User Guide for the BFY Age-3 Data Collection

[September 27, 2023]

Table of Contents

INTRODUCTION:

About this User Guide

This User Guide is for secondary analysts using the Age-3 follow-up data collected by the Baby's First Years project. For full information on the study's design and Baseline data collection, please see the "User Guide for the Baseline Data Collection" available on the study's ICPSR's website. Full information on the Age-1 follow-up is available in the "User Guide for the BFY Age-1 Data collection," while full information on the Age-2 follow-up is available in the "User Guide for the BFY Age-2 Data collection."

We begin this User Guide with a brief description of the Baby's First Years project, followed by additional information about the Age-3 data. The Age-3 data are follow-up data collected approximately 36 months after Baseline data collection, timed to coincide with the focal child's third birthday. We then describe noteworthy features of specific variables important for analysts.

The current Age-3 data deposit includes:

- this User Guide document;
- Age-3 survey instruments in English and Spanish;
- Age-3 data file that contains the survey data;
- STATA .do file that creates the Age-3 data file; and
- ICPSR's electronic and pdf codebooks.

The Age-3 deposit and documentation follow the structures of the Age-1 and Age-2 deposits and documentation but differ from Baseline data documentation. As with Age-1 and Age-2, in this Age-3 deposit, we do not provide a separate codebook as we did at Baseline because ICPSR provides an electronic and pdf codebook. Instead of having a single "Noteworthy Features" section, we distribute this content across several sections that follow the structure and sequence of the survey instrument.

Principal Investigators

Dr. Katherine Magnuson, PhD; Lead PI, social and behavioral science; University of Wisconsin-Madison

Dr. Kimberly Noble, MD, PhD; Lead PI, neuroscience; Teachers College, Columbia University

Dr. Greg Duncan, PhD; University of California, Irvine

Dr. Nathan Fox, PhD; University of Maryland

Dr. Lisa A. Gennetian, PhD; Duke University Sanford School of Public Policy

Dr. Hirokazu Yoshikawa, PhD; New York University

Dr. Sarah Halpern-Meekin, PhD; University of Wisconsin-Madison is the Principal Investigator of the Qualitative Sub-Study

Study Management

Lauren Meyer, Teachers College, Columbia University is the study's National Project Director

Andrea Karsh, University of California, Irvine is the study's Administrative Director

Study Co-Investigators

Sarah Black, PhD; Assistant Professor of Psychology, University of New Orleans, William Fifer, PhD; Sackler Institute for Developmental Psychobiology, Columbia University Medical Center; Michael Georgieff, MD; University of Minnesota; Joseph Isler, PhD; Columbia University Medical Center; Connie Lamm, PhD; University of Arkansas; Dennis Molfese, PhD; University of Nebraska, Lincoln; Victoria Molfese, PhD; University of Nebraska, Lincoln; Jennifer Mize Nelson, PhD; University of Nebraska, Lincoln; Timothy Nelson, PhD; University of Nebraska, Lincoln and Sonya Troller-Renfree, PhD; Teachers College, Columbia University

Scientific Advisory Board

Current (as of October 2022):

Orazio Attanasio, PhD, Professor of Economics, Yale University; Flavio Cunha, PhD, Professor of Economics, Rice University; Kathryn Edin, PhD, Professor of Sociology and Public Affairs, Princeton University; Philip Fisher, PhD, Professor at the Graduate School of Education, Stanford University; Bridget Goosby, PhD, Professor of Sociology, Faculty Affiliate Population Research Center, University of Texas at Austin; Brenda Jones Harden, PhD, Ruth Harris Ottman, Class of '45, Professor of Child and Family Welfare, Columbia School of Social Work; Krista Perreira, PhD, Professor of Social Medicine, UNC School of Medicine, UNC Chapel Hill; Eldar Shafir, PhD, Professor of Psychology, Princeton University; and Catherine S. Tamis-LeMonda, PhD, Professor of Applied Psychology, New York University

Former:

Tom Boyce, MD, Professor of Pediatrics, University of California, San Francisco; Michael López, PhD, Vice President in the Education and Child Development department at NORC, University of Chicago; Bruce McEwen, PhD, Alfred E. Mirsky Professor and Director of Harold and Margaret Milliken Hatch Laboratory of Neuroendocrinology, Rockefeller University

Study Consultants

Alicia Kunin-Batson, PhD, LP, Associate Professor, Department of Pediatrics, University of Minnesota; Charles A. Nelson III, Professor of Pediatrics and Neuroscience, Harvard Medical School; Professor of Education, Harvard University; Richard David Scott Chair in Pediatric Developmental Medicine Research, Boston Children's Hospital; Charles Zeanah, MD; Sellars-Polchow Professor of Psychiatry, Professor of Clinical Pediatrics, Vice-Chair for Child and Adolescent Psychiatry, and Executive Director of the Institute of Infant and Early Childhood Mental Health, Tulane University

Title of the Study

Baby's First Years (abbreviated here as "BFY"; the study is also known as "Household Income and Child Development in the First Three Years of Life")

Funding sources:

Source	PI	Grant number
National Institutes of Health	Duncan, Magnuson, Noble	R01HD087384
Annie E. Casey Foundation	Noble	214.0183

Andrew and Julie Klingenstein Family Fund	Noble	N/A
Arnold Ventures	Noble	21-06173
Arrow Impact	Noble	N/A
Bezos Family Foundation	Noble	N/A
Bill and Melinda Gates Foundation	Noble	OPP1185312
Bill Hammack and Janice Parmelee	Magnuson	N/A
Blue Cross and Blue Shield of Louisiana Foundation	Magnuson	N/A
Brady Education Fund	Duncan	N/A
Chan Zuckerberg Initiative (Silicon Valley Community Foundation)	Noble	2017-177918
Charles and Lynn Schusterman Family Philanthropies	Duncan and Noble	13080
Child Welfare Fund	Noble	13-1624202
Esther A. and Joseph Klingenstein Fund	Noble	N/A
Ford Foundation	Noble	0170-0832
Greater New Orleans Foundation	Magnuson	N/A
Heising-Simons Foundation	Magnuson	542569
Holland Foundation	Noble	542709
Jacobs Foundation	Duncan	102535
JPB Foundation	Noble	1132
J-PAL North America	Duncan	S5341
Lozier Foundation	Noble	N/A
New York City Mayor's Office for Economic Opportunity	Noble	CT1 069 20201415397
Office of Planning, Research, and Evaluation	Duncan, Magnuson, Noble	R01HD087384A
Perigee Fund	Noble	N/A
Robert Wood Johnson Foundation	Noble	75592
Robin Hood	Gennetian	N/A
Russell Sage Foundation	Halpern-Meekin	2104-31401
Sherwood Foundation	Noble	4288
Valhalla Foundation	Noble	N/A
Weitz Family Foundation	Noble	N/A
W.K. Kellogg Foundation	Noble	P3031579
Three anonymous donors	Duncan and Magnuson	N/A

Data Collector

The data were collected by the Survey Research Center (SRC), Institute for Social Research, University of Michigan, Ann Arbor, Michigan, under contracts from the University of California, Irvine, and from Teachers College Columbia University, running from September 2017 through December 2023. Data collection for Baseline data began in May 2018 (exact date suppressed to preserve participant anonymity); data collection for Age-1 began in July 2019, data collection for Age-2 in July 2020, and data collection for Age-3, which is the focus of this user guide, began in July 2021. SRC data collection operations are overseen by: Stephanie Chardoul, Director of Survey Research Operations and Piotr Dworak, Senior Survey Specialist, Survey Research Operations.

PROJECT DESCRIPTION:

Summary Description of the Intervention and its Data Collection Plan

The overall goal of the Baby's First Years study is to assess the causal role played by household income in affecting children's early cognitive, socio-emotional, and brain development. Recent advances in developmental neuroscience suggest that experiences early in life have profound and enduring impacts on the developing brain. Family economic resources shape the nature of many of these experiences, yet the extent to which they affect children's development is unknown. The Baby's First Years project is the first randomized controlled trial to provide estimates of the causal impacts of unconditional cash gifts on the cognitive, socio-emotional, and brain development of infants and toddlers in low-income U.S. families. For a more complete description of the study, see Noble et al. (2021).[1]

Specifically, 1,000 mothers of infants with incomes below the federal poverty line were recruited in 12 birth hospitals in four diverse U.S. communities and began to receive monthly cash gift payments by debit card. Mothers were initially told the gifts would last for the first 40 months of their child's life, but we have secured funding to continue the payments through child-age 6 (i.e., for a total of 76 months). Parents in the high cash gift group (n=400 in the study sample) are receiving a cash gift of $333 per month ($4,000 per year), while parents in the low cash gift group (n=600) are receiving a nominal monthly gift payment of $20 ($240 per year), also for 76 months.

In order to measure the impacts of the unconditional cash gift income on children's development, using measures of EEG, cognitive, language, memory, self-regulation, and socio-emotional development, we are assessing high and low cash gift group differences at age 4. EEG was also gathered at age 1. A small subset of other measures, all through maternal reports, were administered at ages 1, 2 and 3. In order to understand the processes by which child impacts emerge, we are measuring a host of family process measures summarized in our pre-registration chart included in Appendix Table 1.

Our five data collection points are referred to as: "Baseline," "Age-1," "Age-2," "Age-3" and "Age-4."

Preregistration

We preregistered hypotheses for data collected at child ages 1, 2, and 3 with clinicaltrials.gov within a month after recruitment began (May 2018). In September 2018 we preregistered hypotheses with the Registry of Effectiveness Studies. We submitted a revised set of hypotheses in July 2019 – which was just before we began Age-1 impact data collection – to clinicaltrials.gov, the Registry of Effectiveness Studies and the AEA RCT Registry. Additionally, we submitted a revised set of hypotheses in July 2020, June 2021 and July 2022 – which was before we began Age-2, Age-3 and Age-4 data collection, respectively. In December 2022 we made a few additional updates in the registry to indicate changes to the Age-4 protocol that were necessary, such as dropping two measures. A summary of our Age-3 preregistration is given in Appendix Tables 1 and 2.

Our core pre-registered empirical approach for deriving the preregistered impact estimates will use the pooled data across the four sites to compare family processes and child outcomes for the pooled cross-site $333/month group with the $20/month group. Because of the random assignment design, the average of an outcome for the low cash gift group corresponds to the counterfactual state outcome

[1] Noble, K. G., Magnuson, K., Gennetian, L. A., Duncan, G. J., Yoshikawa, H., Fox, N. A., & Halpern-Meekin, S. (2021). Baby's first years: design of a randomized controlled trial of poverty reduction in the United States. *Pediatrics*, *148*(4).

that would have occurred, on average, for the high cash gift group had they not been offered the additional $313/month income supplement. Therefore, differences in outcomes for the high compared with the low cash gift group can be interpreted as estimates of causal treatment effects of the $313/month higher income (regardless of how high cash gift group mothers use the cash gift). These are intent-to-treat estimates. Because randomization took place within each of the four sites, we recommend that all impact regressions include site fixed effects.

UNIVERSE, SAMPLE AND SAMPLING PROCEDURES; SUBSTANTIVE, TEMPORAL, AND GEOGRAPHIC COVERAGE OF THE DATA COLLECTION:

Universe of the BFY Study Sample

Between May 2018 and June 2019, 1,003 mothers with incomes below the poverty threshold in four metropolitan areas in the United States (New York City (abbreviated NY), the greater New Orleans metropolitan area (LA), the greater Omaha metropolitan area (NE), and the Twin Cities (MN)) were recruited in 12 hospitals shortly after giving birth. "Recruited" means that they were deemed eligible based on the information they provided in a screening interview, consented to and participated in a Baseline interview, and were offered and agreed to receive a debit card with a randomly assigned monthly cash gift deposit. All consent forms and data collection instruments for the Baseline, Age-1, Age-2 and Age-3 data collections are available on the study website www.babysfirstyears.com. The Institutional Review Board of Teachers College at Columbia University served as the single IRB of record for most of the study sites. To meet local requirements, stand-alone IRB reviews were conducted in 5 of the 12 recruitment hospitals.

The construction of the sample is detailed in the Baseline CONSORT diagram (Appendix Figure 1) and in the User Guide for the Baseline Data. The final study sample consisted of 1,000 mothers, all of whom were targets of our Age-3 interviewing efforts and will be targets of our data collection efforts at focal child age 4.[2]

Age-3 Follow-up Sample

Between July 2021 and July 2022, we attempted to contact as many of the 1,000 study participants as possible and interview them close to their child's third birthdays. Given COVID conditions, all interviews were conducted over the phone. We completed interviews with 922 participants. Our performance on the Age-1, Age-2, and Age-3 data collection is summarized in the Age-3 CONSORT diagram in Appendix Figure 1.

Across the entire year, 10 mothers refused to be interviewed and 60 mothers were either not found or were unavailable to be interviewed. A small number of mothers were ineligible for the Age-3 follow-up due to the mother's death (3). Adjusting the denominator for ineligibility, our response rate for the Age-3 data collection was 93%.

Participants were pre-paid $50 before the Age-3 interview and received an additional payment of $50 at the end of it. Some mothers were eligible for an additional payment as an incentive to complete the interview.

[2] Target samples for follow-up waves of data collection may vary if participants are excluded from the study. These cases are detailed in CONSORT diagrams. They could include cases of mothers or children who died or mothers who decided to be excluded from the study and stop receiving the cash gift.

Age-3 Sample Group Equivalence

As we did when collecting Baseline, Age-1 and Age-2 data, we sought to create equivalent high and low cash gift groups when we collected the Age-3 data. At the end of the Age-3 data collection (n=922), the completion-rate gap was 4.7 percentage points, but the fact that the completion rate in both gift groups was very high – 90.3% and 95.0% for the low and high cash gift groups respectively – leaves relatively little room for differential nonresponse bias. We investigated whether the low and high cash gift group members that were successfully interviewed as part of the Age-3 survey were similar on the same set of baseline measures used to assess baseline equivalence. Appendix Table 3 shows that of the 26 individual tests, three were statistically significant at the $p<.05$ level. The p-value of a joint test of orthogonality across all baseline predictors is 0.195, which means that we cannot reject the null hypothesis of no group differences. Just as the high and low cash gift groups were very similar at baseline, the two groups appear to be very similar at Age-3.

AGE-3 DATA AND DOCUMENTS

Age-3 Files

This *User Guide* provides basic documentation for the Age-3 data collection. Relevant files on deposit are:

- Age-3 data, STATA .dta file: *BFY_age3clean_public.dta*
- STATA script file (.do file) that cleans the Age-3 data: *BFY_age3cleanpublic.do*
- Age-3 survey instrument, .pdf file: *BFY_Age 3 Instrument.pdf*

Other data collection instruments:

- The Baseline, Age-1, and Age-2 survey instruments can be found at the ICPSR website (https://www.icpsr.umich.edu/web/DSDR/studies/37871/datadocumentation#).
- All data collection instruments can be found on the study website (www.babysfirstyears.com). This includes the Baseline screening instrument, consent forms and the qualitative sub-study Interview Guide.

The Age-3 survey instrument was administered by telephone, with responses recorded by the interviewer on a laptop computer. All prompts used for items are described underneath the item in the Age-3 survey instrument.

Users should note that some sections of the STATA script cannot be run by analysts because they involve personally identifiable information that cannot be made publicly available. The purpose of releasing the code is to provide as much transparency, clarity, and reproducibility as we can.

Version Dates

There were minor changes in the Age-3 survey throughout its administration. These are reflected in the variable `versiondatea3`, which contains five unique values corresponding to five versions of the Age-3 instrument. The minor differences in the versions of the Age-3 instrument are described in Appendix Table 4, though some dates may not correspond exactly.

Weighting

There was no oversampling of population subgroups and assignment to cash gift group was predetermined by an algorithm embedded in our computer-assisted interview software. Thus, no weighting is necessary to obtain Intent-To-Treat estimates for births to low-income women in the study's twelve hospitals distributed across four sites. The timing of interviewer shifts in the 12 hospitals was based on administrative convenience rather than any attempt to obtain a formal population sample from the hospitals. The random nature of births across interviewer shifts contributes to the population diversity of the sample, although the attained sample is not formally representative of any well-defined population.

Unit(s) of Analysis/Observation

The unit of analysis may be the mother, the focal child, the mother/child dyad, or the household, depending on the analysis and the variables.

Frequency Distributions and Weighting

Unweighted univariate frequency distributions are provided for all variables in the ICPSR electronic codebook.

GENERAL AGE-3 DATA CONVENTIONS, DECISIONS, AND PROCESSING

All respondents were asked all questions, unless the questions did not apply to their situation (e.g., mothers who reported that they did not know the identity of the father of their baby were not asked about the father or the father-mother relationship). As with Baseline, Age-1, and Age-2 data, variables in the Age-3 data file are of two types – **raw** and **generated**.

The first type of variables is considered **raw** because these variables are direct outputs from the Age-3 electronic survey programmed by SRC, the partner organization directing collecting data. The raw data are otherwise unprocessed, except for two minor adjustments: suffix "a3" is added to the end of the variable *name* in the dataset (described in more detail in the "Item Naming Conventions" section below) and prefix "[raw]" is added at the start of the variable *label*. The values of these variables follow the conventions in our previous data. For example, SRC has a long-standing practice of coding yes/no responses as yes=1 and no=5. Those 1/5 values are used in the first, "raw" portion of the data file. Exact question wording can be found in the Survey Instrument, with the variable name listed under each question or checkpoint. The order of the variables in the dataset generally follows the order in which the questions were asked in the instrument.

The second – "**generated**" – type of variables are what we generated using the raw data. In addition to simple recoding of values (e.g., yes/no responses are recoded to yes=1 and no=0), we generated pre-registered variables and additional variables that require careful understanding and quality checks of the raw data. We provide important details on the data generation process in the sections below. Finally, some variables serve as indicators for types of open-ended responses in the dataset but do not appear in the questionnaire.

