for that year. You can determine whether a secondary caregiver was in the home at the time of the CDS-II interview with variable RELOCG02 on the CDS-II Demographic data file. Table 1 is an example “map file” output.

Table 1: Excerpt from the CDS Caregiver Map File

	ER30001	ER30002	PCGID_97	PCGPN_97	OCGID_97	OCGPN_97	PCGID_02	PCGPN_02	OCGID_02	OCGPN_02
368	655	30	655	172	655	3	655	172	655	3
369	655	31	655	172	655	3	655	172	655	3
370	656	31	656	4	656	172	656	4	656	172
371	656	33	656	5	656	173	656	5	656	173
372	656	34	656	5	656	173	656	5	656	173
373	656	35	656	4	656	172	656	4	656	172
374	664	30	664	8	664	174	664	8	664	174
375	664	32	664	175	664	7	664	175	.	.
376	664	34	664	175	664	7	664	175	.	.
377	664	35	664	8	664	174	664	8	664	174
378	672	172	672	3	672	171
379	672	174	672	3	672	171	672	3	672	171
380	673	32	673	172	673	1	673	172	673	1
381	673	33	673	30	673	175	673	30	673	2
382	673	174	673	172	673	1	673	172	673	1
383	674	30	674	5	674	170	674	5	674	170
384	674	31	674	171	674	3	674	171	674	3

In the example above, the primary caregiver for the first child listed, “0655”-“030” is “0655”-“172” in both CDS-I and CDS-II. The secondary caregiver for child “0655”-“030” is “0655”-“003” in both waves as well. There are two ways to determine the relationship of the caregivers to the child. The first is to use the variables on the CDS Demographic data files “RELPCG97/02” and “RELOCG97/02”. This information was reported by the PCG during the CDS data collections. The second approach is to use the “Relationship to Head” variables on the PSID individual files to map a relationship of the child to the caregiver through both of their relationships to the person designated as the “Head” of the family.

To use the CDS caregiver map file for merging data about the caregivers onto CDS child-level file, you need to create data files about each of the caregivers, merge onto the map file, then merge onto the CDS child-level file. First, identify the variables about the PCG and the OCG that are of relevance to your analysis. The cross-year index in the Data Center (<http://simba.isr.umich.edu/VS/i.aspx>) is an excellent tool for selecting measures within specific substantive domains across multiple waves of the PSID. Keep in mind that some characteristics of the PCG and OCG may be in the PSID Individual File, and some may be in the PSID Family Files if these individuals are Heads (and Wives) of PSID families. Chapter 12 of the CDS-II User Guide provide information on PSID respondent roles and data structure – this is a good place to start to get acquainted with the PSID content prior to selecting data.

Once a data file about the PCG and a data file about the OCG are created, merge these files onto your CDS map file by using ID68 and PN for each of the caregivers as the primary keys for merging. The last step is to merge these files onto the CDS data file you have downloaded from the PSID-CDS Data Center. Use ER30001 and ER30002 as the primary keys for merging. The end result is a data file with all of the CDS variables that you have selected for your analyses and the identifiers for the caregivers.

Family Identification Mapping

As mentioned earlier, caregivers are usually, but not always, biological or adoptive parents. There may be times when you want to know about biological or adoptive parents only or grandparents. In the fall of 2006, PSID incorporated a new data extraction tool on the PSID-CDS website (<http://psidonline.isr.umich.edu/FIMS>) called the “Family Identification Mapping System” or “FIMS”. FIMS facilitates the process of constructing a data file with identifiers for pre-specified relationships to PSID individuals. For example, by using the FIMS, you can easily create a customized inter-generational identification map with identifiers for parents, grandparents, or great grandparents.

This section describes how the FIMS constructs Inter-Generation Identification (GID) data files. The FIMS uses the *Parent Identification File* or “PID”, a supplemental PSID data file that summarizes information about parent-child relationships in its “behind-the-scenes” extraction process. The PID contains relevant identifiers for the child and his or her birth and adoptive parents for all individuals in the PSID (not just the CDS children), including both sample and nonsample individuals. More information about the original construction of PID and its updating process is available in the “Parent Identification Introduction” documentation (see URL: <http://simba.isr.umich.edu/Zips/ZipSupp.aspx#MH>). A brief description of the file and how individual family members are followed across time and generations in the PSID is available in Chapter 12 of the CDS-II User Guide.

To illustrate the FIMS process for creating a GID data file, we will use an example for selecting *biological parents and grandparents of the CDS children*. The final output file will have one data record per individual with biological parent and grandparent identifiers, where available, on each record. Let's get started.

Step 1: Getting to FIMS

The FIMS extraction tool is part of the data sharing resources on the PSID-CDS website. You can access FIMS from <http://psidonline.isr.umich.edu/FIMS>. To begin the process of constructing a GID file with parent and grandparent identifiers of CDS sample members, select "Inter-generation (GID)" under "Select Map Type", the first selection box on the FIMS home page. Options for further specification will then appear.

Step 2: Select Parent-Type

Under "Parents" drop box, select "biological parents".

The genealogical sample design of the PSID means that the parents, grandparents, and siblings of respondents are respondents themselves in some waves of the data. This feature supports unique analyses of the inter- and intra-generational connections and relationships within families. To facilitate creation of data extracts that include information on parents, grandparents, and siblings of respondents, we have designed the tool on this web page.

To understand how the tool works, review the User Help examples.

Inter-generation (GID) <input type="button" value="v"/>	Select User Help... <input type="button" value="v"/>
Options	
Parents:	Output Options
Biological Only <input type="button" value="v"/>	<input type="radio"/> SAS V9 System File
Generation Map:	<input type="radio"/> Excel
--Select a value <input type="button" value="v"/>	<input type="radio"/> ASCII With SAS Statements
Map Type:	<input type="radio"/> ASCII With SPSS Statements
--Select a value <input type="button" value="v"/>	<input type="radio"/> ASCII With Stata Statements
File Format:	<input type="button" value="Submit"/>
--Select a value <input type="button" value="v"/>	

FIMS uses in its "behind the scenes" extraction. While the FIMS draws upon the PID as the initial source of information in creating GID data files, there are important differences between what the PID and the FIMS offer. These differences are highlighted as we describe how the FIMS tool works. The FIMS relies on PID variables that contain the 1968 Family Identifier ("interview number") and Person Number for all PSID individuals and their parents of interest (i.e., variables PID1-2; PID3-4; PID5-6; PID17-18; PID1920). As described earlier, the combination of the 1968 Family Identifier and Person Number unique identify each individual in the PSID. There are 78,902 individuals in the PID data file, Data Release #4. Recall that there are 3,563 CDS-I participants enrolled at the baseline interview in 1997.

The example in this document focuses on selecting *biological parents and grandparents of the CDS children*. The PID contains information about adoptive parents and likewise, in the FIMS, you can specify adoptive parents for your output identification map file. For our example, however, we are focusing on biological parents. Table 2 below shows the PID input data for biological parents that the FIMS would draw upon. The data are shown for PSID family "2069" and are current as of the 2003 wave of the PSID. Data points with a value of "0" on the PID

represent a lack of available information about that individual or in the case of adoptive parents, possibility that the individual does not have an adopted mother / father. From the table below, we see that Family “2069” has 20 individual members. There are identification numbers for biological mothers for 14 of the individuals and identification numbers for biological fathers for nine of the individuals.

Table 2: Data Values from PID for PSID Family 2069

1968 INTERVIEW NUMBER OF INDIVIDUAL	PERSON NUMBER OF INDIVIDUAL	1968 INTERVIEW NUMBER OF BIRTH MOTHER	PERSON NUMBER OF BIRTH MOTHER	1968 INTERVIEW NUMBER OF BIRTH FATHER	PERSON NUMBER OF BIRTH FATHER
2069	1	2069	913	2069	912
2069	2	2069	176	2069	907
2069	3	2069	2	2069	1
2069	4	0	0	0	0
2069	5	2069	4	0	0
2069	30	2069	2	2069	171
2069	31	2069	3	2069	175
2069	170	2069	915	2069	914
2069	171	2069	906	2069	905
2069	172	2069	170	0	0
2069	173	2069	170	0	0
2069	174	0	0	2069	171
2069	175	2069	911	2069	910
2069	176	0	0	0	0
2069	177	0	0	0	0
2069	178	0	0	0	0
2069	179	0	0	0	0
2069	901	2069	170	0	0
2069	903	2069	170	0	0
2069	904	2069	170	0	0

Next, FIMS selects information for all individuals in the PID who have ever been a part of a PSID family unit. These are the individuals about whom PSID has collected data. Recall that the purpose of FIMS is to facilitate the development of analytic files for exploring intergenerational patterns on a wide range of topics measured in the PSID and CDS. If an individual never resided in a PSID family and data were never collected about them, then we would not be able to include them in analysis. Hence they are not included in FIMS. They are, however, included on the PID. We can identify these individuals by their Person Number: range of 900-990. From the PID codebook, person numbers are defined as:

PN Code	Code Description
1 - 19	Individuals living in core sample families at the time of the 1968 interview; individuals living in Latino sample families at the time of the 1990 (or 1992) interview; individuals living in Immigrant sample families at the time of the 1997 (or 1999) interview
20	Husband of a core sample 1968 Head living in an institution at the time of the 1968 interview; spouse of a Latino sample 1990 Head living in an institution at the time of the 1990 interview; spouse of an Immigrant sample 1997/1999 Head living in an institution at the time of the 1997/1999 interview

- 21 - 26 Children or stepchildren of a 1968 core sample Head who were under 25 and living in institutions at the time of the 1968 interview; children of a Latino sample 1990 (or 1992) Head who were under 25 and living in institutions at the time of the 1990 (or 1992) interview; children of an Immigrant sample 1997 (or 1999) Head who were under 25 and living in institutions at the time of the 1997 (or 1999) interview.
- 30 - 169 Individuals born into the core sample since the 1968 interview, the Latino sample since the 1990 (or 1992) interview, or the Immigrant sample since the 1997 (or 1999) interview and who have at least one sample parent; these persons are sample members
- 170 - 226 Individuals who moved into a core family for the first time after the 1968 interview or who have no sample parent; individuals from the Latino sample, other than the spouse of a 1990 (or 1992) Head, who moved out in the year preceding the 1990 (or 1992) interview or who moved in for the first time after the 1990 (or 1992) interview; individuals from the Immigrant sample, other than the spouse of a 1997 (or 1999) Head, who moved out in the year preceding the 1997 (or 1999) interview or who moved in for the first time after the 1997 (or 1999) interview. These individuals are nonsample unless they were born after the initial wave and have at least one sample parent.
- 227 Spouse of a core sample 1968 Head who moved out in the year prior to the 1968 interview; spouse of a Latino sample 1990 Head who moved out in the year preceding the 1990 (or 1992) interview; spouse of an Immigrant sample 1997 (or 1999) Head who moved out in the year preceding the 1997 (or 1999) interview. These individuals are nonsample.
- 228 Spouse of a core sample 1968 Head who died in the year prior to the 1968 interview; spouse of a Latino sample 1990 (or 1992) Head in the year prior to the 1990 (or 1992) interview; spouse of an Immigrant Latino sample 1997 (or 1999) Head in the year prior to the 1997 (or 1999) interview. These individuals are nonsample.
- 229 - 399 Individuals who moved into a core family for the first time after the 1968 interview or who have no sample parent; individuals from the Latino sample, other than the spouse of a 1990 (or 1992) Head, who moved out in the year preceding the 1990 (or 1992) interview or who moved in for the first time after the 1990 (or 1992) interview; individuals from the Immigrant sample, other than the spouse of a 1997 (or 1999) Head, who moved out in the year preceding the 1997 (or 1999) interview or who moved in for the first time after the 1997 (or 1999) interview. These individuals are nonsample unless they were born after the initial wave and have at least one sample parent. This series is simply a continuation of the 170-226 group.
- 900-990 Child who has never been in any sample family.

The Person Number range of 1 to 399 (actual data range 1-231 for PID Release # 4) represents individuals that have resided in family units and are thus included in FIMS. There are 78,902 individuals listed in the PID, but 65,332 have a person number under 900 and are thus included in FIMS.

When we select CDS-I baseline participant data records (n=3,563) from the PID, we find that all CDS individuals have a person number in the range of 1-231 and are thus included in FIMS. In terms of biological parents, there are four CDS children whose biological mothers have a PN greater than 900 and have thus never resided in a PSID family and 64 CDS children for whom we do not have data on biological mothers (PN=0). There are seven CDS children whose biological fathers have a PN greater than 900 and 833 CDS children for whom we do not have data on

biological fathers (PN=0). The net result is that as of the 2003 PSID Data Release, there are biological mother identifiers for 3,495 CDS-I baseline participants and biological father identifiers for 2,723 CDS-I baseline participants in the FIMS.

To illustrate, table 2 shows the individuals in the PID for Family 2069 who would be excluded from FIMS since they never resided in a PSID family unit. These excluded individuals with a PN greater or equal to 900 are represented below as “.”. There are three individuals, four biological mothers, and five biological fathers. We would not expect to find these individuals in the FIMS output, nor would we expect to find the biological mothers and fathers who are coded as “0” in PID.

Table 3: Input Data Values Adjusted in FIMS

	1968 INTERVIEW NUMBER OF INDIVIDUAL	PERSON NUMBER OF INDIVIDUAL	1968 INTERVIEW NUMBER OF BIRTH MOTHER	PERSON NUMBER OF BIRTH MOTHER	1968 INTERVIEW NUMBER OF BIRTH FATHER	PERSON NUMBER OF BIRTH FATHER
20796	2069	1
20797	2069	2	2069	176	.	.
20798	2069	3	2069	2	2069	1
20799	2069	4	0	0	0	0
20800	2069	5	2069	4	0	0
20801	2069	30	2069	2	2069	171
20802	2069	31	2069	3	2069	175
20803	2069	170
20804	2069	171
20805	2069	172	2069	170	0	0
20806	2069	173	2069	170	0	0
20807	2069	174	0	0	2069	171
20808	2069	175
20809	2069	176	0	0	0	0
20810	2069	177	0	0	0	0
20811	2069	178	0	0	0	0
20812	2069	179	0	0	0	0
20813	.	.	2069	170	0	0
20814	.	.	2069	170	0	0
20815	.	.	2069	170	0	0

Step 3: Select Generation Map

The next step is to select a map that will provide identifiers for CDS children, their biological parents, and their biological grandparents. Under “Generation Map” drop box, select “Individual to Grandparents”. In this scenario, “grandparent” is the farthest reaching generation. The output file will provide identifiers for the individuals, their parents, and grandparents.

The genealogical sample design of the PSID means that the parents, grandparents, and siblings of respondents are respondents themselves in some waves of the data. This feature supports unique analyses of the inter- and intra-generational connections and relationships within families. To facilitate creation of data extracts that include information on parents, grandparents, and siblings of respondents, we have designed the tool on this web page.

To understand how the tool works, review the User Help examples.

Inter-generation (GID)

Options

Parents:
 Biological Only

Generation Map:
 Individual to Grandparents

Map Type:
 --Select a value

File Format:
 --Select a value

Output Options

SAS V9 System File

Excel

ASCII With SAS Statements

ASCII With SPSS Statements

ASCII With Stata Statements

The process for identifying grandparents is similar to the process described above for identifying parents. Essentially, FIMS iterates the extraction, this time focusing on *parents of parents*, or grandparents. Because we have selected “Biological Only” under “Parents” drop box, our example is focused on biological parents and grandparents.

To execute this request, FIMS looks at each individual in the PID and finds biological parents and grandparents. Table 4 on the next page shows the generational linkage that FIMS would output for PSID Family “2069”. Each individual from family “2069” is listed under “ID68.PN” – a combination variable of ER30001 and ER30002. Take a look back at Table 3 to see the correspondence between the input data and output data.

The next variable in Table 4 is the Generation Number. Note that some individuals in Family “2069” are third generation individuals. The next column, “Ancestor to Individual Link Path” shows the lineage. Looking back at individual “2069” “030”, we see the mother and grandmother of individual 2069.030 are listed on the first line, and the parents (mother first, father second) listed on the next two lines. For Family “2069”, there are eight individuals with identifiers for at least one biological parent, and three individuals with identifiers for at least one biological grandparent.

Table 4: Individual, GID and Generation Linkage and Parent Type in FIMS for Family “2069” Out of the CDS-I baseline participants, there are 3,495 biological mothers, 2,723 biological fathers, and 3,043 biological grandparents with identifiers in the FIMS and data in the PSID archive.

#	ID68.PN	GID	Ancestor to Individual Link Path
1	2069.001	1	2069.001 - (I)
2	2069.002	2	2069.176, 2069.002 - (M, I)
3	2069.003	3	2069.176, 2069.002, 2069.003 - (M, M, I)
		2	2069.001, 2069.003 - (F, I)
			2069.002, 2069.003 - (M, I)
4	2069.004	1	2069.004 - (I)
5	2069.005	2	2069.004, 2069.005 - (M, I)

			2069.177, 2069.005 - (AF, I)
6	2069.030	3	2069.176, 2069.002, 2069.030 - (M, M, I)
		2	2069.002, 2069.030 - (M, I)
			2069.171, 2069.030 - (F, I)
7	2069.031	4	2069.176, 2069.002, 2069.003, 2069.031 - (M, M, M, I)
		3	2069.001, 2069.003, 2069.031 - (F, M, I)
			2069.002, 2069.003, 2069.031 - (M, M, I)
		2	2069.003, 2069.031 - (M, I)
			2069.175, 2069.031 - (F, I)
8	2069.170	1	2069.170 - (I)
9	2069.171	1	2069.171 - (I)
10	2069.172	2	2069.170, 2069.172 - (M, I)
11	2069.173	2	2069.170, 2069.173 - (M, I)
12	2069.174	2	2069.171, 2069.174 - (F, I)
13	2069.175	1	2069.175 - (I)
14	2069.176	1	2069.176 - (I)
15	2069.177	1	2069.177 - (I)
16	2069.178	1	2069.178 - (I)
17	2069.179	1	2069.179 - (I)

Step 4: Select Generation Map Type and Format

The last step is to select a map type and file format. Under “Map Type” drop box, you can select “balanced” or “unbalanced” map. A generation map with non-missing data from any generation is called balanced generation map. A generation map with missing data from the higher generation selected is called unbalanced generation map. For the output example on the next page, we selected “unbalanced”.

Under “File Format”, select a wide or long file, depending on whether you want to work with a file that has one observation per individual (wide), or one observation per relationship (long). If you are linking with your CDS data file, select WIDE and then check the box that says “Merge with CDS Map”. This will give you relationships only for CDS-I baseline participants. Under “Output Options”, you can select your analysis program format – SAS, SPSS, Stata, or Excel.

The genealogical sample design of the PSID means that the parents, grandparents, and siblings of respondents are respondents themselves in some waves of the data. This feature supports unique analyses of the inter- and intra-generational connections and relationships within families. To facilitate creation of data extracts that include information on parents, grandparents, and siblings of respondents, we have designed the tool on this web page.

To understand how the tool works, review the User Help examples.