Personally Identifiable Information (PII)

Personally identifiable information (PII; e.g., date of birth) or potentially PII (e.g., child development measure items specific to child age in months) are protected under the Health Insurance Portability and Accountability Act (HIPAA). We refer to HIPAA protected information as PII. We collect PII with the survey, so we have excluded these items in the data file that we deposit to ICPSR. In order to

protect PII, these variables have either been removed or converted into a dummy variable that indicate that the mother provided a response. Some of these variables may become available in the future under more restrictive terms. However, as some of these variables can be essential for analysts, in some cases, we generated new variables that partially or completely mask the sensitive information. These variables are HIPAA compliant and useful for analysis (see Table 1). Some of these variables are described in the table below.

Table 1. Masked Personal Information in Age-3 Public Release

Sensitive information	Variable Name(s)	Description
Child's age at interview	`cagea3`	Child's age at the time of the Age-3 interview recoded to a binary indicator for whether the child was at least three-years-old.
Interviewer ID	`interviewera3`	Randomly generated interviewer identification number. (Not linked across waves)

Missing Data

We use the following coding conventions for missing data for both the raw and generated variables:

.d – don't know

.r – refused

.i – index/scale assigned missing because too many items were missing

. – valid skip

Index/Scale Missing Data Conventions

Generally, a generated scale or index has a value of ".i" if at least half of the individual items are missing or if there are less than 3 non-missing items. If all the items are missing, the generated variable has a value of ". ".

Item Naming Conventions

- Raw variables in the data file generally match the variable names in the survey instrument, which are listed underneath each survey item in the instrument file.[3]
- Raw variables use both "mother" and "mom" interchangeably in the variable name.
- Raw variables have a "[raw]" in the variable label.
- For generated variables, we added one of the following prefixes to the variable name:

  m – Mother item

[3] The variable names in the survey instrument exclude the two characters at the end of the variable name that indicate which data collection wave (e.g., a3 for Age-3).

c – Child item

hh – Household item

d – Father item

- Generated variables use "mother" and "mom" interchangeably in variable names and use "father" and "dad" interchangeably as well.

Analysts are advised to take as many of their variables as possible from our set of "generated" variables because they tend to be cleaner and easier to use.

- When we recoded or reverse coded raw items (often part of a composite), we added "- recoded" (or "- reverse coded") at the end of the *variable label*. We also update the *value label* accordingly. When recoding dichotomous items, we followed the convention of a dummy variable (i.e., 1= yes, 0= no) and did not update the *variable label*.

- See the STATA script *BFY_age3cleanpublic.do* for all recoding decisions.

- The last two characters of all variables indicate the data collection wave. These suffixes were added to all variables. They are:

  a0 – Child Age-0 (i.e., gathered at Baseline shortly after the birth of the child; already available on the ICPSR website)

  a1 – Child Age-1 (gathered around child age 1; available on the ICPSR website)

  a2 – Child Age-2 (gathered around child age 2; available on the ICPSR website)

  a3 – Child Age-3 (gathered around child age 3; this data deposit)

Treatment of Outliers

In general, we did not change values of variables that appear implausible but not impossible. In the rare occasions when we adjusted outliers, we generated new variables and documented the decisions (see survey specific sections below) so that a secondary analyst will always have both the unadjusted and adjusted variables.

Analysts should check for extreme values on key variables to ensure that they do not drive model estimates.

Check All That Apply Questions

Some sections of the survey include a "check all that apply" type of question. Responses to these questions were stored using a particular method. The survey program created as many placeholder variables as there were options to check, then stored mother's *n* checked responses into the first *n* placeholder variables. For example, mothers were asked to check which social service programs they participated in and could select up to 5 options. The survey program creates placeholder values for each

possible program (e.g., servicessupport_s_1a3, servicessupport_s_2a3, etc.). The first placeholder variable contains a value that corresponds to the first social service program the mother indicates that she receives, and the second contains the value that corresponds to the second, and so on. If the mother responds with only one social service program, there will be only one placeholder value. This logic is used for all "check all that apply" type of questions.

AGE-3 SAMPLE VARIABLES

In addition to the survey data, the dataset includes administrative variables created by SRC that describe the Age-3 sample.

- Anyiwdonea3 indicates that the mother completed the Age-3 survey and is the indicator of the final Age-3 sample.
- treat indicates treatment group status (1 = high cash gift group; 0 = low cash gift group).
- site is a categorical variable indicating the 4 sites.
- Iwstartdatema3 has the date when the survey application was first opened (i.e., the day the survey started) and iwenddatema3 has the date when the survey application was last opened (i.e., the day the survey ended).
- cagea3 is the masked age of the child at the time of the Age-3 interview. It indicates whether the child was at least 36 months.
- finaltaskrulea3 is an indicator of the participant's final case status for Age-3, including reason for non-completion for the 78 mothers who did not complete an Age-3 survey. The variable finaltaskrulea3 was constructed by SRC to determine each mother's eligibility for participation and was based on the mother's responses in Section A of the instrument (see details in later sections) and information collected in the process of trying to reach participants. The initial values in the raw data for finaltaskrulea3 included "AcceptedComplete," "AcceptedPartial," "ChildDeceased," "MotherDeceased," "NotAvailable" and "Refusal." The categories allow analysts to crosswalk case status with the Consort Diagram in Appendix Figure 1.

IMPORTANT DATA DETAILS BY SURVEY INSTRUMENT SECTION

What follows are notes about raw and generated variables in the data, organized by survey section as seen in the instrument file. For ease of use with the Age-3 survey instrument, we follow the organization in the instrument, which does not always follow the alphabetical sequence of the section labels.

We recommend that users read both these notes and the instrument for a complete description of the survey data. We also recommend that users review the Stata script for even more precise details. We assume users will refer to the pre-registration tables in Appendix Tables 1 and 2 of this user guide for details about how pre-registered outcome measures were constructed and what the source instruments were, so we do not re-iterate the same details in this section.

We also refer users to Appendix Table 5, which lists pre-registered measures and summarizes changes in measures across waves of data collection. In the Age-2 User Guide, we often identified these

changes in the survey section summary. To ease use of comparison across all three publicly-released waves, we intend this table to assist in quickly identifying differences.

Below is a list of survey sections:

Age-3 Survey Sections

Section A: Intro, Voluntary Statement
Section C: Household Roster
Section XX: Marriage Questions
Section D4: AudioCASI (including items from Section D: Romantic Relationships; Section O: Discipline Strategies; Section Q: Maternal Health; Section X: Everyday Discrimination Scale)[4]
Section E: Residential History, Housing Quality
Section F: Neighborhood (including subsections and Happiness and Life Satisfaction)
Section R: Parent Child Activities
Section Mask: Masking Questions
Section G: Childcare
Section H: Child Health
Section I: Child Development
Section M1: Maternal Health
Section K: Income, Receipt of Public Program Benefits, Debt, and Net Worth
Section L: Expenditures and Economic Stress
Section M2: Maternal Health (same questions as in M1; randomized to be administered after questions on economic stress for half of mothers)
Section Q: Maternal Health
Section V: Closing

[4] Items considered sensitive were grouped together in the AudioCASI portion of the interview, which is Section D4 in the instrument. Because sensitive items may have originated in various survey sections, Section D4 contains items related to Section D (romantic relationships), Section O (discipline strategies), and Section Q (maternal health). In the User Guide, we combine all AudoCASI items into Section D4, following the organization of the instrument.

Section-by-Section Variable Details

Section A: Intro, Voluntary Statement

- **About this Section:** This section determined which version of the survey instrument, if any, would be administered to the mother. The full version of the survey is "Instrument A" and the shortened version that excluded child-focused items is "Instrument B."
    - Instrument A was administered to the mother if the child was well (i.e., not hospitalized) and lived with the mother at the time of the Age-3 interview. Instrument B was administered to the mother if the child had an out-of-home placement (i.e., in foster care). Instrument B was administered to just two participants at Age-3 (see variable `instselecteda3`). Please refer to the survey instrument to see which items were excluded from Instrument B.
    - The survey instrument was not administered if the mother or child passed away, the child lived with adoptive parents, the child was hospitalized or otherwise not well, or if the mother was institutionalized/incarcerated. Analysts should consult the consort diagram (Appendix Figure 1) and the variable `finaltaskrulea3` as described in section AGE-3 SAMPLE VARIABLES of the User Guide for more information.

Section C: Household Roster

- **About this Section:** The household roster section of the Age-3 survey asked the mother to provide basic information about each household member that she had named during the previous year's Age-2 survey, and about each new member that was added to her household since Age-2 that currently lives with her at Age-3. Basic information about each member included name, gender, date of birth, relationship to the mother, whether they contribute to household income, and, when applicable, whether they still currently live in that household. In this section of the user guide, we describe the structure and processing of raw household data and our generated household variables. Personally identifiable information (e.g., name or date of birth) from the roster have been removed from the publicly released dataset.
- **Raw Variables:** Age-3 Survey Instrument's Section C carefully documents what household roster questions were asked, in which order, and with what skip logic. As such, we refer the users to the instrument documentation for complete details of each survey item. However, survey programming of the household roster section was complex, so the raw data exported from the survey program requires further explanations for users interested in using the raw data or in examining our cleaning procedures. Users who want to focus on understanding and using our generated household variables can skip to the subsection "Generated Variables" below.

  There are three key phases of the household roster survey programming relevant to processing the raw data: (1) preloading of a subset of Age-2 roster information (the Age-3 survey asks questions about previously named household members from Age-2); (2) administration of Age-3 rostering (in which the Age-3 survey questions are asked of old and new household members); and (3) storing of the pre-loaded and newly collected roster information into variables with some variable naming conventions. Each phase presented some complications that required our attention for cleaning, which we describe in detail. We begin by describing the naming conventions of the raw variables and how the Age-3 roster information were stored into variables by the survey program.
    - ***Distinguishing variables for different household members*:** The survey program repeated the household roster question sequence for each old and new member (apart from the mother and child) and assigned each member a roster position which was used in the variable name. For example, the first person on the household roster would be assigned position 1, and so numeric "1" would be used at the end of each household variable name (e.g., `hhmemname_1a3` is short for household member name of the person in position 1, `hhmemgender_1a3` refers to the gender of the person in position 1, and so forth). The survey program at Age-3 used positions 1-25 separately for adults and for children such that there is position 1 for adults and position 1 for children. Not all positions get used, and these positions are not fixed across rosters (explained below).

- ***Distinguishing variables from different rosters*:** Table 6 presents a crosswalk of the pre-loaded Age-2 household variables and the newly collected Age-3 household variables for each piece of the household roster information. To distinguish between preloaded Age-2 survey variables, Age-3 survey adult variables, and Age-3 child variables, the program sometimes used a different "stem" in the variable name and/or a different range of roster position numbers (e.g., 26-50 instead of 1-25), or what we will refer to as the position "suffix." For example, to distinguish the household member relationship to mother variables between the preloaded Age-2 roster (adult and child) and Age-3 adult roster, the program used position suffix 1-25 (e.g., `hhmemrel_1a3`) and 26-50 (e.g., `hhmemrel_26a3`) respectively; to distinguish between member name in the Age-3 survey adult roster and child roster, variable stems `hhmemrel` and `childhhmemrel` were used, respectively. Different rules were applied to distinguish the 3 rosters, depending on the roster information. See Table 6 for a detailed display of the rules. For example, while adult and child rosters for the Age-3 survey have a different variable stem for household member name (hhmem and hhchild, respectively), they share variable stems for member month of birth. Also, you will notice that only a single roster of adults and children was preloaded from the Age-2 survey. This is because the adult and child rosters were merged and forced to use the same variable naming conventions during the pre-loading process. This merging has important implications for understanding roster positions. As an important side note, the name variables had been masked even prior to most data cleaning procedures and was made available only to a select staff in secure servers. The variable itself is made available instead with a binary indicator for whether the surveyed mother provided a response.
- ***Relative roster positions*:** As can be seen from Table 6, some Age-3 variables use position suffixes that range from 26-50. The relative positions of 26-50 map onto positions 1-25 such that position 26 corresponds to position 1, position 27 to 2, 28 to 3, and so forth. For example, `childhhmemname_1a3` and `dob_mo_26a3` are variables that belong to the same member.

  Variable stems `hhothadult` and `hhothchild` range only from 1-25. These variables refer to the question where the survey administrators asked the participating mother if there are other adults or children in the household who were unnamed in the Age-2 roster. The first time this question is asked, the information is stored in `otheradultsinhh` or `otherchildinhh`, which do not have a position suffix. These are dichotomous variables indicating if there are (Yes-1) or are not (No-0) any new adults or children in the household.

  When Age-2 survey roster data were pre-loaded for the Age-3 survey, the adult and child rosters were merged into a single roster which used the same variable naming conventions. The child roster was appended to follow the adult roster. For example, if there were 3 adult members and 2 child members in the household at Age-2 and there were no changes to the household by the Age-3 survey, the preloading process placed the 3 adults into positions 1-3 and the 2 children into positions 4-5 of the preloaded Age-2 roster. Except for a few special cases, these relative positions in the preloaded Age-2 roster are retained through the Age-3 survey adult and child rosters. See Table 7.a for a stylized example. In the presented example, Age-2 child roster positions 1-3 remain empty because the children's starting position was 4.
- ***Tracking existing members that leave and new household members*:** During the Age-3 survey administration, for each named member in the Age-2 household roster the survey program asked the mother whether that member was still in the household or if the member had left the household. At the end of reviewing the Age-2 household roster, the survey program asked whether there were other adults or children in the household living with her now at Age-3 who did not live with her at Age-2. If there were other members, the survey proceeded with the sequence of household roster questions and then repeated until the mother confirmed that there were no other adults or children in the household. Table 7.b shows a stylized example of how the information of leavers and new members were stored. When an existing member from the preloaded Age-2 survey was reported as having left the household by the Age-3 survey, that member lost his or her position. For example, if Adult C in position 3 of the Age-2 roster left the household and a new adult, Adult D, entered the household,

position 3 of the adult roster was not reserved for this adult that left and Adult D was slotted into position 3 of the adult roster. This procedure of handling leaving and entering household members was identical for children.

- ***Survey programming that affected raw variable creation*:** There were a couple of survey programming decisions that affected the preloading of the Age-2 roster and implementation of the Age-3 survey rostering based on the nature of the Age-2 roster information. First, older children who became legal adults (i.e., 18 years old) by the Age-3 survey data collection were part of the preloaded Age-2 roster but placed into the Age-3 adult roster, as opposed to retaining their position in the Age-3 child roster. In the following sections, we describe how we dealt with these programming decisions to get to an accurate count of household members.
- ***Counting*:** The immediate objective of understanding and processing the raw variables was to create various household roster counts, such as the number of household members, adults, or children: at Age-2, at Age-2 that stay through Age-3, at Age-2 that leave before Age-3, at Age-3, and at Age-3 that are new to the household since Age-2. There are other variants of household roster counts (e.g., number of members contributing income) which were generated, and they all rely on the same cleaning procedures and methods described in the following sections.
- ***Basic method of counting*:** The basic strategy of counting was two-fold. First, we created "individual member flags"—a 0/1 binary variable for each roster position—to indicate whether the given member fits the criteria of interest (e.g., is an Age-2 member who stayed through Age-3). This created a set of 25 flags, one for each roster position. Second, we took the sum of the 25 flags to generate a subtotal count of interest. Typically, we created counts for adults and children separately, then summed across the two counts to get the total household counts.
    - *Age-2*: Starting with the preloaded Age-2 roster, we created individual member flags, indicating whether a member was part of the Age-2 roster. The rule was to assign a 1 if there was any non-missing value in the household roster variables (e.g., name, gender, relationship to mother, month of birth, etc.) in the given position number. We took the sum of the flags to directly count the total number of household members (i.e., adults and children combined) at Age-2. We were not able to directly count adults and children separately because the two rosters were merged during the preloading process. However, we were able to derive this number, because we could count how many of the household members from Age-2 stayed through the Age-3 survey, or left before the Age-3 survey, separately for adults and children.
    - *Age-2 that stayed through Age-3*: We created individual member flags, indicating whether a member was part of the Age-2 roster and stayed through Age-3. The rule was to assign a 1 if the mother reported that the member was still living with her. We summed across the flags to separately count the number of children and adults that stayed.
    - *Age-2 that left before Age-3*: We created individual member flags, indicating whether a member was part of the Age-2 roster and left before Age-3. The rule was to assign a 1 if the mother reported that the member was NOT still living with her. We summed across the flags and then separately counted the number of children and adults.
    - *Age-3*: We created individual member flags, indicating whether a member was part of the Age-3 roster using a similar method as the Age-2 roster with two modifications to the rule. The rules were to assign a 1 if there was any non-missing value in a selection of the Age-3 household roster variables (i.e., gender, relationship to mother, contributes income, and age; for the child roster, whether employed) in the given position number; and did NOT report having left the household since Age-2. The reason for looking at select Age-3 household roster variables was because some variables were non-missing for household members who were not present at Age-3. The name of the member at Age-2 was loaded into the Age-3 member name variable. When date of birth information of the member was missing at Age-2, the survey program asked the mother to provide this information during the Age-3 survey even if that member was no longer living there. The reason for conditioning on whether the member was NOT reported to have left the household since Age-2 is that some household

roster questions were asked about the biological father of the focal child regardless of whether he was still present in the household at the Age-3 survey. We summed across the flags to separately count the number of children and adults that stayed.

- *Age-3 that is new*: We created individual member flags indicating whether a member that was part of the Age-3 roster was a new member. The rule was to assign a 1 if the mother reported that there was someone else living in the household, which indicates that this member in the corresponding position was a new member unnamed in the Age-2 roster. We summed across the flags that separately counted the number of children and adults.