Inter-generation (GID) <input type="button" value="v"/>	Select User Help... <input type="button" value="v"/>
Options	
Parents: Biological Only <input type="button" value="v"/>	Output Options <input type="radio"/> SAS V9 System File
Generation Map: Individual to Grandparents <input type="button" value="v"/>	<input type="radio"/> Excel
Map Type: Unbalanced Map <input type="button" value="v"/>	<input checked="" type="radio"/> ASCII With SAS Statements
File Format: Wide - 1 obs. per individual <input type="button" value="v"/>	<input type="radio"/> ASCII With SPSS Statements
	<input type="radio"/> ASCII With Stata Statements
	<input type="button" value="Submit"/>
<input checked="" type="checkbox"/> Merge with CDS Map -- Wide Format	

Among the variables that will be outputted in your GID file from FIMS are the identifiers for biological parents and grandparents. Note that maternal and paternal grandparents are identified separately.

```
ER30001 LABEL="CHILD: 1968 INTERVIEW NUMBER"
ER30002 LABEL="CHILD: PERSON NUMBER 68"
ER30001_P_F LABEL="1968 INTERVIEW NUMBER /PARENT /FATHER"
ER30002_P_F LABEL="PERSON NUMBER 68 /PARENT /FATHER"
ER30001_P_M LABEL="1968 INTERVIEW NUMBER /PARENT /MOTHER"
ER30002_P_M LABEL="PERSON NUMBER 68 /PARENT /MOTHER"
ER30001_GP_FF LABEL="1968 INTERVIEW NUMBER /GRANDPARENT /FATHER OF FATHER"
ER30002_GP_FF LABEL="PERSON NUMBER 68 /GRANDPARENT /FATHER OF FATHER"
ER30001_GP_FM LABEL="1968 INTERVIEW NUMBER /GRANDPARENT /FATHER OF MOTHER"
ER30002_GP_FM LABEL="PERSON NUMBER 68 /GRANDPARENT /FATHER OF MOTHER"
ER30001_GP_MF LABEL="1968 INTERVIEW NUMBER /GRANDPARENT /MOTHER OF FATHER"
ER30002_GP_MF LABEL="PERSON NUMBER 68 /GRANDPARENT /MOTHER OF FATHER"
ER30001_GP_MM LABEL="1968 INTERVIEW NUMBER /GRANDPARENT /MOTHER OF MOTHER"
ER30002_GP_MM LABEL="PERSON NUMBER 68 /GRANDPARENT /MOTHER OF MOTHER"
```

For GID data files that are merged with the CDS Map, you will also get the primary and secondary caregiver ID68 and PN variables. You are then ready to link your map file with your CDS child-level file – the linkage should be done with ER30001 and ER30002 combined, the unique identifiers per CDS individual and will follow the same steps as described above for caregiver map file.

Chapter 15 – A Few Tips on Using the PSID-CDS Data Resources

Admittedly, the PSID-CDS data are complex data resources that may appear difficult to use at first glance. The key to using the CDS and PSID data, however, lies in understanding the basic data structure before jumping in and downloading variables across multiple years of collection for multiple individuals in the child's family.

Remember that CDS children are individuals in the PSID panel of families where we collected more in-depth information about them, individually, where the main PSID data collection does not. This means that the CDS kids have data in the PSID main archive (not just the parents or caregivers) as well as in the CDS-specific data collections (including CDS-I, CDS-II, and TA 2005).

In the PSID, the data are collected at a family level, with some individual-level details. But the bulk of the data revolve around the persons who are Heads and Wives of the households. In the CDS, the data are collected mainly at the individual-level, revolving around the children. Some information, however, is collected about the caregivers / parents, who may be Heads or Wives in the PSID. A first tip to using the data is to review the chapters on PSID and CDS data structure in this guide to get a firm grasp on data structure for both data collections. It is well-worth the time investment to become familiar with basic concepts prior to attempting data downloads so you know what to expect and how to merge.

The CDS itself has multiple modules, collected from multiple individuals in the family. The [User Guides](#) provide survey and sample design information, historical context of the study, documentation about specific scales and indices that the CDS staff generated for public use, and data structure of the PSID. Questionnaires are also posted for both CDS and PSID. For detailed listing of measures by module, and for a comparison of measures with other large-scale public use datasets on children and adolescents, see:

<http://psidonline.isr.umich.edu/CDS/researchdesign.html>

Questionnaires are provided on web pages:

<http://psidonline.isr.umich.edu/CDS/quescodetable.html> (for CDS)

and <http://psidonline.isr.umich.edu/Data/zipCore.aspx> for PSID questionnaires.

We additionally encourage you to review the “Frequently Asked Questions (FAQ) for both PSID and CDS: <http://psidonline.isr.umich.edu/Guide/FAQ.aspx>. This webpage gives a quick overview of some key concepts in both projects.

Next, take the opportunity to review the tutorials posted on the PSID website. These [tutorials](#) provide step-by-step instructions on downloading and analyzing the data in a variety of ways, including: cross-sectional analyses, tracking longitudinal changes, creating and analyzing a balanced panel, creating and analyzing a customized subset of CDS modules, and merging PSID and CDS data together for intergenerational analysis. These tutorials also help to understand

PSID file structure, identification numbers, and how to obtain caregiver data about the CDS target children/youth. They can be found at:

http://psidonline.isr.umich.edu/Guide/Tutorials/tutorials_home.html.

The Data Center is the most popular means for obtaining PSID data, and it delivers about 10,000 customized data files to researchers and quantitative social science students each year. The Data Center is fully automated and allows for user-specified subsetting criteria when downloading and merging main family and main individual data files. ASCII or SAS data files can be generated, along with SAS, SPSS, and Stata data definition statements.

In addition to the Data Center, data and documentation can also be obtained from the PSID Web site in the form of prepackaged files. Data files and SAS and SPSS data definition statements are available via the documentation page and Anonymous FTP. Data from the Geocode files, Death files, and Medicare files are available by special contract. See the PSID Overview page or contact PSID Help for more information.

Each time you go to the Data Center to download data, visit the Data News page <http://psidonline.isr.umich.edu/Guide/DataNews.aspx> to see if any modifications to existing files or new files have been posted. Sign up for email alerts when new data or data sharing tools are released.