- ***Checking internal consistency*:** We used two types of methods to check the reliability of our counts. First, as described above, we have direct counts from the snapshot of the household roster at the time of the Age-2 survey and Age-3 survey and the household flow counts that track members who leave, stay, and newly enter the household in between the two snapshots. Algebraically, we can use a different combination of flow counts and one of the snapshot counts (e.g., count of Age-2 members) to derive the other snapshot count (e.g., count of Age-3 members) to check how internally consistent our counts are. For example, the number of Age-3 household members can be derived by taking the preloaded Age-2 roster counts, subtracting the number of members that leave, and adding the number of new members that enter the household. Likewise, we can add the number of members who stayed between the Age-2 and Age-3 surveys and the new members in the Age-3 survey. We used different algebraic combinations to check the internal consistency of our household-level counts.

  Second, we checked the reliability of our household flow indicators (i.e., stay, leave, enter) at the household member level by comparing the household roster position information across the two survey snapshots. As explained above (see subsection Relative Roster Position), the relative roster positions were stable and permanent, excluding a few exceptions. For example, an Adult at Position 1 in the Age-2 roster that had been flagged as having "stayed" in the household should reappear in Position 1 of the adult roster in the Age-3 survey (see Table 7.a for a stylized example). Likewise, Adult 1 who had been flagged as having "left" the household should not reappear in the Age-3 survey and Adult 1 roster in the Age-3 survey should remain empty (see Table 7.b for a stylized example). Using a series of several Boolean expressions, we could verify whether flow counts agreed with what the comparison of snapshot records suggested.

  We found disagreements for a small proportion of our records because of three peculiar types of cases that we could systematically adjust for: (1) children who turned 18 years old and became a legal adult by the Age-3 survey, (2) members that were preloaded more than once, and 3) cases where the focal child had passed away.

- ***Systemic adjustments*:** When child members in the roster at Age-2 turned 18 and became legal adults, the survey program treated them differently during survey administration. They were still loaded as a child in the preloaded Age-2 roster but slotted into the adult roster for the Age-3 survey implementation (this is not visible in the stored raw data). Mothers were asked the household "flow" questions about these cases, treating them as adults. Another issue was that in some cases, the same member from Age-2 was preloaded into Age-3 more than once. These cases mostly occurred when two household members shared the same first or full name. We identified these cases by looking for duplicates on every characteristic and removing duplicates from the dataset, after corroborating that they were indeed duplicates and not simply twins. After executing systematic adjustments, we ran the internal reliability checks for all of our counts.

- **Generated Variables:** The above sections describe in detail how the household roster raw data were processed to generate counts and descriptions of the household structure. We anticipate most users will only use the cleaned, generated variables which we describe in this section. Variables containing snapshot counts of adult members, child members, and all members (adult + child) during the Age-2 survey (based on data preloaded into the Age-3 data) are:
  - `hhcount_age2adulta3`

- hhcount_age2childa3
- hhcount_age2alla3

Variables containing snapshot counts of adult members, child members, and all members during the Age-3 survey are:

- hhcount_age3adulta3
- hhcount_age3childa3
- hhcount_age3alla3

Variables containing counts of adult members, child members, and all members who stayed, left, or were new between the Age-2 survey and the Age-3 survey are:

- Stayed between Age-2 and Age-3
  - hhcount_adultstaya3
  - hhcount_childstaya3
  - hhcount_allstaya3
- Left between Age-2 and Age-3
  - hhcount_adultleavea3
  - hhcount_childleavea3
  - hhcount_allleavea3
- New between Age-2 and Age-3
  - hhcount_otheradultnewa3
  - hhcount_otherchildnewa3
  - hhcount_otherallnewa3

There was a special case of household members: those who become legal adult members between the Age-2 and Age-3 survey straddle the two categories (i.e., child and adult). They were treated as "children" and only contributed to "child counts" (i.e., did not contribute to "adult counts") for the Age-2 snapshot count and the count of children who stayed/left between Age-2 and Age-3. However, they were treated as adults for the Age-3 snapshot count. Additionally, the following additional household structure variables were created from the household roster variables:

- lives with no other adults in Age-3 (hhnoadultsa3)
- lives with romantic partner in Age-3 (hhromanticpa3)
- lives with unrelated adults in Age-3 (hhunrelatedadultsa3)
- count of unrelated adults in Age-3 (hhcount_unrelatedadultsa3)

Finally, the month of birth variables (dob_mo) and names have been masked for data confidentiality reasons.

Section XX: Marriage Questions

- **About this Section:** This section on maternal marital status was new to the Age-3 survey. There is some overlap between the variables in Section XX of the Age-3 survey and the variables in Section D3 of the Age-2 survey. Section D3 was not included in the Age-3 survey.
- **Raw Variables:** legallymarried2a3 asks whether the mother is legally married. If mothers said "yes", lengthmarriageya3 asks how long the mother has been legally married (in years). If the mothers said "less than 1 year", lengthmarriagema3 asks how long the mother has been legally married (in months). marriagepartnera3 asks whether the mother is legally married to the biological father.
- **Generated Variables:** We recoded the yes/no raw variables on maternal marital status into generated binary dummy variables, including mmarrieda3 and mmarriedtodada3.

Section D4: AudioCASI (including items from Section D: Romantic Relationships; Section O: Discipline Strategies; Section Q: Maternal Health; and Section X: Everyday Discrimination Scale)

- **About this Section:** This section contains potentially sensitive items pertaining to romantic relationships, child discipline strategies, and maternal health substance use. Had the Age-3 interviews been conducted in-person, this section would have been administered using Audio Computer Assisted Self Interview (ACASI) to provide additional privacy. Because the entire Age-3 wave of interviews occurred over the phone due to the COVID-19 pandemic, none of the mothers' interviews used the ACASI format. Instead, the interviewers administered the survey questions. Prior to beginning this section, the interviewers asked the mothers to not be on speaker phone and to be somewhere private, if possible. More information about the script is available in the instrument file. Although the ACASI measures were not actually gathered using that method in the Age-3 interviews, we continue to refer to them as "ACASI measures."
  Below we describe each of the ACASI subsections in turn. Analysts should note that the romantic relationships subsection was only answered by those mothers who identify as being in a relationship at the time of the survey.
  Romantic Relationships Subsection: This subsection contains one pre-registered outcome: poor relationship quality. Mothers were prompted to think about their relationship with their most recent partner in the past year, and items were asked to all mothers. Only mothers who volunteered that they were not in a relationship in the last year (N = 121) were not asked these items.
  - **Raw Variables:** The relationship quality index is an additive index of 11 items, with each item on a 3-point scale (Often, Sometimes, Never). Note that at Age-1, one of the items, "Has your partner ever threatened to spank or slap your child or children?" (`pviolenta1`), was deemed to be too sensitive to be on a 3-point scale and had to be collapsed into a binary (yes/no) indicator on the survey, and therefore was excluded from the index. However, in later versions, the question was asked on a 3-point scale and therefore was included in the Age-3 relationship quality index.
  - **Generated Variables:**
    - The relationship quality index (`mrelationquality_11itema3`) is an additive index of 11 items that ask how often the participant's partner was fair and willing to compromise, expressed affection or love, insulted or criticized the participant for ideas, made the participant feel down or bad about herself during an argument, encouraged or helped her to do things that were important to her, isolated the participant, hurt her physically, sexually abused her, listened to her, made her feel afraid, or threatened or hurt her children. The positive relationship items were reverse coded (`mpcompromisea3`, `mpaffectiona3`, `mpencouragea3`, and `mplistena3`) such that higher values indicated more positive relationship quality. Note that this index includes 11 items, rather than 10 items as is the case for the Age-1 variable, for reasons described above. A 10-item version (`mrelationqualitya3`) was also created for consistency between waves.
    - The pre-registered binary outcome of a poor-quality relationship (`mrelationquality_da3`) is a dichotomous indicator of current or recent relationship quality, where poor quality is defined as 1 if the mother is in a relationship and has a score of 26 or below on the relationship quality scale and a 0 either if the mother is not in a relationship or is in a relationship and has a relationship quality index score of 27 or above. This variable was generated using the 10-item version for reasons already described.
- **Discipline Strategies Subsection (Section O):** This subsection includes one pre-registered outcome about whether the mother has used spanking as a discipline strategy in the past month. The generated preregistered variable (`hhspanka3`) is a dummy that indicates the use of spanking.
- **Maternal Health—Smoking and Alcohol Use Subsection (Section Q):** This subsection includes four questions regarding each mother's use of cigarettes, alcohol, and opioids in the past year. Two preregistered outcomes are generated from this section: `maclciga3`, created from summing a scale of use in the past year for both alcohol and cigarettes from 0 to 4 (never in the past year, less than once per month, several times per month, several times per week, every day), and `mopioida3`, a measure of the frequency of opioid use in the past year using the same scale.
- **Everyday Discrimination Scale Subsection (Section X):** This subsection focused on the mothers' day-to-

day experience of discrimination due to race and/or ethnicity. These questions were not included at Age-2 interviews. Each mother's perceptions of racial/ethnic discrimination were assessed by adapting the 6-item Everyday Discrimination Scale (Williams et al., 1997)[5] used by Halim, Moy and Yoshikawa (2017)[6] in a study of perceived ethnic and language-based discrimination among Latina immigrant women. The scale used in this subsection includes an additional item on participants' perceptions of being treated poorly by medical providers and expanded the response options from three to six for all but one item (see description below).

- **Raw variables:** This subsection includes 7 items in total. The first six items are on a 6-point scale measuring the frequency of discrimination experiences (Almost every day, At least once a week, A few times a month, A few times a year, Less than once a year, Never), and the additional seventh item (`discjoba3`) is on a 4-point scale (Often, Sometimes, Rarely, Never). Three items related to the mother's work and job search experience at the time of the survey included a not applicable option coded as 7 for `disccoworkera3` and `discsupervisora3`, and as 5 for `discjoba3`.
- **Generated variables:** We generated versions of each raw item where the response options were reverse coded so that higher values reflected a higher frequency of experiencing discrimination. Non-response due to a not applicable (N/A) item was recoded as missing ("`.`"). Recoded items start with an "m" in their variable name, and it is noted in the variable label. An everyday discrimination total score (`meverydaydiscriminationa3`) was generated as an additive index of the six 6-point scale items (`mdiscfreqa3 mdiscservicea3 mdisccoworkera3 mdiscsupervisora3 mdiscpolicea3 mdiscdoctora3`), with higher values indicating more discrimination. Note that the item `mdiscjoba3` was excluded from the computation of the total scale score because this item did not reflect discrimination that could happen on a daily basis. If three or more items were missing for a participant's response, then the total score was coded as missing.

Section E: Residential History, Housing Quality

- **About this Section:** This section of the survey includes items for two pre-registered outcomes: (1) Excessive Residential Mobility and (2) Homelessness.
- **Generated Variables:**
  - The pre-registered variable `hhexcessivemovea3` is a binary indicator for whether the mother moved three or more times in the past 12 months.
  - We previously pre-registered homelessness to be an additive index of two items (`rhomelessa3`: ever homeless in the past 12 months, and `rgroupsheltera3`: ever in a group shelter in the past 12 months). We have since updated this outcome to be a binary indicator of whether mothers were ever homeless or in a group shelter in the past 12 months (`hhhomelessorsheltera3`). The pre-registration table included in Appendix Table 1 reflects this change. Analysts wanting to use the additive index can create it with variables `rhomelessa3` and `rgroupsheltera3`.

Section F: Neighborhood

- **About this Section:** This section asks mothers about neighborhood safety and includes one pre-registered outcome: Index of Perceptions of Neighborhood Safety.
- **Raw Variables:** Mothers rated neighborhood safety on a four-point scale during the day (`neighborhooddaya3`) and at night (`neighborhoodnighta3`).
- **Generated Variables:** The pre-registered variable `hhneighbsafetya3` is a 2-item additive index of

[5] Williams, D.R., Yu, Y., Jackson, J.S., et al. (1997) Racial differences in physical and mental health: Socio-economic status, stress and discrimination. *Journal of Health Psychology* 2(3): 335–351. https://doi.org/10.1177/135910539700200305.

[6] Halim, M. L., Moy, K. H., & Yoshikawa, H. (2017). Perceived ethnic and language-based discrimination and Latina immigrant women's health. *Journal of Health Psychology*, *22*(1), 68–78. https://doi.org/10.1177/1359105315595121

perceptions of neighborhood safety during the day and at night, with higher values indicating more safety. We also generated two dummy variables to indicate whether mothers felt "safe or very safe" during the day and at night (hhneighborhoodday_da3 and hhneighborhoodnight_da3, respectively).

Section R: Parent-Child Activities

- **About this Section:** This section asks mothers about activities they perform with the child and includes one pre-registered outcome: Self-Report of Parent-Child Activities.
- **Raw Variables:** Parent-child activities were measured using the Parent Child Activities Index, which is a self-report of how often parents engage in a series of activities with their child. The four items were asked on a four-point scale: (1) rarely or not at all, (2) a few times a month, (3) a few times a week, (4) every day. Analysts should note that there is an additional item added from Age-1, which used four items (pretendplaya3). However, one item was dropped from Age-2 (playgroupa2) due to Covid, so the total number of items is four.
- **Generated Variables:** The pre-registered parent child activities index (mparentchildacta3) was generated using an additive index of five items, with higher values indicating higher frequency of parent-child activities.

Section Mask: Covid-19 Pandemic Mask Questions

- **About this Section:** This section asks mothers about their use of masks, along with their child's use of masks, both indoors and outdoors over the past year. This section was added later (see Appendix Table 4 for date). There are no variables generated from this section.

Section G: Childcare, Mother's Training/Education, and Employment

- **About this Section:** This section of the survey asked about various "life events" at the time of the Age-3 interview. Life events recorded at the time of the Age-3 interview included childcare, employment, and mother's education and training attainment. This section includes three pre-registered outcomes: (1) cost of paid childcare; (2) use of center-based care, and (3) mother's education and training attainment.
- **Raw Variables:** When responding to the questions, mothers were asked to think of the typical 7-day week in the last month, which we will refer to as the "time of the Age-3 interview."
- **Generated Variables:** We generated several variables for internal purposes but leave them for secondary analysts in case they are useful (e.g., hours worked across all jobs). For employment, we operationalized full-time as working 35 hours or more per week across all jobs (see mworkparttimea3 and mworkfulltimea3). We refer analysts to the Stata do-file for the full details of non-preregistered generated variables. The pre-registered variables are as follows:
    - Mothers reported the child had spent 5 or more hours in a childcare or day care center last week (hhdclastweeka3).
    - Mothers reported how much money they spent on out-of-pocket childcare arrangements in the last week (hhpaidcccosta3) for the focal child. We asked the mother to confirm whether that amount was for the study focal child or for other children. When the mother reported that the amount was for other children as well, we created a new variable that adjusted the cost of childcare by the number of children involved (hhadjpaidcccosta3).
    - The variable medjobtraina3 indicates whether the mother participated in education (everattenda3) or job training (jobtraina3) activities in the last 12 months.

Section H: Child Health

- **About this Section:** There are two subsections (Child Sleep and Child Health) with two pre-registered outcomes (1) Sleep Disturbance and (2) Child Health
- **Child Sleep Subsection**

- **Raw Variables:** The PROMIS sleep disturbance index in an additive index of four items on a five-point scale (1: Never, 2: Almost Never, 3: Sometimes, 4: Almost Always, 5: Always). By mistake only three of the questions were asked. Mothers responded to the three items that asked in the past 7 days how often the focal child had difficulty falling asleep, slept through the night, had a problem with his/her sleep, and had trouble sleeping. Sleeping through the night was reverse-coded.
- **Generated Variables:** The pre-registered outcome `cPROMISa3` is an additive index of 4-items, with higher scores indicating more sleep problems. The 4-item scale `cPROMISa3` is the originally intended pre-registered outcome that matches the source scale, but by mistake, we left off one of the items in the pre-registration document and effectively pre-registered a 3-item scale. Now, at Age-3, we mistakenly only asked 3 of the 4 items (`csleptthrua3`, `csleepproba3`, and `csleeptroub_b_2a3`). These 3 items do not overlap with the 3 items from Age-2. It was possible at Age-2 for secondary analysts to create the 3-item scale by excluding `csleeptroub_b_2a2`. However, that is no longer possible, since `csleeptroub_b_1a3` is missing. Removing `csleeptroub_b_2a3` will simply create a 2-item scale.

- **Child Health Subsection**
  - **Raw Variables:** This section asks questions about the child's overall health, and we use six of them to generate a pre-registered additive index of child overall health. The six raw variables that are used to create the additive index of child poor health include: `chealtha3`, `cdocsicka3`, `cdochurta3`, `csickera3`, `certimesa3`, and `cdisabilitya3`. In addition to these six items, there are additional raw variables that include further information on child diagnoses, medications, vaccination status, missed medical care, and whether the focal child had received any early intervention services (e.g., speech therapy, physical therapy, or occupational therapy).
  - **Generated Variables:** The pre-registered outcome `csickhealtha3` is an additive index of the 6-items that represents a child's overall poor health, with higher scores indicating poorer health. These original six items use various scales. Please see the Stata do-file for details. We pre-registered these items as an additive index subject to factor analysis. Our analysis suggests that this index does <u>not</u> conform to a one or two factor structure, and we suggest that users consider this in their work with these data.

Section I: Child Development

- **Parents' Evaluation of Developmental Status (PEDS):** This section includes the pre-registered measure of maternal concern for child language development, socioemotional development, and general development: the Parents' Evaluation of Developmental Status (PEDS). The PEDS has eight questions that address maternal concern for the development of various developmentally relevant skills (i.e., expressive and receptive language, fine and gross language skills, behavioral problems, social emotional development, independence, and learning) and two open ended questions that ask mothers to share any concerns about their child's development.
  - **Raw Variables:** For the eight skill-specific questions, mothers reported whether they had concern for their child's development of each skill. Each item was scored as 0 = "No", 1 = "A little", or 2 = "Yes." For the two open-ended questions, mothers' responses were reviewed and coded. When mothers indicated a concern about any of the eight primary areas of development in their open-ended responses, the specific item was scored "1" if the mother had not already indicated concern for this item when responding to the respective question. As per PEDS scoring instructions, responses to these questions were also coded for fit within the categories of "Global Cognition" and "Other/Health."
  - **Generated Variables:** To generate the PEDS language concerns score, socioemotional concerns score, general concerns score, and predictive concerns score, we recoded the items in the following ways:
    - The eight developmental concern items were recoded to be on a scale of 0 to 1 with 0= "No"

and 1= “A little” or “Yes.”

- *PEDS Language Development Concerns Score* (`clanguagedelaya3`): We summed the responses to the two items pertaining to child expressive and receptive language development. Participants had to answer both questions to be assigned a total score.
- *PEDS Socioemotional Development Concerns Score* (`cbehavioralproblemsa3`): We summed the responses to the two items pertaining to child behavioral development and social interaction. Participants had to answer both questions to be assigned a total score.
- *PEDS Total Concerns Score* (`cpedsindexa3`): We summed the responses to all ten items (the eight concern items and two additional items generated through the open-ended questions regarding global cognition and other/health) to create a total concerns index. Participants had to have a score for at least five items to be assigned a score.
- *PEDS Total Predictive Concerns Score* (`cpredictiveconcernsa3`): Next we summed the total number of predictive concerns that the mother endorsed according to the items that the PEDS has deemed as predictive of developmental delay at age three. These areas included concern surrounding global/cognitive, expressive language, receptive language, gross motor, and other/health concerns. Participants had to have a score for at least two items to be assigned a score.

- **Child Behavior Checklist (CBCL):** This section includes the pre-registered measure of maternally reported child socioemotional development. Four subscales of the CBCL were administered, together comprising 41 items. These subscales were the: Anxiety/Depression (eight items), Aggressive Behavior (19 items), Attention Problems (five items), and Emotionally Reactive (nine items) subscales.
  - **Raw Variables:** We asked 41 items from four subscales from the full CBCL. For each item, mothers reported the extent to which each behavior was a problem for their child in the preceding two months (0 = “not true”, 1 = “somewhat or sometimes true”, 2 = “very true”).
  - **Generated Variables:**
    - Items from each subscale were summed and converted into subscale-specific standardized t-scores according to a CBCL-provided conversion table.
    - *CBCL Total Behavioral Problems Score* (`ccbclindexa3`): The 41 raw items were summed to create a total score comprised of items from all four subscales. This is the pre-registered outcome. Participants had to answer at least 20 items to be assigned a score.
    - *CBCL Anxiety/Depression Subscale Score* (`ccbclanxdepta3`): This score was generated using the standardized t-score for the eight Anxiety/Depression subscale items. Participants had to answer at least four items to be assigned a score.
    - *CBCL Aggressive Behavior Subscale Score* (`ccbclaggbehaviorta3`): This score was generated using the standardized t-score for the 19 Aggressive Behavior subscale items. Participants had to answer at least nine items to be assigned a score.
    - *CBCL Attention Problems Subscale Score* (`ccbclattentionta3`): This score was generated using the standardized t-score for the five Attention Problems subscale items. Participants had to answer at least two items to be assigned a score.
    - *CBCL Emotionally Reactive Subscale Score* (`ccbclemoreactiveta3`): This score was generated using the standardized t-score for the nine Emotionally Reactive subscale items. Participants had to answer at least four items to be assigned a score.

Section M1: Maternal Health

- **About this Section:** This section includes two pre-registered outcomes: (1) global happiness and (2) maternal agency. Placement of this section in the survey administration was randomized for participants. Half of the mothers were asked these items before the income items (Section K), and the other half were asked these items after income questions.

- **Raw Variables:** Because of the randomization in survey order, mother's responses were stored in two sets of variables, one set for each randomization block. For example, the response to the item about maternal happiness is stored in two variables `mhealth1a3` and `mhealth2a3`. In addition, the responses were stored on different values for the different randomization blocks. For example, the three-point response options to happiness (Not happy, Pretty happy, Very happy) were stored in values 1-3 for the first randomization block and stored in values 4-6 in the second randomization block. This is also true of the maternal agency items. We adjust the response options accordingly.
- **Generated Variables:** The pre-registered outcome global happiness (`mhappya3`) is a one-item outcome on a three-point response scale (0: Not too happy; 1: Pretty happy; 2: Very happy). The pre-registered outcome maternal agency (`mHOPEa3`) is an additive index of eight items on a five-point response scale (1: Definitely false; 2: Mostly false; 3: Sometimes true and sometimes false; 4: Mostly true; 5: Definitely true), with higher values indicating more agency. In addition to these two pre-registered variables, we also generated a dummy variable to indicate the mother felt pretty happy or very happy (`mhappyda3`).

Section K: Income, Receipt of Public Program Benefits, Debt and Net Worth

- **About this Section**: This section includes three subsections: (1) income; (2) receipt of public benefits; and (3) debt and net worth.
- **Income Subsection**: To estimate total household income, this subsection asks participants to report five components of income: (1) mother's earned income, (2) spouse's earned income (if living with a spouse or other romantic partner determined by the mother's responses in the household roster section), (3) others' earned income (if living with other adults determined by the mother's responses in the household roster section), (4) government income, and (5) all other income (such as money from any businesses, help from friends or relatives, child support, and any other money income). Income reports correspond to the entire calendar year preceding the year of the interview. If the survey interview was conducted in 2022 or 2021, the values of these components correspond to the annual total earned in 2021 or 2020, respectively.
    - **Raw Variables**: Mothers are first asked to report a dollar value for each component of their income. If they do not provide a dollar value, then the value is estimated through a series of "unfolding questions" that approximate the income component amount. The dollar values for each income component are then stored for all mothers (i.e., those who provided a dollar value and those whose amount was estimated through the unfolding sequence) in the following raw variables: `totearnedincomea3`, `totspouseincomea3`, `totothhhmemincomea3`, `totgovtincomea3`, and `totallotherincomea3`. The five components are automatically added up by the survey program and stored in another raw variable (`combinedincomea3`), and then mothers are asked if this total income value is about right for their household income. If the answer is "no" or "don't know" (in variable `calculatedincomea3`), they are asked to provide a best estimate, including all sources (`estimatedincomea3`). The unfolding questions are coded the same way as at Age-2:
        - If mothers said "yes" to the last unfolding question "is it $45k or more?", they were assigned a value of $50,000.
        - If mothers said that they did not know or they refused to provide a value (or to answer the unfolding questions), they were assigned a value of $2,500.

      Four questions were added to the government income component of the income subsection at Age-3. Mothers were asked whether they received a stimulus payment and how they used the stimulus payment. They were also asked whether they received child tax credit payments and whether anyone in the household filed an income tax return. `receivestimulusa3` asks whether anyone in the household received the most recent $1,400 stimulus payment from the Federal Government. If mothers said "yes," `stimulususea3` asks whether they spent the stimulus payment, saved it, used it to pay off debt, or something else. `childtaxcreditsa3` asks whether anyone in the household received additional money from the government since July 2021 for children. `filetaxesa3` asks

whether anyone in the household filed a federal income tax return for the most recent tax year. None of the four variables are used in household income calculations.

- **Generated Variables**: For all generated income variables, we adjusted amounts to 2019-dollar values, thereby deflating the amounts reported. We generated variables to represent the total for each of the five income components. The generated income variables by component are: `hhmomearneda3`, `hhspouseearneda3`, `hhothersearneda3`, `hhgovtincomea3`, and `hhotherincomea3`. There are two key differences from the versions calculated in the raw variables: (1) we maintain missing values (.d, .r, and .) instead of assigning $2,500 when mothers did not know or did not report a value, and (2) if mothers said it was more than $45,000 (i.e., the last bracket in the unfolding sequence), instead of assigning $50,000, we assigned the median value of individuals who gave an amount above $45,000 in the first place. This is:
  - $79,466.66 for other household members' earnings (based on 1 observation)
  - No mothers reported more than $45,000 for maternal earnings, spouse earnings, government income, or other sources of income.

  We generated alternative variables for two of these components (`hhspouseearnedexpa3` and `hhothersearnedexpa3`) that replace missing values with 0 for mothers who report no spouse or no other household members, respectively. To estimate total household income (`hhrevisedincomea3`), we applied the following rules:
  - If the mother confirmed the "combined income" value as being correct (`calculatedincomea3 == 1`), then the household income is the sum of each component (replacing sums less than $5,000 with $2,500).
  - If the mother did not confirm the "combined income" value and provided a new estimate (`estimatedincomea3`), then the household income (`hhrevisedincomea3`) is the new estimated income that the mother provided.
  - If the mother did not confirm the "combined income" and then did not offer an alternative estimate from the one calculated from her previous responses, then this household income variable is missing.

  We also generated an alternative version of the total household income variable (`hhincomea3`) where we use the combined income value that mothers did not confirm (`calculatedincomea3 == 5`) for the mothers who did not provide an alternative estimate to minimize missing values for these mothers.

  We recoded the yes/no raw variables on stimulus and child tax credit receipt into generated binary dummy variables, including `hhreceivestimulusa3`, `hhchildtaxcreditsa3`, and `hhfiletaxesa3`. Additionally, we generated an alternative variable for stimulus payment use (`hhstimulususeexpa3`) that replaces missing values with 0 for mothers who report no stimulus payment receipt.

- **Receipt of Public Benefits Subsection**: This subsection asks whether the mother receives public benefits and includes the pre-registered outcome social services receipt index.
  - **Raw Variables**: Mothers are asked whether they receive 5 social services (i.e., food stamps/SNAP/EBT, Women, Infants, and Children (WIC), State Unemployment, Medicaid coverage for self, Housing Assistance). Mothers were asked about these social services using a checklist. Please see section "Check all that apply" for details on how raw variables get stored for these types of questions.
  - **Generated Variables**: The pre-registered outcome social services receipt index (`hhsocialservicesa3`) is an additive index summing the 5 social service items.
- **Debt and Net Worth Subsection**: This subsection was new to the Age-3 survey and has three components: (1) savings, (2) debt, and (3) net worth.
  - **Savings**:
    - **Raw Variables**: `savingsa3` asks whether the mother has any money saved up currently. If

mothers said "no," savingsaccta3 asks whether the mother has any type of checking, savings, or bank account. If mothers said "yes," they are first asked to report a dollar value for their savings (savingsamta3). If they do not provide a dollar value, then the value is estimated through a series of "unfolding questions" that approximate the savings amount (savingsamt5000a3, savingsamt1000a3, savingsamt25000a3).

- **Generated Variables**: The generated savings variable is hhsavingsa3. The key difference from savingsa3 is the imputation of missing values using the unfolding questions as follows:
  - If mothers did not know or did not report a value for the savings amount or the unfolding questions, we maintained missing values (.d, .r, and .) .
  - If mothers said the savings amount was less than $1,000, we assigned a value of $500.
  - If mothers said the savings amount was above $1,000 and below $5,000, we assigned a value of $3,000.
  - If mothers said the savings amount was above $5,000 and below $25,000, we assigned a value of $15,000.
  - If mothers said the savings amount was more than $25,000, we assigned the median value of individuals who gave an amount above $25,000 in the first place. This is equal to $61,250 (based on 2 observations). We generated an alternative variable for savings (hhsavingsexpa3) that replaces missing values with 0 for mothers who report no savings.

- **Debt**:
  - **Raw Variables**: debta3 asks whether the mother or any other household member has any debt currently. If mothers said "yes," they are first asked to report a dollar value for their debt (debtamta3). If they do not provide a dollar value, then the value is estimated through a series of "unfolding questions" that approximate the debt amount (debt5000plusa3, debt1000plusa3, debt25000plusa3). Mothers who reported having debt were also asked whether they hold 8 types of debt (credit card or store card balances, student loans, medical debt, pawn shops, pay day loans, child support debt, criminal justice debt, or other debt). Mothers were asked about these debt types using a checklist. Please see section "Check all that apply" for details on how raw variables get stored for these types of questions.
  - **Generated Variables**: The generated debt variable is hhdebta3. The key difference from debta3 is the imputation of missing values using the unfolding questions as follows:
    - If mothers did not know or did not report a value for the debt amount or the unfolding questions, we maintained missing values (.d, .r, and .).
    - If mothers said the debt amount was less than $1,000, we assigned a value of $500.
    - If mothers said the debt amount was above $1,000 and below $5,000, we assigned a value of $3,000.
    - If mothers said the debt amount was above $5,000 and below $25,000, we assigned a value of $15,000.
    - If mothers said the debt amount was more than $25,000, we assigned the median value of individuals who gave an amount above $25,000 in the first place. This is equal to $74,392 (based on 2 observations).

  We recoded the placeholder raw variables on debt types into generated binary dummy variables, including hhccarddebta3, hhstudentdebta3, hhmeddebta3, hhpawndebta3, hhpaydaydebta3, hhcsupportdebta3, hhlegaldebta3, and hhotherdebta3.

- **Net Worth**:
  - **Raw Variables**: networtha3 asks the mother to imagine selling all major possessions,

turning all assets into cash, and paying all outstanding debts. Mothers are then asked whether they would have something left over, break even, or be in debt. If mothers said, "something left over," they are first asked to report a dollar value for the amount left over (`leftovera3`). If they do not provide a dollar value, then the value is estimated through a series of "unfolding questions" that approximate the leftover amount (`leftover5000a3`, `leftover1000a3`, `leftover25000a3`). If mothers said "be in debt," they are first asked to report a dollar value for the amount left over (`beindebta3`). If they do not provide a dollar value, then the value is estimated through a series of "unfolding questions" that approximate the debt amount (`indebt5000a3`, `indebt1000a3`, `indebt25000a3`).

- **Generated Variables**: The generated continuous net worth variable is `hhnetwortha3`. This variable is equal to `leftovera3` or `beindebta3`, with missing values imputed using the unfolding questions as follows:
    - If mothers did not know or did not report a value for the net worth amount or the unfolding questions, we maintained missing values (.d, .r, and .).
    - If mothers said they would have a leftover amount more than $25,000, we assigned a value of $30,000.
    - If mothers said they would have a leftover amount above $5,000 and below $25,000, we assigned a value of $15,000.
    - If mothers said they would have a leftover amount above $1,000 and below $5,000, we assigned a value of $3,000.
    - If mothers said they would have a leftover amount less than $1,000, we assigned a value of $500.
    - If mothers said they would break even, we assigned a value of $0.
    - If mothers said they would have a debt amount less than $1,000, we assigned a value of -$500.
    - If mothers said they would have a debt amount above $1,000 and below $5,000, we assigned a value of -$3,000.
    - If mothers said they would have a debt amount above $5,000 and below $25,000, we assigned a value of -$15,000.
    - If mothers said they would have a debt amount more than $25,000, we assigned a value of -$30,000.

  We generated an alternative categorical variable for net worth (`hhnetworthcata3`) with categories corresponding to the bullet points above.

Section L: Expenditures and Economic Stress

- **About this Section:** This section asks about child-focused expenditures**,** economic stress, food expenditures, food insecurity, and assets and expenditures. It includes the following pre-registered outcomes: (1) index of child-focused expenditures in the last 30 days, (2) food insecurity, and (3) index of economic stress.
- **Child-Focused Expenditures Subsection:** This subsection measures child-focused expenditures using purchases made in the last 30 days.
    - **Raw Variables:** For the five items of the index of child-focused expenditures in the last 30 days, mothers are first asked if they purchased any of the items. If mothers say yes, then they are asked how much money they spent on the item. Due to the skip pattern, these raw variables have missing values if mothers said "no" to having purchased the item in the last 30 days.
    - **Generated Variables:** We generated a set of five item-level variables where the amount spent on each item is equal to 0 (instead of missing) if they have not purchased the items. Then, we generated another set of five variables that truncate each amount using the 99th percentile. The pre-registered

index of child-focused expenditures in last 30 days (hhchildexpense30daysa3) is an additive index of the total amount spent on all five items, using the non-truncated item-level variables.

- **Economic Stress Subsection:** This subsection asks mothers about their experiences with economic worries and hardships, and it includes five of the nine items that are used to construct our index of economic stress. The other four items for this index were asked in Section E, Section H, and in the food insecurity subsection below. The items were dichotomized and reverse coded as described in the pre-registration table. The pre-registered index of economic stress (hheconstressa3) is an additive index of all nine items, with higher values indicating more economic stress.
- **Food Expenditures Subsection:** This subsection asks mothers how much they and their household spend on food and how much is received in food stamps. Whether the mother receives food stamps is asked in Section K. In this subsection, mothers are asked whether someone else in their household receives food stamps (foodstampsa3). If mothers reported receiving food stamps themselves (in Section K) or if anyone else in their household received them (responded "yes" to foodstampsa3), they were asked how much the household received in food stamps (foodstampamta3). They were then asked if they spent money out-of-pocket on food in addition to what they buy with food stamps (ofoodamta3). All mothers received one of two versions of the question, "How much do you and everyone else in your family spend on food that you use at home in an average week?" (ofoodamtwka3). For mothers who reported that they or someone else in their household receives food stamps, the question was preceded by the preamble, "In addition to food stamp benefits." For these mothers, their answers should not include their food stamp benefits. "Amount spent eating out in an average week" is also asked in this section. There are no generated variables for these items.
- **Food Insecurity Subsection:** This subsection asks mothers about their experiences of food insecurity. The USDA food security 6-item short form module was used to measure this construct. The module contains five questions with a conditional sixth question. If mothers respond "yes" to cutting meals (hhcutmealsizea3), then they are asked how often this occurs (hhcutmealfreqa3). Our pre-registered food insecurity scale (hhfoodinsecuritya3) in the Age-3 data is an additive index of all six items (recoded as binary indicators as described in the pre-registration table), with higher values indicating more food insecurity. Two additional questions regarding utilities are asked to the mother and used in the index of economic stress.