Appendices

Appendix A: Measurement Resource Table for CDS Modules

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
PCG-Child Interview				
Birth Weight		A8	National Longitudinal Survey of Youth (NLSY)-Mother/Child Supplement	PCG reports child's weight in pounds and ounces at time of birth – asked at CDS baseline only.
Breastfeeding	A2-A2a	A10-A10a	NLSY	PCG reports whether child was breastfed as an infant and age of child in months when stopped. Items remained identical in wording and sequencing as in CDS baseline.
Health Care Utilization	A3-A3b A5-A8b A10	A12, A20 A22-A25 A27	NLSY	PCG reports: frequency of overnight hospital stays, date of most recent stay, reason for this hospital stay, month and year of last doctor visits for illness, injury, and mental health and reason for each visit; and status of shots and immunizations. Items remained identical in wording and sequencing as in CDS baseline.
Chronic conditions: epilepsy, asthma, diabetes, ear infections, speech impairment, hearing impairment, sight impairment, mental retardation, mental health condition, anemia, lead, orthopedic impairment, developmental disability, autism, ADHD, allergies.	A4a-p	A21a-s	NLSY	PCG reports whether doctor ever said child had each of 16 chronic illnesses (plus "other"). This is an expanded list from the NLSY. In CDS-II, allergies and medications for asthma were added; except "developmental delay" and "learning disability" were combined into one item; all other items remained identical in wording and sequencing as in CDS baseline.
Activity Limitations of Child, parent-report	A9a,b,c	A26	NLSY	PCG reports whether the child has any physical/mental conditions that limit him/her in play, school attendance, or school work. Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
General Health Status of Child, parent-report	A11	A34	NHIS	PCG rates the child's health on a scale of 1 (Excellent) -5 (Poor). Items remained identical in wording and sequencing as in CDS baseline.
Asthma symptoms: wheezing with exercise or playing hard; frequency of wheezing; hospital visits for asthma or wheezing; doctor visits for asthma or wheezing; missed school because of asthma or wheezing in the past 12 months	A12-A16		Jerome Nriagu at the University of Michigan	In these new items to CDS-II, PCG reports the child's experience with asthma-related symptoms in the past 12 months.
Health Care Expenses for Insurance and Out-of-Pocket Care	A17-A24	A28-A33	Hofferth / Eccles	PCG reports (in the past 12 months): How much paid for dental and medical insurance and for out-of-pocket dental and medical expenditures, and who paid these expenses. Series modified across CDS-I/CDS-II.
School Expectations	B1-B3	G21	NLSY; National Education Longitudinal Survey of 1988 (NELS:88)	Question items adapted from NLSY and NELS to tap PCG expectations for child's educational attainment. Items remained identical in wording and sequencing as in CDS baseline.
School Enrollment and Costs	B4-B6 B11-B22 B28	G2-G3 G10-G20	National Health Examination Survey (NHES95); PSID95; NLSY	Adapted measures on enrollment in school and special/gifted programs, school characteristics, fees, and tuition; age began kindergarten, ever suspended, ever drop out. Items remained identical in wording and sequencing as in CDS baseline.
Missed School Days Due to Illness or Injury	B7	G4-G5	NLSY	PCG reports the number of days in the past 12 months that the target child missed school because of illness and injury. Items remained identical in wording and sequencing as in CDS baseline.
Enrollment in Federal Breakfast/Lunch Program	B8-B9	G6-G8	Food and Nutrition Survey (USDA)	Items taken from USDA to assess participation in federal breakfast and lunch programs. Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
School Changes	B10	G30	NLSY	Number of times the CDS target child changed school in current school year. Items remained identical in wording and sequencing as in CDS baseline.
Participation in Child's Education	B23a-d B24a-h	G27a-d G28a-k	NELS:88 ECLS NHES	Several educational surveys contributed items to measure various dimensions of parental involvement in the child's schooling. CDS-II items B23a-d are the same as CDS-I G27a-d. CDS-II items B24 are the same as CDS-I G28 except CDS-I items G28f, g, and i are dropped.
Parent – Child Communication about School	B26a-c	G31a-c	NELS:88	These items measure the frequency and nature of communication between parent and child about school. These questions were modeled after questions in the NELS: 88 parent baseline survey. Items remained identical in wording and sequencing as in CDS baseline.
Behavior Problems Index	B29a-ff	G23a-dd G32a-b	NLSY version of the scale developed by Peterson & Zill (1986). See Peterson, J. L., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavioral problems in children. <i>Journal of Marriage and the Family</i> , 48, 295-307.	The Behavior Problems Index (BPI) measures the incidence and severity of child behavior problems. CDS uses the same set of items used in the NLSY. NLSY added items B29aadd to the original Achenbach Behavior Problems Checklist to provide additional measures for the withdrawn behavior subscale. Items B29ee-ff map to items G32a-b in CDS-I. Items remained identical in wording and sequencing as in CDS baseline. There is an overall BPI scale and two subscales that measure externalizing or aggressive behavior and internalizing or withdrawn behavior.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Positive Behavior Scale	B30a-j	G24a-j	Child Trends, JOBS study version of the measure developed for the New Chance Evaluation Study. See Polit, D. (1998). The Positive Behavior Scale. Saratoga Springs, NY: Humanalysis.	The Positive Behavior Scale (PBS) measures positive aspects of children's lives, including self-esteem, social competence, self-control, obedience/compliance, and persistence. The scale used in CDS consists of 10 items that were selected by Child Trends, Inc., for use in the JOBS evaluation, measured on a five-point scale. These 10 items are a subset of the original Polit (1998) scale consisted of 25 items evaluated on a 10-point scale, from not at all like my child to very much like my child. Items remained identical in wording and sequencing as in CDS baseline.
Household Tasks	B31a-n	G25a-m	National Survey of Families and Households (NSFH); NLSY	The series of items about household tasks measures cognitive stimulation in the family environment. Items B31a-d, B31f, B31i-B31j tap activities that could be done in a household with children in ways that are not dependent on purchased toys or equipment. Items B31d B31e; B31gB31h, B31k, B31m tap activities with items that are purchased (and potentially income-biased), including computer/video games and board, card games, and puzzles. CDS-II added one item to the CDS-I set – “do yard work or gardening (together)”. The item “look at books with (CHILD)” was modified in CDS-II to additionally include “...or talk with (CHILD) about books (he/she) has read”. Other items remained identical in wording and sequencing as in CDS baseline.
Parental Monitoring	B33-B34	G33-G34	NLSY	These items obtain information about PCG knowledge of child's friends and activities. Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Household Rules	B35a-B35r1	A31a-I in PCG-HH	Detroit Area Study 1997	Series of questions that query the PCG about the frequency in which (s)he sets rules about a number of items such as watching TV, staying up late, and eating sweets. The items were expanded in CDS-II to capture behaviors relevant to adolescents.
Sibling Interaction	B37a-e			Six questions that collect information about social interactions with the child's siblings, including helping and other pro-social behaviors.
Child Care	C1-C39	H1-H40	National Child Care Survey 1990; NLSY Mott, F. L., & Baker, P. (1989). Evaluation of the 1989 Child Care Supplement in the National Longitudinal Survey of Youth. Columbus, Ohio: Center for Human Resource Research, The Ohio State University.	CDS obtains information on the type, frequency of use, and costs of arrangements for children up to and after Kindergarten. Information on arrangements used during weekday, weekend, and summer are obtained separately. Types of arrangements queried were expanded in CDS-II to capture full range of options used through adolescence.
Conflict with Absent Parent	D14a-h D30a-h	J28a-k	NSFH	These items were drawn from the National Survey of Families and Households, supplemented by additional items from other scales. It measures the degree of conflict between the parent with whom the child resides and the absent parent. CDS-II included the same items as CDS-I except two items were dropped: "how the absent parent spends money" and "disciplining (CHILD)".
Absent Parent Interaction	D15a-d D31a-d	J29	NSFH JOBS Child Outcomes Study	These items obtain information about the frequency and types of activities in which absent parents are involved with their children. Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Aggravation in Parenting	E8a-c in PCG-Child; J15a-d in PCG-HH	B11a-e in PCG-Child; A29a-d in PCG-HH	Child Trends, Inc. for the JOBS Child Outcomes Study	The aggravation in parenting scale (Q1E8a-c) and (J15a-d) measures parenting stress that may result from changes in employment, income, and other factors in their lives. The three items in the CDS-II PCG Child interview focus on the specific CDS child and are from the JOBS study. These items are the same as in CDS-I. CDS-II dropped B11a & B11e from CDS-I, which were created by the CDS-I research staff. The four items in the CDS-II PCG HH interview focus on parenting, in general, and are not specific to the CDS child. Items J15a, J15b are from the JOBS study. Two additional items were added in CDS-I (and kept in CDS-II).
Parental Warmth	E13a-g	G37a-f	Child Trends, Inc. for the JOBS Child Outcomes Study	The Parental Warmth scale measures the warmth of the relationship between the child and parent. These measures can also be used with items in the interviewer observation of the home environment in Section K that ask for interviewer report of maternal warmth observed during the visit (CDS-II K9, K13-K15). CDS-II kept items G37b-c; G37e-f but dropped G37a and d. CDS-II added E13e-g.
Extra-Curricular Activities, parent reported	G5-10c H5-H9e		Eccles & Simpkins	These items measure children's participation in after school lessons and sports, community, and religious activities. CDS-II queried PCG about frequency of and costs for involvement during school year and summer months.
Disciplinary Behaviors	F6a-h G15a-h H15a-h	F9a-g	Hofferth & Davis-Kean	PCG report on disciplinary actions taken within prior week. Items remained identical in wording and sequencing as in CDS baseline.
Response to Poor Grades	G16a-j H15a-j	F11a-k	Hofferth & Davis-Kean	PCG report on actions taken in response to poor grades. Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Encouragement of Reading	G18-G20		Eccles & Simpkins	PCG report of encouragement of reading and availability of dictionary and encyclopedia in the home.
Allowance	G21a1-G21b3		Eccles & Simpkins	These items measure frequency, amount, and conditions of an allowance.
Absent Parent Expenditures on Child	H24a-k H25	J30	NSFH; JOBS Study	Supplemental child support items to the PSID core, gathering information absent fathers' contributions.
Household Member Expenditures on Child	H26-H28h		Eccles & Simpkins	These items obtain information on household members' financial support of child for same items queried of absent parents.
Household Member Savings for Child	H29-H32a		Eccles & Simpkins	These items obtain information on household members' savings for child.
HOME Scale Items	See Table 3 for details	See Table 3 for details	NLSY Home SF. Based on: Caldwell, B., & Bradley, R. Home Observation for Measurement of the Environment. Little Rock, AR: University of Arkansas.	Measure of cognitive stimulation and emotional support in the home environment.
PCG Household Interview				
Neighborhood Measures	J1-J5	A1-A7 A11-A12	Used in NLSY; MIDMAC; Denver Youth Study; Human Development in Chicago Neighborhoods (McArthur)	These items assess residential stability (J1), residential satisfaction (J2), neighborhood anonymity (J3), social cohesion (J4a-J4d), neighborhood safety (J5). Items remained identical in wording and sequencing as in CDS baseline.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Participation in Community Activities	J6a-j	A8a-i	Human Development in Chicago Neighborhoods (McArthur Study)	CDS-II used the same items as CDS-I with three modifications: (1) added J6e “physical exercise, such as aerobics, running, or lifting weights” ; (2) changed J6h to read “community center” rather than “institution”; (3) CDSII asked the questions for caregivers and CDS children; CDS-I queried just for the caregivers.
PCG Religiosity	J7-J8	A9-A10	Detroit Area Study	These items reflect PCG report of importance of religion and spirituality in her/his life.
Rosenberg Self-Esteem Scale	J9a-j	A14a-j	Rosenberg, M. (1986). <i>Conceiving the Self</i> . New York: Basic Books.	This scale assesses PCG self-esteem. CDS-II used same items as CDS-I, and as the NLSY 1986-1990.
Pearlin Self-Efficacy Scale	J10a-d	A20a-g	Pearlin, L. I., Lieberman, M. A., Menaghan, E. G., & Mullan, J. T. (1981). The stress process. <i>Journal of Health and Social Behavior</i> , 22, 337-356.	CDS-II shortened the CDS-I scale version based on factor analysis and reliability assessment.
Household Tasks	J11a-n	A21a-p	NSFH	NSFH queried the respondent about the number of hours each HH member spends per week doing list of tasks. CDS queried the respondent about who usually is responsible for these tasks. CDS-II used the same items as CDS-I with two modifications: (1) CDS-II combined CDS-I A21a&b and (2) dropped A21i “bathing children, changing diapers”.
Child Rearing Values and Rules	J12	A22-A25	Alwin, D. F. (1990). Cohort Replacement and Changes in Parental Socialization Values. <i>Journal of Marriage and the Family</i> , 52, 347-60.	These items appeared in the Detroit Area Study and the General Social Survey. They were used at both waves but shortened for CDS-II.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Parenting Attitudes and Opinions	J16a-d	A33-A36	Hofferth & Davis-Kean	This set of items was developed for CDS-I to supplement the items on parental attitudes and values in child rearing. Item series was shorted for CDS-II.
Gender Roles	J17a-s	A37a-cc	Hofferth & Davis-Kean Based on items in the NSFH, supplemented with items from Being a Father Scale (Pleck, 1997) and “Role of the Father” study (Palkowitz, 1984).	These items measure attitudes towards gender roles in parenting and was constructed for CDS-I based on items from several established studies. Item series was shorted for CDS-II.
Psychological Distress	J18a-h Also-PSID	A38a-j	Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.-L.T., Walters, E.E., & Zaslavsky, A. (2002). Short screening scales to monitor population prevalences and trends in nonspecific psychological distress. <i>Psychological Medicine</i> , 32(6), 959-976.	This measure is the six-item scale developed by Ron Kessler for use in the National Health Interview Survey. CDS-II used the exact scale as NHIS, and CDS-I, although CDS-I also includes the 10-item version. For scoring instructions, see: http://www.hcp.med.harvard.edu/ncs/k6_scales.php
Parent Disagreement	J20a-e	A40a-i	NLSY NSFH	These items were taken verbatim from the NLSY and NSFH for CDS-I. They measure the extent of agreement between parents on daily activities. The set of items was shortened for CDS-II.
Joint Goals	J21a-c	A41a-f	Hofferth & Davis-Kean – developed for CDS-I	These items were designed for CDS-I to measure the extent to which respondent and partner have joint marital/family goals for the future. The set of items was shortened for CDS-II.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Family Conflict	J22a-e	A42a-f	Sweet, J., Bumpass, L., & Call, V. (1988). The Design and Content of the National Survey of Families and Households. NSFH Working Paper No. 1. Madison, WI: Center for Demography and Ecology, University of Wisconsin	These items were taken verbatim from NSFH to examine methods of conflict resolution among family members. The set of items were shortened for CDS-II.
Economic Strain	J25a-o	A53 (circle all that apply)	Conger, R., & Elder, G. H., Jr. (1994). Families in Troubled Times: Adapting to Changes in Rural America. New York: Aldine de Gruyter.	These items were drawn from G. Elder's and R. Conger's work measuring experiences of economic or financial stress and strain and practical responses to such financial pressures (Conger & Elder, 1994). This series of YES/NO items was used at both waves but shortened for CDS-II.
Media Use	J36a-d J37a-f		Eccles	Set of items added to CDS-II concerning electronic media use in the household.
PCG work schedule	J42-J48		Current Population Survey (CPS) Supplement	Modified sequence to capture the PCG's typical work schedule.
Social Support	J49a-f		Eccles & Simpkins	Set of items measuring PCG social support network.
Child Assessments				
Body Mass Index (BMI)	Q24BMI WTIND02 Q24A1- Q24A4c	PCG A2- A3	CDC	In CDS-II, both height and weight were measured by the interviewer. In the situations whereby the parent or child refused, the interviewer obtained PCG report of height and weight at the last doctor's visit. BMI is calculated from items A1-A4c. In CDS-I, children were measured but the weight was reported by the PCG. In children, BMI-for-age is plotted on age & gender specific growth charts (see http://www.cdc.gov/growthcharts). The body weight indicator variable that classifies the child as under or overweight, normal, or obese is labeled "WTIND02".