Section M2: Maternal Health (same questions as in M1 except randomized to appear after questions on economic stress)

Section Q: Maternal Health

- **About this Section:** This section asks mothers about their physical and mental health. There are five preregistered outcomes in this section: (1) depression; (2 & 3) anxiety; (4) sleep; and (5) perceived stress.
- **Maternal Depression Subsection:** Mothers' depressive symptoms were measured using the Patient Health Questionnaire (PHQ-8) scale, which includes eight common symptoms of depression. The raw variables use a scale from 1 to 4 (1: not at all; 2: several days; 3: more than half of days; 4: every day). We generated recoded variables that use a scale from 0 to 3 (0: not at all; 1: several days; 2: more than half of days; 3: every day). The pre-registered generated index (mphq8a3) is an additive index of the eight recoded PHQ-8 items, with higher values indicating more depressive symptoms.
- **Maternal Anxiety Subsection:** Mothers' anxiety symptoms were measured using both the Generalized Anxiety Disorder scale (GAD-7) and the Beck Anxiety Inventory. Which measure they received first was decided randomly.
    - *GAD-7*: Mothers were asked the eight items of the GAD-7, which includes seven items about the frequencies of common anxiety symptoms and one item about the difficulty of symptoms. The seven raw variables about frequencies of symptoms use a scale from 1 to 4 (1: not at all; 2: several days; 3: more than half of the days; 4: nearly every day). We generated recoded variables that use a scale from 0 to 3 (0: not at all; 1: several days; 2: more than half of the days; 3: nearly every day). The

pre-registered generated index (mgada3) is an additive index of these seven recoded GAD-7 items, with higher values indicating more anxiety symptoms. The raw variable for difficulty of symptoms used a scale from 1 to 4 (1: not difficult at all; 2: somewhat difficult; 3: very difficult; 4: extremely difficult), and we generated a recoded variable (mcripplingfeelingsa3) that uses a scale from 0 to 3 (0: not difficult at all; 1: somewhat difficult; 2: very difficult; 3: extremely difficult).

- *BAI*: Mothers' anxiety symptoms were measured using the Beck Anxiety Inventory (BAI), which includes 21 common anxiety symptoms. The raw variables use a scale from 1 to 4 (1: not at all; 2: mildly; 3: moderately; 4: severely bothersome), and we generated recoded variables that use a scale from 0 to 3 (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome). The pre-registered generated index (mbecka3) is an additive index of the 21 recoded BAI items, with higher values indicating more anxiety symptoms.

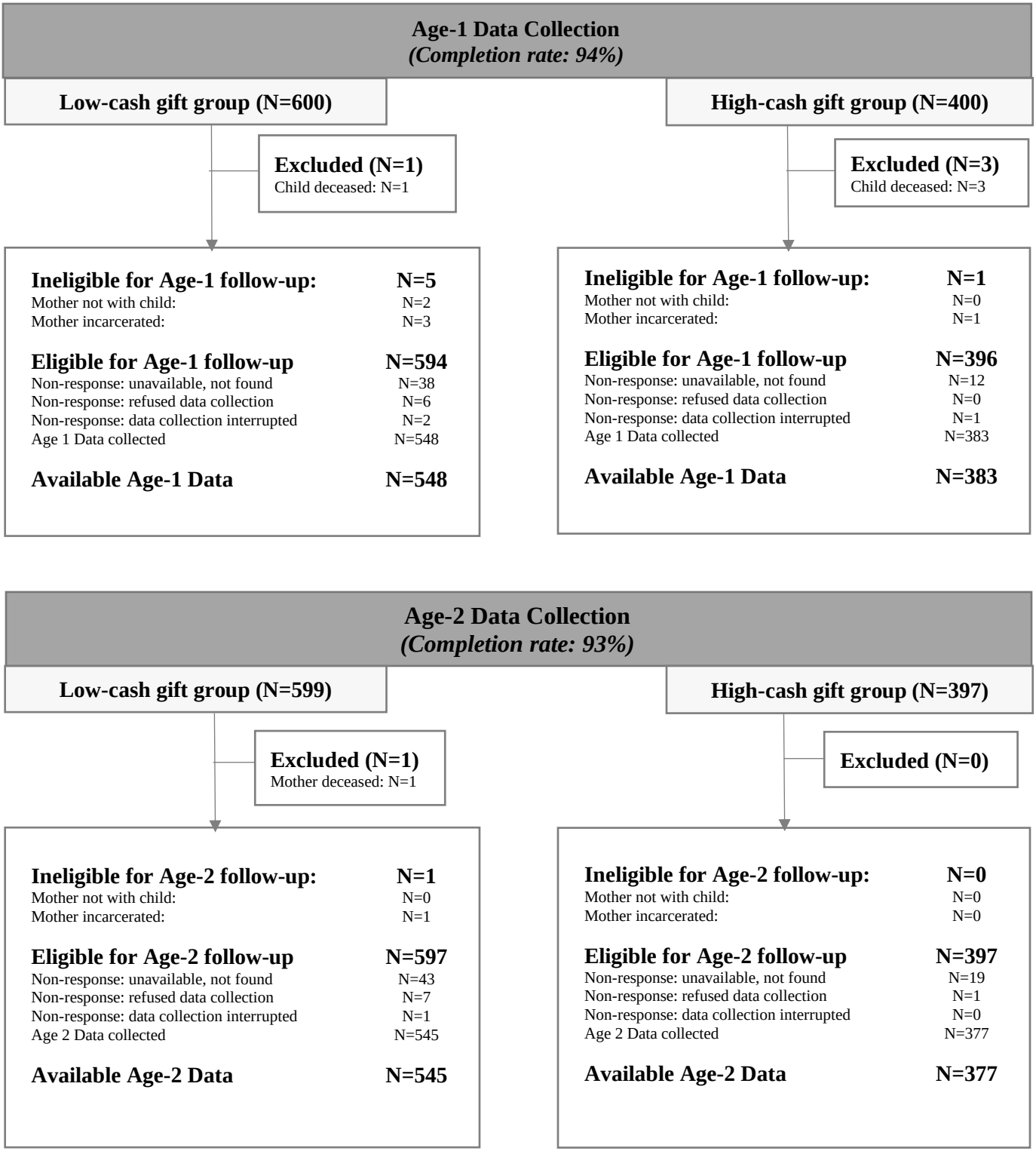
- **Maternal Sleep Subsection:** Mother's sleep (msleepa3) was measured by summing up three items with a 5-point response scale, 2 of which were reverse coded (mhardtosleepa3 and mmtireda3)
- **Perceived Stress Subsection:** Mothers' perceived stress was measured using the Perceived Stress Scale (PSS), which assesses the degree to which the respondent has perceived situations as stressful within the last month. The ten raw variables use a scale from 1 to 5 (1: Never; 2: Almost never; 3: Sometimes; 4: Fairly often; 5: Very often). We generated recoded variables using a revised scale from 0 to 4 (0: Never; 1: Almost never; 2: Sometimes; 3: Fairly often; 4: Very often). Four of the items were reverse scored (0: Very often; 1: Fairly often; 2: Sometimes; 3: Almost never; 4: Never). The pre-registered generated scale (mperceivedstressa3) is an additive index of the ten items, with higher values indicating more perceived stress. Note: This is different from the Age-2 survey, because one item was erroneously omitted, leaving a total of 9 items drawn from the intended 10-question scale. This was fixed at Age-3.

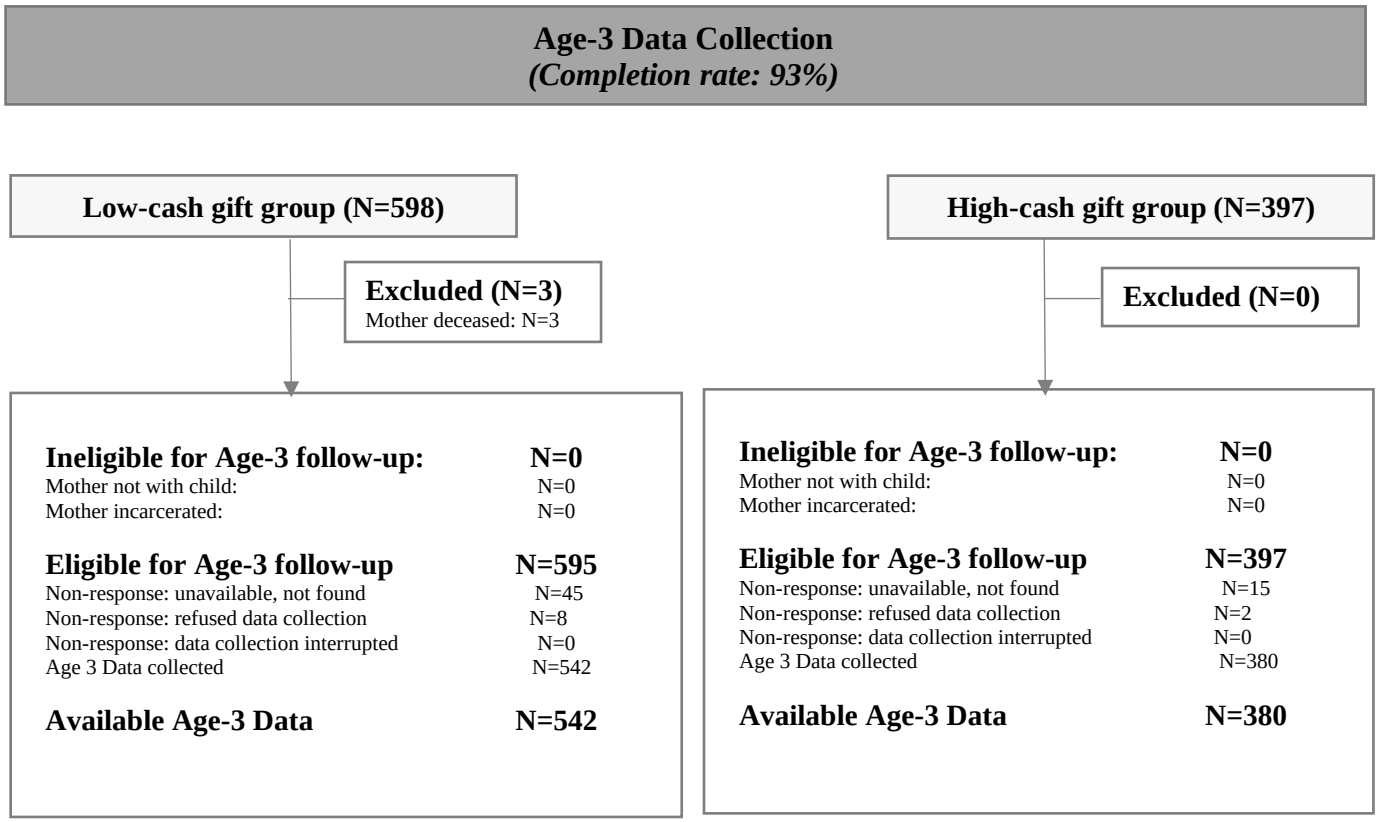
Section V: Closing

- **About this Section:** This is a short section to close out the survey. It asks for the mother's current address and whether the mother revealed the gift amount to the person on the phone giving the survey. No variables were generated from this section.

APPENDIX TABLES AND FIGURES

Appendix Figure 1. Age 1, Age 2, and Age 3 Consort Diagram




Age-3 Data Collection
(Completion rate: 93%)
Low-cash gift group (N=598)
Excluded (N=3)
Mother deceased: N=3
Ineligible for Age-3 follow-up: N=0
Mother not with child: N=0
Mother incarcerated: N=0
Eligible for Age-3 follow-up N=595
Non-response: unavailable, not found N=45
Non-response: refused data collection N=8
Non-response: data collection interrupted N=0
Age 3 Data collected N=542
Available Age-3 Data N=542
High-cash gift group (N=397)
Excluded (N=0)
Ineligible for Age-3 follow-up: N=0
Mother not with child: N=0
Mother incarcerated: N=0
Eligible for Age-3 follow-up N=397
Non-response: unavailable, not found N=15
Non-response: refused data collection N=2
Non-response: data collection interrupted N=0
Age 3 Data collected N=380
Available Age-3 Data N=380

Appendix Table 1. Maternal and Family Focused Pre-Registered Hypotheses

Domains (in gray) and sub-domains	Measure/ Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Household Economic Hardship					
Index of economic stress	MTO; Kling, Liebman, Katz, 2007			1, 2, 3, 4	Additive index of dichotomous variables (higher score=more stress): 1. worried about expenses? (0: occasionally or never; 1: frequently or more) 2. whether spent more than income? (0: no; 1: yes) 3. missed rent or mortgage (0 if homeless or not missed; 1 if missed rent or mortgage) 4. Set aside rainy day funds for 1 mo (0: Yes 1: No) 5. Ability to cover expenses for 1 mo with loss of income (0: Yes; 1: No) 6. in past 12 mos, missed payments for water, gas, oil, electricity? (0: no or not applicable; 1: yes) 7. in past 12 mos, gas, water, electricity ever shut off? (0: no; 1: yes) 8. Since child's birth, have you ever been evicted or forced to leave? (0: No; 1: Yes).*changes to "in the past 12 months" for surveys at ages 2 through 4 9. needed medical or dental care and did not get it? (0=no; 1=yes) *item 9 dropped at age 4 owing to survey time constraint
Household Poverty rate	US Census Bureau			1, 2, 3, 4	Measured using the Census Bureau's poverty thresholds by size of family and number of children
Index of food insecurity*	Economic Research Service, USDA, 2012			1, 2, 3, 4	Additive index of 6 dichotomized items (higher score=more food insecurity): 1. Food didn't last, no $ for more (0: Never true, 1: sometimes or often true) 2. Can't afford balanced meals (0: Never true, 1: sometimes or often true) 3. Cut size or skip means (0: No; 1: Yes) 4. If yes to (3), how often? (0: only one or two months; 1: almost every month or some months) 5. Eat less than should (0:No; 1: Yes) 6. Hungry^{+} (0:No; 1: Yes)
Social Services Receipt					

Number of Benefits received by mother	Study PIs	1, 2, 3	Additive index of dichotomized items (higher score=more benefits received): 1. Food stamps SNAP (0: not currently receiving; 1: currently receiving) 2. Free or reduced childcare* 3. Early Head Start or HS* 4. Women, Infants and Children (WIC) 5. State Unemployment 6. Cash assistance/TANF* 7. Medicaid coverage for self 8. Housing assistance 9. LIHEAP / heat/AC assistance* *Indicates benefits that were not asked about at age 3. Note: Age 4 benefit index was not pre-registered because of the availability of administrative records for some of the benefits
Mother's Labor Market and Education Participation			
Time to labor market reentry from birth	Current Population Survey	1	Continuous outcome: # of months until mom's reentry into labor market from birth of child derived from the following items: 1. did you ever work for pay since child's birth? 2. in what months did you work for pay?
Time to full-time labor market reentry from birth	Current Population Survey	1	Continuous outcome: # of months until mom's full-time reentry into labor market from birth of child derived from the following items: 1. did you ever work full time since child's birth? 2. in what months did you work full time?
Mother's education and training attainment	Current Population Survey	1, 2, 3	Dichotomous variable indicating that mother participated in education and/or job training activities since birth* *changes to "in the past 12 months" for surveys at ages 2 and 3
Mother's Labor Market Participation	Current Population Survey	4	Dichotomous variable indicating whether mother is participating in the labor market using the item "do you currently work for pay?"
Maternal Earnings	PSID	4	Mother's Earnings in the previous calendar year
Child-Focused Expenditures			
Index of child-focused expenditures (since birth)	Lugo-Gil, Yoshikowa, 2006	1	Additive index of the following dichotomous items (higher score=more purchased): Since child's birth, purchased… 1. Crib? 2. Car seat? 3. High chair? 4. Safety covers for outlets? 5. Latches for cabinets? 6. Gate? 7. Smoke detector? 8. books (yes/no)?