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Woodcock-Johnson Revised Tests of Achievement for Reading and Math	Q24LW1 Q24AP60 Q24LWSS Q24PCSS Q24APSS Q24BRSS	ASM LW1-AP60	Woodcock, R. W., & Johnson, M. B. (1989). Tests of Achievement, Standard Battery. Chicago, IL: Riverside Publishing.	The WJ-R measures achievement in several domains. CDS-II used three subtests to measure reading and math skills (the Letter-Word, Passage Comprehension, and Applied Problems tests). In CDS-I, these tests were chosen due to the ease of administration as well as their brevity, and then repeated in CDS-II. When applicable, the Spanish version of the WJ-R (Batería-R, Form A). National norms for English-speaking and Spanish-speaking respondents are available for these tests. A standardized score is provided for each of the three tests as well as a combined “Broad Reading” score.
WISC Digit Span for Short-Term Memory	ASM C1-C15c	ASM B1-B15c	Wechsler, D. (1974). Wechsler Intelligence Scales for children-Revised. New York: The Psychological Corporation.	Also used in the NLSY-Child Study. A total test score is provided as well as scores for each of the forward and backward item sets.
Child Interview				
Ability Self-Concepts of Children	E1-E20 (C1-C20 in CDS-I)		Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender difference in children's self-and task perceptions during elementary school. <i>Child Development</i> , 64, 830-847.	These items reflect a set of scales developed and validated by Jacquelyne Eccles for the domains of math and reading. The same items were used both in CDS-I and CDS-II, and are asked of children 8 years and older.
Children’s Expectations about Family and Future	E21a-c		ADD Health	Three items adapted from ADD Health Study that measure the child’s concern over getting a good job as an adult, the child’s future, and family finances.
Connectedness to School	E22a-d		Eccles	Four items developed for CDS that measure the degree of inclusiveness, closeness, and happiness with, as well as safety at school.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Electronic Media Use	Section F		Eccles	These items were developed for CDS-II to collect information about children's use of electronic media.
Peer Bullying/ Victimization	H1 a-d		Kochenderfer, B.J. & Ladd, G.W. (1996). Peer victimization: Cause or consequence of school maladjustment? Child Development, 67, 1305-1317.	Set of four items designed to assess frequency of peer victimization.
Self-Rated symptoms of: feeling really sick; waking up tired; having skin problems; dizziness; chest pains; headaches; muscle soreness; and stomachache	H2a-h		ADD Health; NHIS	Series of items that query the child about a number of physical symptoms.
Closeness to Others	H4a-i H5a-h H6		Eccles & Simpkins	H4 assesses the frequency in which the child converses with mother, father and friends about school, future plans, and friends. H5-6, in turn, queries the child about his/her closeness with family members and other adults.
Race	J1		Eccles & Simpkins	Adolescents' self-reported race.
Religiosity and Spirituality	J2-J5a		MSALT	Series of questions developed for CDS-II about the child's religious and spiritual life.
Employment	J6-J33		Eccles	Set of question items developed for CDS-II that collect information about part-time and summer employment. We ask the child about job title/ type of work, job characteristics, job satisfaction, hours, and wages, as well as future aspirations for work and career.
Future Expectations	J34-J36 L9-L10		ADD Health	Questions adapted from ADD Health to assess the child's self-rated probability of on family formation, schooling, employment, and life-expectancy.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
School Courses	J37-J39		Eccles	Detailed series of items for which Children in grades 6-12 provided the name of the class and the grade received in the class for six instructional areas (English, math, science, social science, foreign language, vocational) for two prior terms. We ask for up to three "mentions" for each instructional area.
After School and Community Activities	K1-K7		Eccles	Parallel set of questions to the PCG that queries the Child about time spent in after school and community activities and involvement of friends.
General Health Status of Child – Self-Report by Child	K8		ADD Health	Single item that asks the child to rate his/her own health on a scale of 1 (Excellent) -5 (Poor).
Weight Maintenance	K9-K12		ADD Health	CDS youth aged 10 years and older are asked in ACASI module to self-report strategies used in the past seven days to gain or loose weight.
Dietary Habits	K13-K14		ADD Health	These items in ACASI module collect information from the youth aged 10 years and older about foods consumed from each of the major food groups.
Exercise	K15-K17		ADD Health	These items in ACASI module collect information from the youth aged 10 years and older about amount of exercise in school and outside of school during the prior week.
Limitations-Self-Report by Youth	K18-K19		ADD Health	These items in ACASI module collect information from the youth aged 10 years and older about frequency in the last month they missed school or social/recreational activity due to health or emotional health problem.
Sleep	K20-K22		ADD Health	Youth aged 10 years and older are asked in ACASI module time they go to sleep and amount of sleep they obtain on weeknights. They are also asked their opinion if they get sufficient amount of sleep.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Perseverance	K26a-e		Furstenberg, Frank F. Jr.; Thomas D. Cook; Jacquelynne Eccles, Glen H. Elder, Jr. and Arnold Sameroff. (1999). <i>Managing to Make It: Urban Families and Adolescent Success</i> . University of Chicago Press.	Developed for the MacArthur Network on Adolescent Development; asked of youth aged 10 and older in ACASI.
Pro-Social Behaviors	K24a-f		Eccles	Developed for CDS-II, these questions address helping and social support.
Friend Characteristics	K25a-n		MSALT	Detailed list of positive and negatives behaviors of the child's friends.
Subscale Global Self Concept	K27-K32 (C21-C28)		Marsh, H. (1990). <i>Self-Description Questionnaire Manual</i> . Macarthur, Australia: Univ. of Western Sydney.	Six items that assess the child's self-reported self-concept.
Savings and Expenditures	L1-L6		Eccles & Simpkins	Detailed set of questions about allowance, savings habits and expenditures.
Dating	L7-L8		ADD Health	Adolescents are asked whether parents allow them to date-supervised and unsupervised and the age of first date.
Deviant Behaviors	L11a-j		NLSY79 http://www.bls.gov/nlsy79cyaguide/nlsy79cusg.htm	Adolescents are asked about frequency of norm-breaking behaviors in the past six months that range in severity from staying out past curfew to harming others badly enough to require medical attention. NLSY Items "ever used marijuana" and "ever got drunk" are asked in an earlier section of the CDS ACASI and are part of this overall measure.
Tobacco, Alcohol, and Drug Use	L12-L15		ADD Health	Questions that collect information about experiences with tobacco, alcohol, and drugs. Adolescents are asked if the ever tried, age first tried, frequency of use, and peer use.
Children's Depression Inventory (CDI)	L16a-j		Kovacs, Maria. <i>Children's Depression Inventory (CDI)</i> . http://www.pearsonassessments.com/tests/cdi.htm	The CDI assesses youth self-report symptoms of depression during the prior two weeks to the interview.

Measure	CDS-II Items	CDS-I Items	Source/Original Authors	Notes
Acceptance Subscale from Child Report of Parent Behavior Inventory (CRPBI)	L17 and L18: c, e, f, h L19a-d		Barber, B.K. & Olsen, J.A. (1997). Socialization in context: Connection, regulation, and autonomy in the family, school, and neighborhood, and with peers. <i>Journal of Adolescent Research: Special Issue: Adolescent Socialization in Context: Connection, regulation, and autonomy in multiple contexts, Part II, 12(2), 287-315.</i>	This set of items adapted from the CRPBI and work by B. Barber intends to measure acceptance by each parent, friends, and boy/girl friends. Four out of the original ten items were retained for CDS-II, based on factor analyses.
Psychological Control Scale-Youth Self Report	L17 and L18: a, b, d, g, i, j		Barber, B.K. (1996). Parental psychological control: Revisiting a neglected construct. <i>Child Development, 67, 3296-3319.</i>	This set of items adapted from work by B. Barber intends to measure psychological control, as reported by the child. Six out of the original eight items were retained for CDS-II, based on factor analyses.
Social Initiative	L22a-e		Barber, B.K. (1996). Parental psychological control: Revisiting a neglected construct. <i>Child Development, 67, 3296-3319.</i>	This set of items adapted from Bachman et al (1993) and work by B. Barber intends to measure social initiative in school context. Five out of the original 13 items were retained for CDS-II, based on factor analyses.
Monitoring	L23-L25 L26-L28		Stattin & Kerr	Monitoring: Parents Knowledge; Child Disclosure
Sexual Experiences	L29-L36		ADD Health	Set of questions querying the youth about experiences with sexual intercourse and pregnancy.
Subjective Well Being	L37a-c L37d-h L37k-n		MIDUS http://midmac.med.harvard.edu/	This set of items adapted from MacArthur MIDUS Youth and work by Corey Keyes intends to measure Subjective Well-Being among adolescents. Some of the original items from the MIDUS were retained for each of these subscales for the CDS-II, based on factor analyses.

Appendix B: HOME Scale Mapping to NLSY and PHDCN

HOME Scale Item	NLSY Subscale ²⁷	PHDCN ²⁸ Subscale	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
1. How often child gets out of the house	C	--	C1							S	
2. How many books child has read	C	ARM	B16	B16	B16	B16	E12	E12	E12	S	
3. How often mother takes child to grocery store	C	--	B15	B15 ³⁰	B15					S	
4. How many cuddly or role-playing toys child has	C	--	C2							S	
5. Mother's belief about how child learns best	C	--	C4							S	
6. How many push or pull toys child has	C	--	C3							S	
7. How often mother talks to child while working	E	--	C5							S	
8. Mother reads to child	C	ARM	B14	B14	B14	B14	F8	-	-	S	

²⁷ E=Emotional Support; C=Cognitive Stimulation

²⁸ Project on Human Development in Chicago Neighborhoods (PHDCN). Scale abbreviations are: ARM=Access to Reading Materials; EG=Exterior Home; IH=Interior Home; LS=Learning Stimulation; O/A=Outings/Activities; PA=Parental Activity; PEC=Parental Engagement with Child; PLH=Parental Lack of Hostility; PVS=Parental Verbal Skills; PW=Parental Warmth

²⁹ 'S' indicates 'self-report'; 'O' indicates interviewer observation

³⁰ An underlined question number indicates that the item was asked of the cohort, but is not part of the NLSY subscale

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
9. How often child eats with mother and father	E	--	B5	B5	B5	B5	E5	E5	E5	S	
10. Child's play environment appears safe	C	IH	L14	L14	L14	L14	K31	K31	K31	O	In CDS, the item is worded 'is play environment safe'; in PHDCN as 'house or environment free of potentially dangerous structural or health hazards.' ³¹
11. Mother conversed with child at least twice	E	PW	L2	L2	L2	L2	K7	K7	K7	O	Scale is 1 to 5 (never=1;4 or more times=4) instead of yes/no.+ Different scale than 1997. ^{*32}
12. Mother answered child's questions/speech verbally	E	PW	L3	L3	L3	L3	K8	K8	K8	O	Scale is 1 to 5 (never=1;4 or more times=4) instead of yes/no.+ Different scale than 1997.*
13. Mother caressed, kissed, or hugged child at least once	E	PW	L4	L4	L4	L4	K9	K9	K9	O	Scale is 1 to 5 (never=1;4 or more times=4) instead of yes/no.+ Different scale than 1997.*
14. Mother slapped or spanked child at least once	E	PLH	L5	L5	L5	L5	K10	K10	K10	O	Scale is 1 to 5 (never=1;4 or more times=4) instead of yes/no.+ Different scale than 1997.*
15. Mother physically restricted child	E	--	L6	L6	L6	L6	K11	K11	K11	O	Different scale than 1997.*
16. Mother provided toys or interesting activities	C	--	L7	L7	L7	L7	K12	K12	K12	O	Different scale than 1997.*
17. Mother kept child in view	E	--	L9	L9	L9	L9				O	Different scale than 1997.*
18. How many magazines family gets	C	PA	HH-A27	HH-A27	HH-A27	HH-A27	J13	J13	J13	S	

³¹ +Wording difference from PHDCN

³² *Wording difference from 1997 to 2002

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
19. Does child have record/tape/CD & at least 5 records/tapes/CDs	C	--		D1			F1			S	
20. How often child was taken to a museum	C	O/A		D2	E5	F5	F2	G12	H12	S	
21. How often child is taken on an outing	C	O/A		D3			F3			S	
22. Child is encouraged to learn numbers	C	LS		D4a			F4a			S	
23. Child is encouraged to learn alphabet	C	ARM		D4b			F4b	G4		S	In 2002, G4 asks if child has toys that help him/her learn the alphabet OR words+
24. Child is encouraged to learn colors	C	LS		D4c			F4c			S	
25. Child is encouraged to learn shapes and sizes	C	LS		D4d			F4d			S	
26. Home interior is dark or monotonous	C	IH	L10	L10	L10	L10	K28	K28	K28	O	Scale is 1 to 5; (not at all monotonous=1; very monotonous=4) instead of yes/no+ 1997 scale was reversed.*
27. All visible rooms are minimally cluttered	C	IH	L11	L11	L11	L11	K29	K29	K29	O	Scale is 1 to 5; (not at all cluttered=1; very cluttered=4) instead of yes/no+; 1997 scale was reversed.*
28. All visible rooms are reasonably clean	C	IH	L12	L12	L12	L12	K30	K30	K30	O	Scale is 1 to 5; (not at all clean=1; very clean=4) instead of yes/no+; 1997 scale was reversed.*
29. Amount of choice child has in selected food	E	--		D5			F5			S	

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
30. Mother's response to hitting: hit child back	E	--		D6			F7			S	
31. Mother's response to hitting: spank child	E	--		D6	E11	F10	F7k	G17b	H17b	S	
32. Mother's response to hitting: Other	E	--		D6	E11	F10	F7b	G17b	H17b	S	
33. Number of hours per day TV is on in home	E	--	HH-A26a	HH-A26a	HH-A26a	HH-A26a	J27a	J27a	J27a	S	
34. Mother's voice showed positive feeling about child	E	PW	L8	L8	L8	L8	K13	K13	K13	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no+; Different scale than 1997.*
35. Mother introduced interviewer to child by name	E	--		L17	L17	L17	K16	K16	K16	O	
36. How often child spanked in past week	E	--			E9b	F9b	F6b	G15b	H15b	S	
37. How often child reads for enjoyment	E	--			E1	F1		G1	H1	S	
38. Does family encourage hobbies	C	--			E2	F2		G2	H2	S	
39. Is there a musical instrument child can use at home	C	LS			E3	F3		G3	H3	S	G3/H3 asks if they have access to a musical instrument. CDS does not ask if they have access to a musical toy.+ G3a/H4 asks how often they play (every day, several times a week, several times a month, several times a year, or never).

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
40. Does child get special lessons or activities	C	--			E4	F4		G5-G10c	H5-H8c	S	In 2002, CDS does not have a question about general activity involvement. CDS has several questions about children's involvement in specific activities.*
41. How often child was taken to theater	C	O/A			E6	F6		G13	H13	S	
42. Do parents discuss TV programs with child	C	PEC			E7	F7	J35	J35	J35	S	In CDS, the question asks if primary or other caregiver discuss TV with child. CDS does not ask about the last 2 weeks specifically+
43. Does family get daily newspaper	C	PA	HH-A28	HH-A28	HH-A28	HH-A28	J14 & J14a	J14; J14a	J14; J14a	S	J14 asks if the family received the daily newspaper. J14a asks how many days a week they read the newspaper+
44. How often child spends time with father	E	--	B3	B3	B3	B3	E4a	E4a	E4a	S	In 2002, the wording was changed from "time spent with father" to "time spent with father in indoor activities."*
45. How often child spends time with father outdoors	E	--	B4	B4	B4	B4	E4	E4	E4	S	
46. How often family gets together with friends or relatives	E	--	B7	B7	B7	B7	E6	E6	E6	S	
47. How often child expected to make own bed	E	--			E8a	F8a		G14a	H14a	S	
48. How often child expected to clean own room	E	--			E8b	F8b		G14b	H14b	S	

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
49. How often child expected to clean up spills	E	--			E8c			G16c		S	
50. How often child expected to bathe self	E	--			E8d			G14d		S	
51. How often child expected to pick up after self	E	--			E8e	F8f		G14e	H14f	S	
52. How often child expected to help manage own time	E	--				F8e			H14e	S	
53. How often child keeps shared living areas clean	E	--				F8c			H14c	S	
54. How often child does routine chores	E	--				F8d			H14d	S	
55. Child included in family hobby	--	O/A					E15	E15	E15	S	
56. Parent engages in regular outdoor recreation w/ child	--	O/A					E2	E2	E2	S	E4 discusses father involvement in outdoor activities w/child & E5 discusses mother involvement in outdoor activities w/child. The two can be used to gauge parent involvement.+
57. Buildings in the face block in good condition	--	EH					K37	K37	K37	O	
58. Street in the face block in good condition	--	EH					K38	K38	K38	O	
59. Bi garbage or broken glass in street or sidewalk	--	EH					K39	K39	K39	O	

HOME Scale Item	NLSY Subscale ²⁷	PHDCD Subscale ²⁸	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
60. No drug, alcohol, or cigarette litter in street/sidewalk	--	EH					K40	K40	K40	O	
61. How many books are in the house	--	ARM					J39	J39	J39	S	
62. How many dictionaries are in the house	--	ARM					--	G19	G19	S	CDS asks how many dictionaries are in the home instead of yes/no.+
63. Parent encourages child to contribute	--	PW					K18	K18	K18	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
64. Parent mentions skill of child	--	PW					K17	K17	K17	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
65. Parent praises child twice during visit	--	PW	L13	L13	L13	L13	K14	K14	K14	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+ Different scale than 1997.*
66. Parent uses diminutive for child's name	--	PW					K20	K20	K20	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
67.											
68.											
69. Parent responds positively to praise of child	--	PW					K19	K19	K19	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
70. Parent's speech is distinct, clear, audible	--	PVS					K1	K1	K1	O	Scale is 1 to 5 (never=1; Often=5) instead of yes/no.+
71. Parent initiates verbal interchanges	--	PVS					K4	K4	K4	O	Scale is 1 to 5 (never=1; Often=5) instead of yes/no.+

HOME Scale Item	NLSY Subscale ²⁷	PHDCD ²⁸ Subscale	1997 Questions				2002 Questions			S/O ²⁹	Notes on Changes
			0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.		
72. Parent expresses ideas freely and easily	--	PVS					K3	K3	K3	O	Scale is 1 to 5 (never=1; Often=5) instead of yes/no.+
73. Parent appears to understand questions	--	PVS					K2	K2	K2	O	Scale is 1 to 5 (never=1; Often=5) instead of yes/no.+
74. Parent does not shout at child during visit	--	PLH					K22	K22	K22	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
75. Parent does not express annoyance with child	--	PLH					K23	K23	K23	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
76. Parent does not scold or criticize child	--	PLH					K21	K21	K21	O	Scale is 1 to 5 (never=1; 4 or more times=4) instead of yes/no.+
77. Home has 100 square feet space per person	--	IH					K32	K32	K32	O	
78. Rooms are not overcrowded with furniture	--	IH					K33	K33	K33	O	Scale is 1 to 5 (not at all overcrowded=1; very overcrowded=5) instead of yes/no.+
79. Home is not too noisy (from noise in house)	--	IH					K34	K34	K34	O	Scale is 1 to 5 (never=1; Often=5) instead of yes/no.+
80. Home is not too noisy (from noise outside)	--	IH					K35	K35	K35	O	Scale is 1 to 5 (not at all noisy=1; very noisy=5) instead of yes/no.+
81. No signs of alcohol or illegal drug use	--	IH					K36	K36	K36	O	Scale is 1 to 5 (none=1; yes, just about everywhere=5) instead of yes/no.+