Index of child-focused expenditures (in past 30 days)	Lugo-Gil, Yoshikowa, 2006	1, 2, 3, 4	Continuous dollar amount of age-relevant items*: Past 30 days, total $ amount spent on... 1. books 2. toys 3. clothes 4. diapers 5. videos for age 1; 1. books 2. toys 3. clothes 4. activities 5. videos for ages 2, 3, and 4
Cost of paid child care	National Study of Early Care and Education	1, 2, 3, 4	Out of pocket spending on child care last week. 1. altogether, about how much money did you spend out-of-pocket on all of [CHILDNAMEF]'s child care arrangements last week? Note: Age 4: dropped wording "out-of-pocket"
Use of center-based care	National Study of Early Care and Education	1	1. Has child spent any time in childcare or day care? (Y/N)
		2, 3, 4	1. Has child spent 5 or more hours in a child care or day care center last week? (Y/N)
Housing and Neighborhoods			
Index of perceptions of neighborhood safety	MTO; Kling, Liebman, Katz, 2007	1, 2, 3	Additive index of two items (higher score=feels more safe). 1. how safe during day? (3: very safe, 2: safe, 1: unsafe, 0: very unsafe)2. how safe during night? (3: very safe, 2: safe, 1: unsafe, 0: very unsafe)
Index of housing quality	MTO; Kling, Liebman, Katz, 2007	1	Additive index of 7 items (higher score=higher quality): 1. Bad walls (0: big problem; 1: small problem; 2: not problem) 2. bad plumbing 3. rodents 4. cockroaches 5. bad windows 6. bad heat 7. overall condition (3: excellent, 2: good 1: fair, 0: poor)
		2	Additive index of 9 items (higher score=higher quality): 1. Bad walls (0: big problem; 1: small problem; 2: not problem) 2. bad plumbing 3. rodents 4. cockroaches 5. bad windows 6. bad heat 7. bad air condition 8. bad locks˜ 9. overall condition (3: excellent, 2: good 1: fair, 0: poor)
Homelessness	MTO; Kling, Liebman, Katz, 2007	1, 2, 3	Dichotomous indicator of whether the mother has ever been homeless or in a group shelter (age 1 "since birth", age 2-3 "in the past 12 months"): 0: No 1: Yes

			4	Dichotomous indicator of whether mom experienced "homelessness, eviction, or sudden loss of housing in the past 12 months".
Excessive Residential mobility	MTO; Kling, Liebman, Katz, 2007		1, 2, 3	Moved three or more times since birth of baby* (Y/N) *changes to "in the last 12 months" for surveys at ages 2 and 3
Neighborhood poverty	Kling, Liebman, Katz, 2007		1, 2, 3, 4	# of residents below poverty line in census tract divided by total number of residents in census tract
Family and Maternal Perceived Stress				
Perceived stress	Cohen et al., 1994, 1983	alpha: .86	1, 2	Perceived Stress Scale (PSS): additive index of 9 items (0: never; 1: almost never; 2: sometimes; 3: fairly often; 4: very often) 1. upset because of something unexpected 2. felt unable to control important life things 3. felt nervous and stressed 4. confident in ability to handle personal probs (reverse coded - rc) 5. couldn't cope with all things to do 6. control of irritations in life (rc) 7. "on top of things" (rc) 8. angered bc of things outside control 9. could not overcome difficulties
			3	Perceived Stress Scale (PSS): additive index of 10 items (0: never; 1: almost never; 2: sometimes; 3: fairly often; 4: very often) 1. upset because of something unexpected 2. felt unable to control important life things 3. felt nervous and stressed 4. confident in ability to handle personal probs (reverse coded - rc) 5. couldn't cope with all things to do 6. control of irritations in life (rc) 7. "on top of things" (rc) 8. angered bc of things outside control 9. could not overcome difficulties 10. felt things were going "your way" (rc)

Parenting stress	Items 1-4: Project GAIN Items 5-7: PSID-Child Development Supplement		1, 2 (originally also registered for age 4 and then dropped)	Aggravation in Parenting Scale: additive index of 7 items (0: Strongly agree-5: Strongly disagree): 1. confidence in parenting abilities 2. feels good about parenting abilities 3. thinks good parent 4. kids will say she was wonderful 5. giving up more for kids than ever expected 6. feels trapped (rc) 7. unable to do different things bc of kids (rc) Note: Index dropped from age 4 survey owing to time constraints
Maternal Happiness and Optimism				
Global happiness	The General Social Survey from NORC		1, 2, 3	One-item with 3-point response scale "Taken altogether, how happy are you these days?" (0: not happy; 1: pretty happy; 2: very happy)
Maternal Agency	Snyder et al., 1991	alpha: .86 test-retest: .81	1, 2, 3	HOPE Scale: additive index of 8 items with 5-point response scale (0: definitely false; 5: definitely true) 1. think of ways to get out of a jam, 2. energetic pursuit of goals, 3. lot of ways around any problem, 4. ways to get what's important, 5. solves problems, 6. past has prepared me for future, 7. pretty successful in life, 8. meets goals set for oneself
Maternal Physiological Stress				
Maternal hair cortisol	Ursache et al., 2017		1, 4	At age 1, we attempted to collect maternal hair cortisol for all in-person visits, prior to the onset of the pandemic (when data collection became limited to phone-based survey administration only). This resulted in a hair sample being collected from 409 of the 605 mothers who participated in an in-person visit, with large racial and ethnic differences in willingness to provide a sample. At age-4, we attempted to improve collection rates following focus groups and the development of informational videos. However, the first several months of data collection again revealed large racial and ethnic differences in willingness to provide a hair sample, due to both cultural and practical reasons. Because of the large amounts of non-random missing data, which would both compromise our statistical power and limit the generalizability of any findings, we dropped hair cortisol from the age-4 data collection procedures on October 25, 2022.
Maternal Mental Resources				

Maternal cognitive resources	Carlson, 2017; Carlson, & Zelazo 2014		4	Minnesota Executive Function Scale
Maternal Mental Health				
Index of maternal depression	Kroenke & Spitzer, 2002		1, 2, 3, 4	PHQ-8: additive index of 8 items (0: not at all; 1: several days; 2: more than half of days; 3: every day) 1. little interest or pleasure doing things 2. feeling down, depressed, hopeless 3. trouble sleeping or sleep too much 4. feel tired and no energy 5. poor appetite or overeating 6. feel like a failure 7. trouble concentrating 8. moving slowly or fidgety
Index of maternal anxiety	Steer & Beck, 1997	alpha: .92 test-retest: .75	1, 3	Beck Anxiety Inventory: additive index of 21 common anxiety symptom items (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome)
	Spitzer et al., 2006	alpha: .92 test-retest: .83	2, 3, 4	GAD-7: additive index of 7 items (0: not at all; 1:several days; 2: more than half the days; 3: nearly every day)
Maternal Substance abuse[×]				
Alcohol and cigarette use	MTO; Kling, Liebman, Katz, 2007		1, 3	Additive index of the following items (0: never in last year; 1: less than 1x per month; 2: several times per month; 3: several times per week; 4: everyday): 1. How often do you smoke cigarettes? 2. How often drink alcohol?
Opioid use	MTO; Kling, Liebman, Katz, 2007		1, 3	Number of times of opioid use in the past year (0: never in last year; 1: less than 1x per month; 2: several times per month; 3: several times per week; 4: everyday):
Chaos in Home				

Index of chaos in the home	Evans et al., 2005	alpha: .77 test-retest: .93	1, 2	Home Environment Chaos Scale: additive index of 20 items (higher score=more chaos): (0: not true; 1: true) 1. can find things (reverse coded - rc) 2. little commotion in home (rc) 3. always rushed 4. can "stay on top of things" (rc) 5. always late 6. "zoo" in home 7. can talk wo interruption (rc) 8. always a fuss 9. family plans don't work out 10.can't hear oneself think at home 11. drawn into others' arguments 12. can relax at home (rc) 13. phone takes up a lot of time 14. atmosphere is calm at home (rc) 15. regular morning routine (rc) 16. eat together during daily (rc) 17. evening routine with child (rc) 18. regular late afternoon routine with child (rc) 19. child goes to bed at regular time (rc) 20. set aside for talking with child daily (rc)
Maternal Relationships				
Physical Abuse	Fragile Families and Child Wellbeing Study		1,2	1. Ever abused? (1: yes; 0: no)
Frequency of Arguing			1,2	1. How often do you argue about things that are important to you? (1: never; 2: rarely; 3: sometimes; 4: often; 5: always)
Relationship quality			1	Additive index of the following items (higher score=higher qual rel) 1. Partner fair and willing to compromise? (3: Often; 2: sometimes; 1: never) 2. partner expressed affection or love? (3: Often; 2: sometimes; 1: never) 3. partner insulted or criticized you or your ideas (0: Often; 1: sometimes; 2: never) 4. partner made you feel down or bad about yourself during an argument? (0: Often; 1: sometimes; 2: never) 5. partner encouraged or helped you to do things that were important to you? (2: Often; 1: sometimes; 0: never) 6. partner isolated you? (0: Often; 1: sometimes; 2: never) 7. partner hurt you physically (0: Often; 1: sometimes; 2: never) 8. partner sexually abused you? (0: Often; 1: sometimes; 2: never) 9. partner listened to you? (3: Often; 2: sometimes; 1: never) 10. partner made you feel afraid? (0: Often; 1: sometimes; 2: never) 11. partner threatened or hurt your child/children?$^{+}$ (0: Often; 1: sometimes; 2: never)

			2, 3	Dichotomous indicator of current or recent relationship quality, where poor quality is defined as 1 if the mother is in a relationship and has a score of 26 or below on the relationship quality scale (approximately the bottom tercile of the low cash gift group distribution of scores) and a 0 either if the mother is not in a relationship or is in a relationship and has a relationship quality index score of 27 or above (approximately in the top two terciles of the distribution).
Maternal Physical Health				
Global health	Idler & Benyamini, 1997		1, 2	One item with 5-point response scale "overall, how would you describe your health…" (1:poor - 5:excellent)
Sleep	Yu et al., 2012		1, 3	Additive index of the following items (higher score=higher qual sleep): 1. Quality of sleep (0: very poor-5: very good) 2. Difficulty falling asleep (0: not at all; 5: very much) (rc) 3. Felt tired (0: not at all-5: very much) (rc)
Mother's BMI	CDC scales		4	Measured by CDC BMI percentile scales
Parent-Child Interaction Quality				
Adult word count	Xu et al (2009), LENA foundation		1	Measured using LENA processing software
Conversational turns	Xu et al (2009), LENA foundation		1	Measured using LENA processing software
Index of mother's positive parenting behaviors	Roggman, et al., 2013; Griffen & Friedman, 2007; Belsky, et al., 2007	inter-rater reliability varies by domain: .69-.80; alpha: .78	1, 4	Measured using PICCOLO coding of parenting behaviors from the total of four sub-scales (affection, responsiveness, encouragement and teaching) with responses ranging from 0: absent, 1: barely, 2: clearly. The total composite score is preregistered. Exploratory analyses will examine differences across the subscales, and factor analysis will be used to confirm the extent to which the four subscales best fit the data. Parent child interaction task and script adapted from the NICHD Study of Early Child Care and Youth Development.
Epigenetic Pace of Aging				

Methylation pace of aging	Belsky et al., 2020; Belsky et al., 2022	4	Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort. Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). In contrast, so-called epigenetic clocks are trained on chronological age. Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. The second iteration (DunedinPACE) takes into account an additional measurement occasion (collected 20 years after inclusion) and only includes the most reliable DNA methylation probes, i.e., probes with little variation between technical replicates.
Maternal DNA Methylation			
DNA methylation	McCartney et al, 2022	4	Salivary DNA-methylation profiles of cognitive functioning, i.e., "Epigenetic-g", can be computed on the basis of weights from a blood-based epigenome wide association study of general cognitive functions (g) in adults (McCartney et al., 2022). General cognitive ability was derived from the first unrotated principal component of logical memory, verbal fluency and digit symbol tests, and vocabulary. Epigenetic-g is conceptually distinct from biological aging.
Frequency of Parent Child Activity			
Self-Report of Parent-child activities	Rodriguez & Tamis-LeMonda, 2011	1	Additive index of 4 items with response scale (higher score=higher frequency of activities): 1. read books (0: rarely or never; 1: a few times/month; 2: a few times/week; 4: everyday) 2. tell stories 3. play together 4. play groups
		2, 3	Additive index of 5 items with response scale (higher score=higher frequency of activities):1. read books (0: rarely or never; 1: a few times/month; 2: a few times/week; 4: everyday) 2. tell stories, 3. play together, 4. play groups (not asked at age 3 due to COVID), 5. play pretend games
Child meal and sleep routine index	Study PIs	4	Additive index of 2 survey items (higher score=more routines): 1. eat meals together (0: 0 days; 1: 1+ days) 2. had regular bedtime (0: no; 1: yes)

Time on mother-focal child activities	Rodriguez & Tamis-LeMonda, 2011	4	Additive index of activities where the number of days reported doing the activity are multiplied by the number of minutes on a given day. Activities are: read books, tell stories, play game/build something, pretend play, learning activities, screen activities. 1. How many days did you participate in [activity]? (0: no days; 1.5: 0-1 days; 4: 3-5 days; 6.5: 6-7 days) 1a. On those days, how many minutes do you do [activity]? (2: 4 minutes or less; 7.5: 5-10 minutes; 15.5: 11-20 minutes; 25.5: 21-30 minutes; 35: more than 30 minutes).
Maternal Disciplineˣ			
Spanking discipline strategy	Reichman et al., 2001	1, 2, 3	Dichotomous indicator using the following item: 1. In past month, have you spanked child due to misbehavior (1: yes; 2:no)

Notes. The previous version of this table referred to "waves" of data collection. For clarity, we have replaced "wave" with "age," with both referring to the age of the baby at planned data collection. Minor, non-substantive changes may be made to the wording of specific items across data collection years.

+ indicates that items were omitted or programmed incorrectly in the age 1 survey administered to mothers and cannot be used to calculate outcomes. These include item 6 from the index of food insufficiency ("hungry"), and item 11 from the relationship quality index ("partner threatened or hurt your child/children? "). These indices were therefore comprised of one less item at Age-1.
ˣindicates outcomes that were not administered at Age-1 once in-person interviews switched to phone interviews due to COVID-19

˜Indicates that item was omitted from previous pre-registrations but was administered to mothers and is being included in the outcome analyses.

*Indicates that the sub-domain was called something different in previous versions of this table. The sub-domain "Food Insecurity" was previously referred to as "Food Insufficiency."

Due to COVID-19, the Age-2 and Age-3 data collection wave is in the form of a phone survey. Thus, sub-domains that were supposed to be measured in-person at Age-2 or Age-3 are being postponed to ages 45-48 months. These domains include: index of mother's positive parenting behaviors, epigenetic age, DNA methylation, BMI, physiological stress, cognitive resources. Additionally, sub-domains that we had not intended to include in pre-registration at Age-3 have been added to the phone survey at Age-3 and to the pre-registration table. These include: self-report of parent-child activities, spanking discipline strategy, anxiety.
Certain sub-domains were pre-registered at Age-3 and are no longer pre-registered because they are not being included in the Age-3 data collection (due to time constraints). These include: global health, physical abuse, index of chaos in the home, parenting stress, index of housing quality.

Bibliography of Maternal and Family Focused Pre-Registered Hypotheses

Measure description	Bibliography	
Preregistered measures	**Source 1**	**Source 2**
Household Economic Hardship		
Index of economic stress	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.	http://www2.nber.org/mtopublic/
Index of food insecurity	https://www.ers.usda.gov/media/8282/short2012.pdf	
Household poverty rate	Fontenot, Kayla, Jessica Semega, and Melissa Kollar, U.S. Census Bureau, Current Population Reports, P60-263, Income and Poverty in the United States: 2017, U.S. Government Printing Office, Washington, DC, 2018.	
Social Services Receipt		
Number of Benefits received by mother	Study PIs	
Mother's Labor Market and Education Participation		
Time to labor market reentry from birth	Current Population Survey, retrieved from: https://www.census.gov/programs-surveys/cps/technical-documentation/questionnaires.html	
Time to full-time labor market reentry from birth		
Mother's education and training attainment		
Maternal Earnings	Panel Study of Income Dynamics https://psidonline.isr.umich.edu/	
Child-Focused Expenditures		
Index of child-focused expenditures	Lugo-Gil, J., Yoshikawa, H. (2006). Assessing expenditures on children in low-income, ethnically diverse, and immigrant families. National Poverty Center Working Paper Series, 06-36.	
Child-focused expenditures		
Cost of paid child care	National Study of Early Care and Education	
Use of center-based care		
Housing and Neighborhoods		
Index of perceptions of neighborhood safety	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.	

Index of housing quality		
Residential mobility		
Homelessness		
Neighborhood poverty		
Family and Maternal Perceived Stress		
Perceived stress	Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. Measuring stress: A guide for health and social scientists.	Cohen, S., Kamarck, T., Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385-396.
Parenting stress	PSID-CDS Aggravation in Parenting Scale https://psidonline.isr.umich.edu/cds/cdsi_usergd.pdf for items 5-7 5. giving up more for kids than ever expected 6. feels trapped (rc) 7. unable to do different things bc of kids (rc)	Project GAIN (Gaining Access to Income Now) https://preventionboard.wi.gov/Pages/OurWork/ProjectGAIN.aspx for items 1-4 1. confidence in parenting abilities 2. feels good about parenting abilities 3. thinks good parent 4. kids will say she was wonderful
Maternal Happiness and Optimism		
Global happiness	The General Social Survey from NORC at the University of Chicago, retrieved from: http://gss.norc.org/Get-Documentation/questionnaires	
Maternal Agency	Snyder, C.R., Harris, C., Anderson, J.R., Holleran, S.A., Irving, L.M., Sigmon, S.T., Yoshinobu, L., Gibb, J., Langelle, C., Harney, P. (1991). The will and the ways: development and vaildation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, *60*(4), 570-585.	
Maternal Epigenetic Pace of Aging		
Epigenetic age	Belsky, W. D. et al. (2020). Quantification of the pace of biological aging in humans through blood test, the DunedinPoAm DNA methylation algorithm. eLife 9:e54870. https://doi.org/10.7554/eLife.54870	Belsky, W. D. et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. eLife 11:e73420. https://doi.org/10.7554/eLife.73420
Maternal DNA Methylation		

DNA methylation	McCartney, D.L., Hillary, R.F., Conole, E.L.S. *et al.* Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biol* **23,** 26 (2022). https://doi.org/10.1186/s13059-021-02596-5	
Maternal Physiological Stress		
Maternal hair cortisol	Ursache, A., Merz, E.C., Melvin, S., Meyer, J., Noble, K.G. (2017). Socioeconomic status, hair cortisol and internalizing symptoms in parents and children. *Psychoneuroendocrinology*, 78, 142-150.	
Maternal Mental Resources		
Maternal cognitive resources	Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale: Test Manual. St. Paul, MN: Reflection Sciences, Inc.	Carlson, S. M. (2017). *Minnesota Executive Function Scale*: *Technical Report, v. 2*. St. Paul, MN: Reflection Sciences, Inc.
Maternal Mental Health		
Index of maternal depression	Kroenke, K. & Spitzer, R.L. (2002). The PHQ-9: a new depression diagnostic and severity measure. Psychiatric annals, 32(9), 509-515.	
Index of maternal anxiety	Steer, R.A. & Beck, A.T., (1997). Beck Anxiety Inventory. In C.P. Zalaquett & R.J. Wood (Eds), Evaluating stress: A book of resources (pp. 23-40). Lanham, MD, US: Scarecrow Education	
Index of maternal anxiety	Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med.* 2006;166(10):1092–1097. doi:10.1001/archinte.166.10.1092	
Maternal Physical Health		
Global health	Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.	
Sleep	Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., ... & Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. Behavioral sleep medicine, 10(1), 6-24.	
Mother's BMI	Kuczmarski, R. J. (2000). CDC growth charts; United States.	
Maternal Substance abuse		
Alcohol and cigarette use	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.	