Appendix C: HOME Scale Recoding Instructions

Recoding Specifications for the CDS-II HOME-SF Items ³³									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.	
1. How often child gets out of the house	C	C1							
2. How many books child has	C	B16	B16	B16	B16	E12	E12	E12	If response=(1, 2) then recode to 0 If response=(3, 4, 5) then recode to 1
3. How often mother takes child to grocery store	C	B15	B15	B15	B15				
4. How many cuddly or role-playing toys child has	C	C2							
5. Mother's belief about how child learns best	C	C4							
6. How many push or pull toys child has	C	C3							
7. How often mother talks to child while working	E	C5							

³³ An underlined question number indicates that the item was asked of the cohort, but is not part of the NLSY subscale and therefore not used in the construction of the summary variable for Emotional Support and Cognitive Stimulation subscales in the CDS-II data. An empty shade cell indicates that the item is not part of the NLSY subscale and therefore not used in the construction of the summary variable for Emotional Support and Cognitive Stimulation subscales in the CDS-II data. An italic question number indicates that the item was asked of the cohort, and it is part of the NLSY subscale, but it was not used to construct the summary variable for Emotional Support and Cognitive Stimulation subscales in the CDS-II data because of a significant number of missing data.

Recoding Specifications for the CDS-II HOME-SF Items33									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs	10+ yrs	
8. Mother reads to child	C	B14	B14	B14	B14	F8	-	-	If response=(1, 2, 3, 4) then recode to 0 If response=(5, 6) then recode to 1
9. How often child eats with both mother and father	E	B5	B5	B5	B5	E5	E5	E5	If response=(1, 2) then recode to 0 If response=(3, 4, 5, 6, 7) then recode to 1
10. Child's play environment appears safe	C	L14	L14	L14	L14	K31	K31	K31	If response=(5) then recode to 0 If response=(1) then recode to 1
11. Mother conversed with child at least twice	E	L2	L2	L2	L2	K7	K7	K7	If response=(1,2) then recode to 0 If response=(3, 4)then recode to 1
12. Mother answered child's questions/speech verbally	E	L3	L3	L3	L3	K8	K8	K8	If response=(1) then recode to 0 If response=(2, 3, 4) then recode to 1
13. Mother caressed, kissed, or hugged child at least once	E	L4	L4	L4	L4	K9	K9	K9	If response=(1) then recode to 0 If response=(2, 3, 4) then recode to 1
14. Mother slapped or spanked child at least once	E	L5	L5	L5	L5	K10	K10	K10	If response=(1) then recode to 0 If response=(2, 3, 4) then recode to 1
15. Mother physically restricted child	E	L6	L6	L6	L6	K11	K11	K11	If response=(1) then recode to 0 If response=(2, 3, 4) then recode to 1
16. Mother provided toys or interesting activities	C	L7	L7	L7	L7	K12	K12	K12	
17. Mother kept child in view	E	L9	L9	L9	L9				
18. How many magazines family gets	C	HH A27	HH A27	HH A27	HH- A27	J13	J13	J13	If response=(0) then code as 0 If response=(1+) then recode to 1
19. Does child have record /tape/CD & at least 5 records/tapes/CDs	C		D1			F1 & F1a			If F1=5 then recode as 0 If F1=1 and F1a < 5 then recode as 0 If F1=1 and F1a =>5 then recode as 1

Recoding Specifications for the CDS-II HOME-SF Items33									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.	
20. How often child was taken to a museum	C		D2	E2	F2	F2	G12	H12	If response=(1) then code as 0 If response=(2, 3, 4, 5) then recode to 1
21. How often child is taken on an outing	C		D3			F3			If response=(1, 2) then recode to 0 If response=(3, 4,5) then recode to 1
22. Child is encouraged to learn numbers	C		D4a			F4a			If response=(5) then recode to 0 If response=(1) then recode to 1
23. Child is encouraged to learn alphabet	C		D4b			F4b			If response=(5) then recode to 0 If response=(1) then recode to 1
24. Child is encouraged to learn colors	C		D4c			F4c			If response=(5) then recode to 0 If response=(1) then recode to 1
25. Child is encouraged to learn shapes and sizes	C		D4d			F4d			If response=(5) then recode to 0 If response=(1) then recode to 1
26. Home interior is dark or monotonous	C	L10	L10	L10	L10	K28	K28	K28	If response=(2, 3, 4, 5) then recode to 0 If response=(0) then recode to -4
27. All visible rooms are minimally cluttered	C	L11	L11	L11	L11	K29	K29	K29	If response=(2, 3, 4, 5) then recode to 0 If response=(0) then recode to -4
28. All visible rooms are reasonably clean	C	L12	L12	L12	L12	K30	K30	K30	If response=(1) then recode to 0 If response=(2, 3, 4, 5) then recode to 1 If response=(0) then recode to -4
29. Amount of choice child has in selecting food	E		D5			F5			If response=(1,2 then recode to 0 If response=(3,4) then recode to 1

Recoding Specifications for the CDS-II HOME-SF Items33									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.	
30. Mother's response to hitting: hit child back	E		D6			F7B			If Respondent indicated harsh reaction to F7: "hit back", "spank", "kick out of house" or "eat soap" then recode to 0 ELSE recode new variable=1
31. Mother's response to hitting: spank child	E		D6	E11	F10	F7K	G17	H17	
32. Mother's response to hitting: kick out of house	E		D6	E11	F10	F7	G17	H17	
33. Mother's response to hitting: eat soap	E		D6	E11	F10	F7	G17	H17	
34. Number of hours per day TV is on in home	E	HH-A26	HH-A26a	HH-A26a	HH-A26a	J27a	J27a	J27a	If response=(0, 1, 2, 3, 4) then recode to 1 If response=(5+) then recode to 0
35. Mother's voice showed positive feeling about child	E	L8	L8	L8	L8	K13	K13	K13	If response=(1) then recode to 0 If response=(2, 3, 4, 5) then recode to 1
36. Mother introduced interviewer to child by name	E		L17	L17	L17	K16	K16	K16	If response=(1) then recode to 1 If response=(5) then recode to 0S
37. Mother encouraged child to the conversation	E					K18	K18	K18	If response=(1) then recode to 0 If response=(2, 3, 4, 5) then recode to 1
38. How often child spanked in past week	E			E9b	F9b	F6b	G15b	H15b	IF Q21E10=5 then recode to 0 If Q21E10=1 & F6b/G15b/H15b=0, 1 then recode to 1 If F6b/G15b/H15b = 2 or more then recode to 0
39. How often child reads for enjoyment	C			E1	F1		G1	H1	If response=(5, 6) then recode to 1 If response=(1, 2, 3, 4, 7) then recode to 0
40. Does family encourage hobbies	C			E2	F2		G2	H2	If response=(1) then recode to 1 If response=(5) then recode to 0
41. Is there a musical instrument child can use at home	C			E3	F3		G3	H3	If response=(1) then recode to 1 If response=(5) then recode to 0

Recoding Specifications for the CDS-II HOME-SF Items33									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10+ yrs.	
42. Does child get special lessons or activities	C			E4	F4		G5- G10	H5- H9c	IF (G5 or G6a or G6c or G7a or G8 or G8c or G8f or G10) =1 Recode new variable=1 Else new variable=0 IF (H5 or H6 or H7 or H7b or H7d or H8 or H9a or H9c) =1 Recode new variable=1 Else new variable=0
43. How often child was taken to theater	C			E6	F6		G13	H13	If response=1 then recode to 0 If response=(2, 3, 4, 5) then recode to 1
44. Do parents discuss TV programs with child	C			E7	F7	J35	J35	J35	If response=(1) then recode to 1 If response=(5) then recode to 0
45. Does family get daily newspaper	C	HH A2 8	HH A28	HH A28	HH- A28	J14	J14	J14	If response=(1) then recode to 1 If response=(5) then recode to 0
46. How often child spends time with father	E	B3	B3	B3	B3	E4 & E4a	E4 & E4a	E4 & E4a	If E4 or E4a = (7, 6) then recode to 1 Else recode to 0
47. How often child spends time with father outdoors	E	B4	B4	B4	B4	E4	E4	E4	If response=(0, 1, 2, 3, 4) then recode to 0 If response=(5, 6, 7) then recode to 1
48. How often family gets together with friends or relatives	E	B7	B7	B7	B7	E6	E6	E6	If response=(1, 2, 3) then recode to 0 If response=(4, 5) then recode to 1
49. How often child expected to make own bed	E			E8a	F8a		G14a	H14a	If at least 3 out of 4 of the items G14a-G14d => 2 Recode new variable = 1 Else new variable=0;
50. How often child expected to clean own room	E			E8b	F8b		G14b	H14b	

Recoding Specifications for the CDS-II HOME-SF Items33									
Item	NLSY79 Subscale	1997 Questions				2002 Questions			Recoding Instructions
		0-3 yrs.	3-5 yrs.	6-9 yrs.	10-12 yrs.	3-5 yrs.	6-9 yrs.	10 + yrs.	
51. How often child expected to ... clean up spills	E			E8c			G14c	--	If at least 3 out of 4 of the items H14a, b, d, e => 2 Recode new variable = 1 Else new variable=0 ; If G14e / H14f => 2 then new variable=1 If G14e / H14f =(1) then new variable=0 ;
52. How often child expected to ... bathe self	E			E8d			G14d	--	
53. How often child expected to ...pick up after self	E			E8e	F8f		G14e	H14f	
54. How often child expected to ...help manage own time	E				F8e			H14e	
55. How often child keeps shared living areas clean	E				F8c			H14c	
56. How often child does routine chores	E				F8d			H14d	