Opioid use		
Chaos in Home		
Index of chaos in the home	Evans, G.W., Gonnella, C., Marcynyszyn, L.A., Gentile, L, & Salpekar, N. (2005). The role of chaos in poverty and children's socioemotional adjustment. *Psychological Science*, 16(7), 560-565.	
Maternal Relationships		
Physical Abuse		
Frequency of Arguing	User's Guide for the Fragile Families and Child Wellbeing Study Public Data, Year 3. (2018). Retrieved from: https://fragilefamilies.princeton.edu/sites/fragilefamilies/files/year_3_guide.pdf#page=84	
Relationship quality		
Parent-Child Interaction Quality		
Adult word count	Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in young children's natural home environment. *LENA Foundation*.	
Conversational turns		
Index of mother's positive parenting behaviors	Roggman, L.A., Cook, G.A., Innoccenti, M.S., Norman, V.J., Christiansen, K. (2013). Observations Linked to Outcomes (PICCOLO) Of Diverse Ethnic Groups. *Infant Mental Health Journal, 34(4), 290-306.*	Griffin, J. A., & Friedman, S. L. (2007). NICHD Study of Early Childcare and Youth Development. *National Institute of Health*
Frequency of Parent Child Activity		
Self-Report of Parent-child activities	Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child development*, *82*(4), 1058-1075.	
Child meal and sleep routine index	Study PIs	
Time on mother-focal child activities	Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child development*, *82*(4), 1058-1075.	
Maternal Discipline		

Spanking discipline strategy	Reichman, N.E., Teitler, J.O., Garfinkel, I., MclAnahan, S.S. (2001). Fragile Families: Sample and design. *Children and Youth Services Review, 23*(4-5), 303-326.

Appendix Table 2. Child Focused Pre-Registration Hypotheses

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Language Development					
Language Milestones	Squires et al., 2009	sensitivity .86 specificity .85		1	Measured using ASQ- Communication Subscale
Vocabulary*	Fenson, 2002; Jackson-Maldonado, 2012	internal consistency .85		2	Measured by short-form versions of the MacArthur Communicative Development Inventories
	Martin & Brownell, 2011		4		Measured by Receptive One Word Picture Vocabulary Test (ROWPVT) We will administer the monolingual (English) or bilingual (English/Spanish) versions as appropriate. Because the two versions of the test are not co-normed, the primary outcome will be a derived "conceptual score," or sum of the raw scores on all individual items that appear on both versions of the test.
Maternal concern for language delay	Glascoe, 1997		3		Measured by the sum of the two questions included in the PEDS on expressive language and articulation and receptive language: 1. Do you have any concerns about how your child talks and makes speech sounds? (0: No; 1: Yes or a little) 2. Do you have any concerns about how your child understands what you say? (0: No; 1: Yes or a little)
Executive Function and Behavioral Regulation					
Executive Function	Diamond & Taylor, 1996; Weiland& Yoshikawa, 2013; Bierman et al., 2008			Originally registered for age 4 and then dropped	Intended to be measured by the pencil tap test. This item was preregistered as an age-4 secondary outcome but was dropped on September 13, 2022, due to evidence of floor effects, and numerous reports from research staff that children were not understanding the instructions.
Executive Function	Carlson, 2017; Carlson, & Zelazo 2014	MEFS: validity .92 test-retest .93	4		Measured by the Minnesota Executive Function Scale.
Socio-Emotional Processing					

Social-Emotional Problems	Briggs-Gowan et al., 2004	internal consistency .65-.79 test-retest reliability .87		1, 2	Measured by the Brief Infant–Toddler Social and Emotional Assessment (BITSEA)
Behavior/Emotional Problems	Achenbach et al., 2000	parent report reliability .80	3, 4		Measured by a shortened version of the Child Behavior Checklist measuring the following areas: emotionally reactive, anxious/depressed, attention problems, and aggressive behavior. At age 3, we will estimate the statistical significance of the entire family of related measures in the Child Socio-Emotional Processing outcome cluster measured during the same wave using step-down resampling methods for multiple testing (see statistical analysis plan for more details; Westfall and Young, 1993).
Social-Emotional Behavior	Roggman et al., 2013; Griffen & Friedman, 2007; Belsky, 2007			Originally registered for age 1 but unable to be coded	Measured using NICHD SECCYD parent-child-interaction task coding scheme, with child codes Positive Mood, Negative Mood, Activity Level, Sustained Attention, Positive Engagement at age 1 and agency, negativity, persistence, affection at age 4. (Due to funding limitations, this was not feasible to code, and we have no immediate plans to do so).
Maternal concern for behavioral and social-emotional problems	Glascoe, 1997		3		Measured by the sum of the two questions included in the PEDS on behavior and social-emotional: 1. Do you have any concerns about how your child behaves? (0: No; 1: Yes or a little) 2. Do you have any concerns about how your child gets along with others? (0: No; 1: Yes or a little)
IQ					

IQ*	Wechsler & Naglieri, 2006	internal consistency .88test-retest reliability .77	Originally registered for age 4 but not able to be calculated	4, for matrices subtest only	The Wechsler Nonverbal Scale of Ability was originally pre-registered as a Primary Outcome. The IQ score is calculated using two subtests -- Matrices and Recognition -- and we began our fieldwork on July 9, 2022 with both. On the basis of preliminary analysis of the first 71 cases, we discovered that 21% of participants scored at the floor of the Recognition assessment. We therefore dropped the Recognition subtest from our data collection instrument on September 30 2022, precluding us from calculating IQ in subsequent participants. Scores on the Matrices subtest, which measures visual processing and abstract spatial perception (not IQ per se), are now registered as an age-4 secondary outcome
Pre-Literacy					
Pre-Literacy	Hutton et al., 2019; Hutton et al., 2021			4	Measured by The Reading House
Resting Brain Function					
Age-1 Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004	n/a		1	Measured by low-density mobile electroencephalography at Age 1: we preregistered group differences in theta, alpha, gamma power.
Age-4 Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004; Troller-Renfree et al. 2022	n/a	4	4	Measured by high-density in-lab electroencephalography: **Age-4 Primary**: Because of limitations in power expected with multiple testing adjustments, we are preregistering a single composite of mid-to-high-frequency whole-brain power summing across alpha, beta, and gamma bands, from 7 to 45 Hz. **Age-4 secondary**: We hypothesize greater frontal gamma power in the high-cash gift group, and plan to analyze a full model of regions nested within bands, with the plan to report all exploratory outcomes. See attached analysis plan. Note: The original preregistration of EEG data collected when children were 12 months old included hypotheses across multiple frequency bands. Please see the history of preregistrations, including analysis plans.
Task-Related Brain Function					

Auditory Discrimination Brain Function*	Choeur et al., 2000; Garcia-Sierra et al., 2011; Kuhl et al., 2005	n/a		4	Measured by mismatch negativity (MMN) ERP with larger differences between standard and deviant stimulus in high-cash gift group compared to the low-cash gift group.
Health: BMI					
Body Mass Index (BMI)	Kuczmarski, 2000	n/a		4	Measured by CDC BMI percentile scales
Health: Physiological Stress					
Physiological Stress	Ursache et al., 2017; Meyer et al., 2014; Davenport et al., 2006	n/a		Originally registered for age 4 but unable to collect	Our original plan was to measure physiological stress using hair cortisol concentration. The first several months of data collection revealed large racial and ethnic differences in willingness to provide a hair sample, due to both cultural and practical reasons. Because of the large amounts of non-random missing data, which would both compromise our statistical power and limit the generalizability of any findings, we dropped hair cortisol from our data collection procedures on October 25, 2022.
Health: Sleep					
Sleep problems	Yu et al., 2012	reliability .9	3	1, 2	Measured by PROMIS Sleep Disturbance- Short Form adapted from ECHO; For ages 1 and 2, additive index of the following items with 5-point answer (0: never; 1: almost never; 2: sometimes; 3: almost always, 4: always): 1. difficulty falling asleep 2. sleeping through night (reverse coded) 3. problem with sleep 4. problem sleeping For Age 3, item 1 was not included in the survey
Health: Other Indicators					

Overall Health, Medical Care, Diagnosis of Condition or Disability	Child's overall health item source: Idler & Benyamini, 1997Halim et al., 2013	n/a	3	1, 2	Additive index of the following items*: 1. Child's overall health? (4: excellent, 3: very good, 2: good, 1: fair, or 0: poor), 2. About how many times in the last year did you take child to a doctor because [he/she] was sick? 0-1 times, 2-5 times, 6+, 3. About how many times in the last year did you take child to a doctor because [he/she] was hurt or injured?, 4. Did you ever have to take child to the Emergency Room because [he/she] was sick, hurt or injured? (Y/N), 5. How many times ER?, 6. Has child been diagnosed with any health condition or disability since birth? (Y/N)*factor analysis of items will be conducted to scale the index
Overall Health, Diagnosis of Health Condition or Disability	Child's overall health item source: Idler & Benyamini, 1997	n/a		4	Additive index of the following items: 1. Child's overall health? (4: excellent, 3: very good, 2: good, 1: fair, or 0: poor) 2.About how many times in the last year was child sick? 0-1 times, 2-3 times, 4-6 times, 7+ 3. Has child been diagnosed with any chronic health condition? (Y/N)
Diagnosis of Developmental Condition					
Diagnosis of Developmental Condition	Study PIs	n/a		4	Has child been diagnosed with any developmental condition, like speech delay, autism, or ADHD? (Y/N)
Child Epigenetic Pace of Aging					

Methylation pace of aging	Belsky et al., 2020; Belsky et al., 2022	n/a	4	Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort. Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). In contrast, so-called epigenetic clocks are trained on chronological age. Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. The second iteration (DunedinPACE) takes into account an additional measurement occasion (collected 20 years after inclusion) and only includes the most reliable DNA methylation probes, i.e. probes with little variation between technical replicates. If a higher quality measure of epigenetic aging at the time of analysis becomes available, we will substitute that instead.
Child DNA Methylation				
DNA methylation	McCartney et al, 2022	n/a	4	Salivary DNA-methylation profiles of cognitive functioning, i.e., "Epigenetic-g", can be computed on the basis of weights from a blood-based epigenome wide association study of general cognitive functions (g) in adults (McCartney et al., 2022). General cognitive ability was derived from the first unrotated principal component of logical memory, verbal fluency and digit symbol tests, and vocabulary. Epigenetic-g is conceptually distinct from biological aging. If a higher quality measure of epigenetic profile of cognitive functioning becomes available at the time of analysis, we will substitute that instead.
Child Nutrition				
Consumption of healthy foods	Los Angeles County WIC Survey, 2017		2	Additive index of the number of times per day consumed the following items*: 1. eat fruits 2. eat vegetables

Consumption of unhealthy foods	Los Angeles County WIC Survey, 2017				2	Additive index of the number of times per day consumed the following items*: 1. juice, soda, chocolate milk or other sweet drinks 2. eat sweets
Parents' Evaluation of Developmental Status (PEDS)	Glascoe, 1997				3	Measured by the total score across categories of components of the PEDS, which includes 10 survey items.
Total "predictive concerns" in the PEDS	Glascoe, 1997				3	Measured by the total number of maternal-reported concerns that are "predictive of developmental delay" in the PEDS
School Achievement & Behavior						
School test scores for target children and siblings	Administrative data	n/a		School age (target child)	School age (siblings)	
Student behavioral data for target children and siblings	Administrative data	n/a			School age (target child and siblings)	

Bibliography of Child Focused Pre-Registration Hypotheses

Language Development	
Language Milestones	Squires, J., Bricker, D. D., & Twombly, E. (2009). Ages & stages questionnaires. Baltimore, MD: Paul H. Brookes.
Language Processing	Golinkoff, R. M., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., & Brezack, N. (2017). *User's Manual for the Quick Interactive Language Screener (QUILS): A Measure of Vocabulary, Syntax, and Language Acquisition Skills in Young Children*. Paul H. Brookes Publishing Company
Vocabulary*	Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. Applied Psycholinguistics, 21, 95 – 116. Jackson-Maldonado, Donna, Virginia A. Marchman, and Lia C. H. Fernald. 2012. "Short-Form Versions of the Spanish MacArthur-Bates Communicative Development Inventories." Applied Psycholinguistics 34 (4): 837–68. Martin, N. A., & Brownell, R. (2011). ROWPVT-4: Receptive One-Word Picture Vocabulary Test.
Maternal concern for language delay	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
Executive Function & Self-Regulation	
	Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "do as I say, not as I do". *Developmental psychobiology*, *29*(4), 315–334. https://doi.org/10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T
Executive Function - Pencil Tap	Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Development and psychopathology*, *20*(3), 821–843. https://doi.org/10.1017/S0954579408000394

	Weiland, C. and Yoshikawa, H. (2013), Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. Child Dev, 84: 2112-2130. https://doi.org/10.1111/cdev.12099
Executive Function - MEFS	Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale: Test Manual. St. Paul, MN: Reflection Sciences, Inc. Carlson, S. M. (2017). *Minnesota Executive Function Scale*: *Technical Report, v. 2*. St. Paul, MN: Reflection Sciences, Inc.
Socio-Emotional Processing	
Social-Emotional Problems	Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of pediatric psychology*, *29*(2), 143-155.
Behavior/Emotional Problems	Achenbach, T. M., & Ruffle, T. M. (2000). The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in* review, 21(8), 265-271.
Social-Emotional Behavior^	Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V., & Christiansen, K. (2013). Parenting interactions with children: Checklist of observations linked to outcomes (PICCOLO) in diverse ethnic groups. Infant Mental Health Journal, 34(4), 290-306. Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care?. *Child development*, *78*(2), 681-701. Griffin, J. A., & Friedman, S. L. (2007). NICHD Study of Early Childcare and Youth Development. National Institute of Health. Adapted script from mother-child-interaction at 15 months.
Maternal concern for behavioral and social-emotional problems	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
IQ	
IQ*	Wechsler, D., Naglieri, J. A. (2006). Wechsler Nonverbal Scale of Ability. San Antonio, TX: Pearson.

Pre-Literacy	
Pre-Literacy	John S. Hutton, Laura Justice, Guixia Huang, Amy Kerr, Thomas DeWitt, Richard F. Ittenbach; The Reading House: A Children's Book for Emergent Literacy Screening During Well-Child Visits. Pediatrics June 2019; 143 (6): e20183843. 10.1542/peds.2018-3843 Hutton, J. S., Dudley, J., Huang, G., Horowitz-Kraus, T., DeWitt, T., Ittenbach, R. F., & Holland, S. K. (2021). Validation of *The Reading House* and Association With Cortical Thickness. *Pediatrics*, *147*(3), e20201641. https://doi.org/10.1542/peds.2020-1641
Resting Brain Function	
Age-1 and Age-4 Resting Brain Function	Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., ... & Kushnerenko, E. (2013). Socioeconomic status and functional brain development–associations in early infancy. *Developmental Science*, *16*(5), 676-687. Otero, G. A., Pliego-Rivero, F. B., Fernández, T., & Ricardo, J. E. E. G. (2003). EEG development in children with sociocultural disadvantages: a follow-up study. Clinical neurophysiology, 114(10), 1918-1925. Marshall, P. J., Fox, N. A., & Group, B. C. (2004). A comparison of the electroencephalogram between institutionalized and community children in Romania. Journal of Cognitive Neuroscience, 16(8), 1327-1338. Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., ... & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences*, *119*(5), e2115649119.
Task-Related Brain Function	
Auditory Discrimination Brain Function*	Cheour, M., Leppänen, P. H., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical neurophysiology*, *111*(1), 4-16. Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., ... & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. Journal of Phonetics, 39(4), 546-557. Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. Developmental science, 8(1), F1-F12.
Health: BMI	
Body Mass Index (BMI)	Kuczmarski, R. J. (2000). CDC growth charts; United States.

Health: Physiological Stress	
Physiological Stress	Ursache, A., Merz, E. C., Melvin, S., Meyer, J., & Noble, K. G. (2017). Socioeconomic status, hair cortisol and internalizing symptoms in parents and children. *Psychoneuroendocrinology*, *78*, 142-150. Meyer, J., Novak, M., Hamel, A., & Rosenberg, K. (2014). Extraction and analysis of cortisol from human and monkey hair. Journal of visualized experiments: JoVE, (83). Davenport, M. D., Tiefenbacher, S., Lutz, C. K., Novak, M. A., & Meyer, J. S. (2006). Analysis of endogenous cortisol concentrations in the hair of rhesus macaques. General and comparative endocrinology, 147(3), 255-261.
Health: Sleep	
Sleep problems	Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., ... & Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. Behavioral sleep medicine, 10(1), 6-24.
Health: Other Indicators	
Overall Health, Medical Care, Diagnosis of Condition or Disability	Halim, M. L., Yoshikawa, H., & Amodio, D. M. (2013). Cross-generational effects of discrimination among immigrant mothers: Perceived discrimination predicts child's healthcare visits for illness. *Health Psychology*, *32*(2), 203. Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37
Diagnosis of Developmental Condition	
Diagnosis of Developmental Condition	Study PIs
Child Epigenetic Pace of Aging	
Methylation pace of aging	Belsky, W. D. et al. (2020). Quantification of the pace of biological aging in humans through blood test, the DunedinPoAm DNA methylation algorithm. eLife 9:e54870. https://doi.org/10.7554/eLife.54870 Belsky, W. D. et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. eLife 11:e73420. https://doi.org/10.7554/eLife.73420

Child DNA Methylation	
DNA methylation	McCartney, D.L., Hillary, R.F., Conole, E.L.S. *et al.* Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biol* **23,** 26 (2022). https://doi.org/10.1186/s13059-021-02596-5
Child Nutrition	
Consumption of healthy foods	Los Angeles County WIC Survey. (2017). Retrievable from: http://lawicdata.org/wp-content/uploads/2014/09/WIC-Parents-Quex-English-FINAL.pdf
Consumption of unhealthy foods	Los Angeles County WIC Survey. (2017). Retrievable from: http://lawicdata.org/wp-content/uploads/2014/09/WIC-Parents-Quex-English-FINAL.pdf
Any Maternal Concern for Developmental Delay	
Parents' Evaluation of Developmental Status (PEDS)	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
Total "predictive concerns" in the PEDS	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.