Appendix D: CDSII Data Groupings

Data File	Description	# variables	Unit	Eligible Participants	Data Records	Response Rate
2002 Demographic File	The “Demographic File” (as labeled in 1997) contains sample weight, family relationship variables, a sibling indicator, interviewer ID, and result status of the individual modules for each child.	19	One record for each child.	3,191	2,907	91%
Primary Caregiver (PCG) Child File	The PCG-Child data were collected about each target child as reported by the primary caregiver. The interview content largely reflects CDS-I PCG interview, with the addition of new content for the adolescents. The file includes interviewer observations of PCG-Child interactions from the HOME Scale. Interviews were conducted by telephone or in person.	829	One record for each child.	3,191	2,907	91%
Primary Caregiver (PCG) Household File	The PCG-Household data contain information about the target child's household, neighborhood, and PCG social/emotional well-being as reported by the PCG. Each PCG completed one household questionnaire regardless of the number of children in the sample for whom they act as primary caregiver. The file includes HOME Scale interviewer observations of physical environment. Interviews were conducted by telephone or in person. A total of 2,019 families participated.	215	One record for each child.	3,191	2891	91%
Assessment Data	CDS conducted in person reading and math achievement tests using the Woodcock Johnson Revised Tests of Achievement, short-term memory tests using the WISC Digit Span, and height and weight measurements.	239	One record for each child.	2,907	2,644	91%
Child File	CDS youth aged 8 years and older were eligible to participate in a personal interview. The interview content was largely new to CDS-II and included an ACASI module for 10+ years.	396	One record for each child.	2,656	2,182	82%
Time Diary Questionnaire	The Time Diary Questionnaire File contains survey administrative data for the diary. This data set has a single record for each child, which includes information for both the weekday and weekend diaries. Time diaries were mailed ahead to the home, and reviewed with the child and/or PCG in person or by telephone.	72	One record for each child.	2,907	2,569	88%

Data File	Description	# variables	Unit	Eligible Participants	Data Records	Response Rate
Time Diary Activity	The Time Diary Activity File contains detailed information about the target child's activities during a twenty-four hour period on both a weekday and a weekend day.	31	One record for each activity.	2,907	99,467	See Time Diary Q'naire
Time Diary Aggregate	The Time Diary Aggregate file "roll-up" individual activities in several ways: at the four-and three-digit code levels, which take all of the time spent in each of the activities and sum them across the diary day, and at the two-digit level, which aggregate time per diary day by 38 activity categories.	1311	One record for each child.	2,907	2,569	88%
Other Caregiver (OCG) Child File	The OCG-Child file contains data about each child as reported by the secondary caregiver (OCG. The items are a subset of the PCG-Child file, administered by mail with telephone follow-up.	78	One record for each child.	2,009	1,686	84%
Other Caregiver (OCG) Household File	The OCG-Household file contains a subset of items from the PCG-Household questionnaire, as reported by the OCG. Each OCG completed only one household questionnaire regardless of the number of children in the sample for whom they act as the other caregiver. This was a paper mail survey with telephone interview follow up on non response.	98	One record for each child.	2,009	1,686	84%
Elementary/Middle School Teacher File	The Elementary/Middle School Teacher File contains teacher reported data about the target child and school environment. 76% of the PCGs who had a child in elementary school also gave permission to contact teachers; 71% of the teachers were interviewed by telephone.	152	One record for each child.	1,305	699	54%
School Environment	The school environment file consist of school or school-district level identifiers extracted from the NCES Common Core of Data for the school/district that the CDS target child attended during the CDS-II data collection. Additionally includes linkages to the NCES Private School Survey. Data available under restricted contract.	6	One record for each child.	2,538	2,203	91%
Media Files	The Center for Research on Interactive Technology, Television and Children (CRITC) at the University of Texas at Austin coded the television programs that were reported in the 1997 and 2002 CDS Time Diaries.	8	One record for each activity.	2,907	11,975	See Time Diary Q'naire

Appendix E: Documentation for CDS Media Files

In 1997, the PSID supplemented its main data collection with additional information on 0-12 year-old children and their parents in a study entitled "The Child Development Supplement" (CDS). The objective was to provide researchers with a comprehensive, nationally representative, longitudinal data base of children and their families with which to study the dynamic process of early human capital formation. The CDS has collected information on this sample at three points in time – the initial wave was in 1997 (CDS-I), a second wave was collected when the children were 5 years older, aged 5-18 in 2002/2003, and a third and final wave was collected in 2007/2008 when the children were 10-18. More information about the study and data are provided in the CDS User Guides: <http://psidonline.isr.umich.edu/CDS/wavesdoc.html>. The time diaries are one of the unique features of the CDS survey design. With the assistance of their primary care giver when needed, CDS children completed a set of time diaries as part of their involvement in the overall CDS data collection effort: one diary for a randomly selected weekday (Monday-Friday) and another for a randomly selected weekend day (Saturday or Sunday). The time diaries provide detailed accounting of the type, number, duration, and location of activities during sampled 24-hour days, beginning at midnight. The time diaries additionally collected information on the social context of the activity by specifying with whom the child was doing the activity and who else was present, but not engaging, and the names of television shows watched by the children.

The Center for Research on Interactive Technology, Television and Children (CRITC) at the University of Texas at Austin coded the television programs that were reported in the 1997 and 2002 CDS Time Diaries. The television coding manual and the data files are now available. This document describes the data structure and coding procedures. For clarification questions, please contact Dr. Elizabeth Vandewater at evandewater@mail.utexas.edu.

Data Structure

The data file is structured at the activity level. There is one record per media activity with a valid content code for each of the weekday / weekend diaries for each child for each wave of data collection (1997 data and 2002 data are kept in separate files). Key identifier variables that will help you link the media codes back to the main time diary activity file are as follows (these are wave specific).

Unique Identifiers for 1997 Data:

TD97M02	1997 Interview Number (corresponds to ER33401)
TD97M03	1997 Sequence Number (corresponds to ER33402)
TD97M04	Type of Diary (corresponds to WDAYWEND)
TD97M05	Start Time (corresponds to COLB)
TD97M06	End Time (corresponds to COLC)

Unique Identifiers for 2002 Data:

- TD02M02 2001 Interview Number (corresponds to ER33601)
- TD02M03 2001 Sequence Number (corresponds to ER33602)
- TD02M04 Type of Diary (corresponds to DIARY_02)
- TD02M05 Start Time (corresponds to COLB_02)
- TD02M06 End Time (corresponds to COLC_02)

Outcome Variables

There are eight variables generated by the media coding project: (1) Format (TD97M09, TD02M09); (2) Intended Audience (TD97M10, TD02M10); (3) Character Age (TD97M11, TD02M11); (4) Genre (TD97M12, TD02M12); (5) Comedy (TD97M13, TD02M13); (6) Science Fiction, Fantasy or Supernatural/Paranormal (SF/F/SP) (TD97M14, TD02M14); (7) Curriculum (TD97M15, TD02M15); and (8) Violence (TD97M16, TD02M16) of each television program. Please consult the codebook for a description of these variables for each wave. Please consult the "TV Coding Manual.pdf" file for a description of the coding process.

Merging the Media Data with Data Center Data

The data contained in the CDS Media supplement is a subset of the same-year CDS time diary records for whom time diary data was collected in 1997 and 2002. A record exists in the CDS Media file only if the corresponding diary record had an activity code of 'TV' (activity code 919 for variable COLA in 1997; activity code 9190 for variable COLA_02 in 2002). Records which had an activity code of 919/9190 but which did not have a codeable response for which program was watched will have missing data for the media variables for that record.

To link the CDS Media file to the unaggregated time diary records from the data center, you must sort and merge on the five unique identifier variables listed above (TD97M02-06 for 1997; TD02M02-06 for 2002). Note: there will be some records in the unaggregated CDS Media file which do not initially appear to have an exact match in the unaggregated main time diary file. These are records which the media coders split into multiple entries based on two or more distinctly different TV programs being recorded in the original activity record. When combined or aggregated, these split records will provide the same overall duration of television viewing for that diary. For example:

Example from the unaggregated media file:

TD02M02	TD02M03	TD02M04	TD02M05	TD02M06
122	36	0	34200	36000
122	36	0	36000	37800
122	36	0	37800	38400

Example from the unaggregated data center file:

ER33601	ER33602	DIARY_02	COLB_02	COLC_02
122	36	0	34200	38400

Note that the start-time (TD02M05) of the first record from the media file and the end-time (TD02M06) of the last record from the media file match the start- and end-times of the record from the data center file (COLB_02 and COLC_02), while the remaining start- and end-times for the media file records define time intervals in between. In this example, an activity record of 4200 seconds duration (70 minutes) in the data center file is divided into three activity records of 1800 seconds, 1800 seconds, and 600 seconds duration (30, 30, and 10 minutes) respectively in the media file.

Note that, in some instances, multiple sets of split records exist for some entries, but when collapsed or cross-referenced as outlined above, these records combine for the same overall duration. For example:

Example from the unaggregated media file:

ER30001	ER30002	TD97M04	TD97M05	TD97M06
209	30	0	33300	34200
209	30	0	34200	36000
209	30	0	38700	39600
209	30	0	39600	40500

Example from the unaggregated data center file:

ER30001	ER30002	WDAYWEND	COLB	COLC
209	30	0	33300	36000
209	30	0	38700	40500

A complete listing of all of the split records from both the 1997 and 2002 CDS Media files as well as the corresponding unaggregated data center records is provided for reference purposes:

td97media-ListOfSplitRecordsInMediaFile.pdf
td97media-ListOfSplitRecordsInDataCenterFile.pdf
td02media-ListOfSplitRecordsInMediaFile.pdf
td02media-ListOfSplitRecordsInDataCenterFile.pdf

If CDS Media file records are aggregated to the diary level, and are merged with aggregated records from the data center, this discrepancy goes away.

Note: to create a cross-year file containing both of these files, first you must merge each media file with time diary data from the data center for the corresponding year (1997 with 1997, 2002 with 2002). Then you can merge the two files using the unique Individual File identifiers: ER30001, ER30002.

Additional Documentation

An aggregated listing of television show titles and the corresponding Curriculum code (TD97M15, TD02M15) assigned to them is provided for reference purposes. Note that the text values of the titles are provided as-is and can contain typos, abbreviations, etc. Also note, these values have been recoded to match the coding scheme of the final release data set.

1997CurriculumTitleList.pdf

2002CurriculumTitleList.pdf