Appendix Table 3. Baseline Balance by High and Low Cash Gift Groups at Age 3 sample (n=922)

	Low Cash Gift Mean (sd)	N	High Cash Gift Mean (sd)	N	Std Mean Difference Hedges' g	Cox's Index	p-value
Child is female	0.502	542	0.479	380		-0.056	0.492
Child weight at birth (pounds)	7.1 (1.059)	541	7.1 (1.012)	379	-0.047		0.482
Child gestational age (weeks)	39.1 (1.268)	538	39.0 (1.244)	380	-0.029		0.668
Mother age at birth (years)	26.9 (5.870)	542	27.4 (5.770)	380	0.080		0.202
Mother education (years)	12.0 (2.777)	535	11.9 (2.957)	378	-0.038		0.597
Mother race/ethnicity: white, non-Hispanic	0.111	542	0.082	380		-0.203	0.104
Mother race/ethnicity: Black, non-Hispanic	0.387	542	0.439	380		0.130	0.080
Mother race/ethnicity: multiple, non-Hispanic	0.044	542	0.029	380		-0.262	0.198
Mother race/ethnicity: other or unknown	0.048	542	0.024	380		-0.435	0.035
Mother race/ethnicity: Hispanic	0.410	542	0.426	380		0.040	0.408
Mother marital status: never married	0.423	542	0.497	380		0.181	0.022
Mother marital status: single, living with partner	0.262	542	0.211	380		-0.172	0.069
Mother marital status: married	0.210	542	0.216	380		0.022	0.815
Mother marital status: divorced/separated	0.046	542	0.029	380		-0.290	0.175
Mother marital status: other or unknown	0.059	542	0.047	380		-0.145	0.422
Mother health is good or better	0.882	542	0.924	380		0.295	0.032
Mother depression (CESD)	0.7 (0.455)	542	0.7 (0.448)	380	-0.003		0.922
Cigarettes per week during pregnancy	5.0 (21.250)	538	3.3 (11.388)	377	-0.099		0.091
Alcohol drinks per week during pregnancy	0.2 (1.705)	540	0.0 (0.391)	379	-0.111		0.053
Number of children born to mother	2.4	542	2.5	380	0.091		0.183

	(1.387)		(1.411)				
Number of adults in household	2.1	542	2.0	380	-0.080		0.231
	(0.986)		(0.975)				
Biological father lives in household	0.402	542	0.347	380		-0.142	0.089
Household combined income	22,484.99	511	20,777.03	353	-0.087		0.189
	(21,904.09)		(15,889.13)				
Household income unknown	0.057	542	0.071	380		0.142	0.401
Household net worth	-1,791.29	480	-2,241.72	339	-0.018		0.770
	(29,910.37)		(12,793.65)				
Household net worth unknown	0.114	542	0.108	380		-0.037	0.783
Joint Test: Chi2(30)= 34.17, p-value= 0.195, n=918.							

Notes: P-values were derived from a series of OLS bivariate regressions in which each respective baseline characteristic was regressed on the treatment status indicator using robust standard errors and site-level fixed effects. The bivariate regressions were also run without site-level fixed effects. The p-values without fixed effects do not appear in the table. The joint test of orthogonality was conducted using a probit model with robust standard errors and site-level fixed effects.
Standardized mean differences were calculated using Hedges' g for continuous variables and Cox's Index for dichotomous variables.
If there were more than 10 missing cases for a covariate, missing data dummies were included in the table and the joint test. If there were less than 10 cases missing, missing data dummies were not included in the table but were included in the joint test.
Chi-square tests of independence were conducted for the two categorical variables: mother race/ethnicity and mother marital status. For both tests, p>0.05
All respondents with missing data on gestational age are in the control group. Thus, this dummy was removed from the joint test due to perfectly predicting failure. This results in a slightly smaller sample for the joint test.

Appendix Table 4. Age 3 Instrument Versions

Date Released	Version	Description
7/21/2021	1 – 7.14.2021	Released to production.
8/3/2021	2 – 8/3/21	1. Added statement "[IF INSTRUMENT B, DISPLAY]: (Given your situation), Some of the questions related to your child will not be asked but we would still like to complete the interview to see how you are doing. The interview will be shorter and take 30-45 minutes." at *Instrument.* 2. Updated page number for M2-M10 to "page 3" 3. Updated page number for Q.39 to "page 11" 4. Updated mis-spelled Spanish text for various questions 5. H8 – [he/she] updated to fill based on child gender in question text 6. Updated Instrument B questions to be excluded (R26-R28a, I36-I88) per spec 7. Removed AdminConsent questions MailAdminConsent field from Blaise
10/25/21	3 – 10/7/21	1. Added Masking questions after Parent/Child Activities. 2. Added vaccination question after H17. 3. Added Iwer instructions to I51, I76, I78 to clarify meaning of words 4. Added routing logic for DK at WhereLiveMost
1/26/22		TCC/Fairview consents updated with new stamps
3/15/22	4 – 3/10/22	Updated HHChildRel question to read: What is [HHCHILD1NAME]'s relationship to you? Are they your… ♦ READ the list of options
3/15/22		NYSPI consents updated with modification to Confidentiality Agreement section.
4/12/22		TCC consents updated with new stamps and modification to Confidentiality Agreement section.
5/25/22		NYSPI consents updated with new stamps.
5/25/22	5 – 5/23/22	The year range has been extended from 2021 to 2022 ChildDOD.ChildDeathY ChildDOB.Child_DOB_Year

Appendix Table 5. BFY Pre-Registered Measure Comparisons Across Waves

Pre-R Measure/ Construct	Age-1 Variable Name	Age-1 Items	Age-2 Variable Name	Age-2 Items	Age-3 Variable Name	Age-3 Items
Index of Economic Stress	hheconstressa1	Additive index of nine items, including: (1) Worried about expenses; (2) Spent more than income; (3) Missed rent or mortgage; (4) Set aside rainy day funds for 1 mo; (5) Ability to cover expenses for 1 mo with loss of income; (6) Missed payments for water, gas, oil, electricity; (7) Gas, water, electricity ever shut off; (8) Ever been evicted or forced to leave; (9) Needed medical or dental care and did not get it Reference period of past 12 months	hheconstressa2	No change from Age-1	hheconstressa3	No change from Age-1
Index of Food Security	hhfoodinsecurityа1	Additive index of 5 items: (1) Food didn't last, no $ for more; (2) Can't afford balanced meals; (3) Cut size or skip meals; (4) How often cut meals; (5) Eat less than should	hhfoodinsecurityа2	Includes an additional item asking about hunger which was erroneously excluded from Age-1	hhfoodinsecurityа3	No change from Age-2
Number of Benefits Received	hhsocialservicesa1	Additive index of 10 social services: (1) food stamps/SNAP, (2) Free or Reduced Child Care, (3) Early Head Start, (4) Head Start, (5) Women, Infants, and Children (WIC), (6) State Unemployment, (7) Cash Assistance/TANF, (8) Medicaid coverage for self, (9) Housing Assistance, and (10) LIHEAP/Heat/AC Assistance	hhsocialservicesa2	Combines Early Head Start and Head Start for an additive index of 9 social services	hhsocialservicesa3	Additive index of 5 social services: (1) food stamps/SNAP; (2) Women, Infants, and Children (WIC), (3) State Unemployment, (4) Medicaid coverage for self, (5) Housing Assistance
Mother's Education and Training	medjobtraina1	Mother participated in education or job training since focal child's birth	medjobtraina2	Mother participated in education or job training in the last 12 months	medjobtraina3	No change from Age-2
Index of Child Focused Expenditures	hhchildexpense30daysa1	Continuous dollar amount of total $ spent on the following items in the last 30 days: (1) Books; (2) Toys; (3) Clothes;	hhchildexpense30daysa2	Replaces "diapers" with "activities"	hhchildexpense30daysa3	No change from Age-2

		(4) Diapers; (5) Videos				
Cost of Paid Child Care	hhpaidcccosta1	Out-of-pocket spending on child care in the last week	hhpaidcccosta2	No change from Age-1	hhpaidccosta3	No change from Age-1
Neighborhood Safety	hhneighbsafetya1	2-item additive index of perceptions of neighborhood safety during the day and at night	hhneighbsafetya2	No change from Age-1	hhneighbsafetya3	No change from Age-1
Use of Center-Based Child Care	hhusecentercarea1	Indicator for whether child has spent any time in center-based childcare since birth	hhdclastweeka2	Indicator for whether child spent 5 or more hours in center-based care in the last week	hhdclastweeka3	No change from Age-2
Index of Housing Quality	hhhousingqualitya1	8-item additive index: (1) bad walls; (2) bad plumbing; (3) rodents; (4) cockroaches; (5) bad locks; (6) bad windows; (7) bad heat; (8) overall condition	hhhousingqualitya2	Incudes a ninth item: bad air conditioning	Not included	Not included
Homelessness	hhhomelessorsheltera1	Ever in a shelter or experienced homelessness since child's birth	hhhomelessorsheltera2	Reference period changes to last 12 month	hhhomelessorsheltera3	No change from Age-2
Excessive Residential Mobility	hhexcessivemovea1	Moved three or more times since child's birth	hhexcessivemovea2	Reference period changes to last 12 month	hhexcessivemovea3	No change from Age-2
Perceived Stress Scale	mperceivedstressa1	Perceived Stress Scale (PSS): additive index of 9 items: (1) upset because of something unexpected; (2) felt unable to control important life things; (3) felt nervous and stressed; (4) confident in ability to handle personal probs (reverse coded - rc); (5) couldn't cope with all things to do; (6) control of irritations in life (rc); (7) "on top of things" (rc); (8) angered bc of things outside control; (9) could not overcome difficulties	mperceivedstressa2	No change from Age-1	mperceivedstressa3	Includes a 10th item: felt things were going "your way" (rc)
Parenting Stress Scale	mparentingstressa1	Additive index of 7 items: (1) confidence in parenting abilities; (2) feels good about parenting abilities; (3) thinks good parent; (4) kids will say she was wonderful; (5) giving up more for kids than ever expected; (6) feels trapped (rc); (7) unable to do different things bc of kids (rc)	mparentingstressa2	No change from Age-1	Not included	Not included

Global Happiness	mhappya1	Mother's global happiness on three-point response scale	mhappya2	No change from Age-1	mhappya3	No change from Age-1
Maternal Agency: HOPE Scale	mHOPEa1	Additive index of 8 items	mHOPEa2	No change from Age-1	mHOPEa3	No change from Age-1
Maternal Depression (PHQ-8)	mphq8a1	PHQ-8: additive index of 8 items (0: not at all; 1: several days; 2: more than half of days; 3: every day): (1) little interest or pleasure doing things; (2) feeling down, depressed, hopeless; (3) trouble sleeping or sleep too much; (4) feel tired and no energy; (5) poor appetite or overeating; (6) feel like a failure; (7) trouble concentrating; (8) moving slowly or fidgety	mphq8a2	No change from Age-1	mphq8a3	No change from Age-2
Maternal Anxiety	mbecka1	Beck Anxiety Inventory: additive index of 21 common anxiety symptom items (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome) 1. numbness 2. felling hot 3. wobbliness in legs 4. restless 5. fear of the worst happening 6. dizzy or lightheaded 7. heart pounding/racing 8. unsteady 9. terrified or afraid 10. nervous 11. feeling of choking 12. hands trembling 13. shaky/ unsteady 14. fear of losing control 15. difficulty breathing 16. fear of dying 17. scared 18. indigestion 19. faintness/ lightheaded 20. face flushed 21. hot/cold sweats	mgada2	GAD-7: additive index of 7 items (0: not at all; 1:several days; 2: more than half the days; 3: nearly everyday) 1. Feeling nervous, anxious, or on edge 2. Not being able to stop or control worrying 3. Worrying too much about different things 4. Trouble relaxing 5. Being so restless that it is hard to sit still 6. Becoming easily annoyed or irritable 7. Feeling afraid, as if something awful might happen	mbecka3; mgada3	Beck Anxiety Inventory: No change from Age-1 GAD-7: No change from Age-2
Alcohol and Cigarette Use	malcciga1	2 item index of mother's self-report of smoking and drinking frequency	Not included	Not included	malcciga3	No change from Age-1
Opioid Use	moipioida1	Frequency of opioid use	Not included	Not included	mopioida3	No change from Age-1

Relationship Quality	mrelationqualitya1	Additive index of 10 items: (1) Partner fair and willing to compromise; (2) Partner expressed affection or love; (3) Partner insulted or criticized you or your ideas; (4) Partner made you feel down or bad about yourself during an argument; (5) Partner encouraged or helped you to do things that were important to you; (6) Partner isolated you; (7) Partner hurt you physically; (8) Partner sexually abused you; (9) Partner listened to you; (10) Partner made you feel afraid Only asked for mothers who completed the survey in person (prior to March 2020)	mrelationqualitya2	Eleventh item added: Partner threatened or hurt your child/children. Sample included only mothers who identified as being in a relationship at the time of the survey	mrelationqualitya3	No change in items from Age-2. All mothers were asked these items and prompted to think about their most recent relationship. ##121 mothers volunteered that they were not in a relationship and thus were not asked these items
Maternal Sleep Quality	msleepa1	Additive index of three items: (1) sleep quality; (2) Difficulty falling asleep; (3) Felt Tired	Not included	Not included	msleepa3	No change from Age-1
Parent-Child Activities	mparentchildacta1	Additive index of 4 items with response scale (0: rarely or never; 1: a few times/month; 2: a few times/week; 4: everyday): (1) read books; (2) tell stories; (3) play together; (4) play groups	mparentchildacta2	Includes a fifth item: Play pretend games	mparentchildacta3	Exclude play groups, and includes play pretend games for a 4 item scale.
Spanking/Harsh Discipline	hhspanka1	Indicator for whether mother spanked child due to misbehavior in the past month	hhspanka2	No change from Age-1	hhspanka3	No change from Age-1
Social-Emotional	cbitseaproblema1	BITSEA: 31 items from the problems scale plus two additional items asked in error	cbitseacompa2 cbitseaproblema2	BITSEA: 11 items from the competence scale and 31 from the problems scale	--	--
Behavior-Emotional		--		--	ccbclindexa3	CBCL
Child Sleep: PROMIS	cPROMISa1	Additive index of the following items: (1) difficulty falling asleep; (2) sleeping through night (reverse coded); (3) problem with sleep; (4) problem sleeping	cPROMISa2	No change from Age-1	cPROMISa3	First item removed from index. Additive index of following 3 items: (1) sleeping through night (reverse coded); (2) problem with sleep; (3) problem sleeping

Child Health Index	csickhealtha1	Additive index of following items, with different response categories: (1) Child's overall health; (2) Number of times child taken to a doctor because [he/she] was sick; (3) Number of times child taken child to a doctor because [he/she] was hurt or injured; (4) Any visit to the ER due to injury or illness; (5) Number of visits to the ER; (6) Child diagnosed with health condition or disability	csickhealtha2	Different response categories for items 3 and 5. At age-1, these were free response answers; at age-2 the options were 0-1, 2-5, and 6 or more.	csickhealtha3	No change from Age-2

Table 6 Crosswalk of the pre-loaded Age-2 variables and newly collected Age-3 rostering variables.

Roster Information	Preloaded Age-2 Survey (Adults)		Age-3 Survey (Adults)		Preloaded Age-2 Survey (Children)		Age-3 Survey (Child)	
	Variable Stem	Position Suffix	Variable Stem	Position Suffix	Variable Stem	Position Suffix	Variable Stem	Position Suffix
Name	hhmemname	1-25	hhmemname	26-50	hhchildname	1-25	hhchildname	1-25
Gender	hhmemgender	1-25	hhmemsex	1-25	hhchild1sex	1-25	hhchild1sex	1-25
Relationship to the mother	hhmemrel	1-25	hhmemrel	26-50	childhhmemrel	1-25	childhhmemrel	1-25
Contributes to household income	hhmemcontr	1-25	hhmemcontr	26-50	hhchildcontr	1-25	hhchildcontr	1-25
Month of Birth	month	1-25	dob_mo	1-25	dob_mo		dob_mo	26-50
Year of Birth	year	1-25	dob_yr	1-25	dob_yr		dob_yr	26-50
Age	hhmemage	1-25	hhmemage	26-50	childhhmemage	1-25	childhhmemage	1-25
Member still lives in household	n/a	n/a	hhmemlivingwith	1-25	n/a	n/a	hhchild	1-25
There are other members	n/a	n/a	otheradultinhh + hhothadult	1-25	n/a	n/a	otherchildinhh + hhothchild	1-25
Employed in the last month	n/a	n/a	n/a	n/a	n/a	n/a	hhchildjob	1-25

Note: Each of these variables has "_a_*a3" after it, with * being the position suffix. For example, "hhmemname_a_1a3" or "dob_yr_a_1a3". The only exceptions are the flags otheradultsinhh and otherchildinhh, which simply have an "a3" at the end.

Table 7.a Stylized example of household roster's relative position and member list when household roster does not change between Age-2 and Age-3 survey.

Preloaded Age-2		Age-3 Survey Adult		Age-3 Survey Child	
1	Adult A	1	Adult A	1	(empty)
2	Adult B	2	Adult B	2	(empty)
3	Adult C	3	Adult C	3	(empty)
4	Child A	4	(empty)	4	Child A
5	Child B	5	(empty)	5	Child B

Table 7.b Stylized example of household roster's relative position and member list when existing members leave, and new members enter between Age-2 and Age-2 survey.

Preloaded Age-2		Age-3 Survey Adult		Age-3 Survey Child	
1	Adult A	1	Adult A	1	(empty)
2	Adult B	2	Adult B	2	(empty)
3	Adult C	3	~~*Adult C*~~ **Adult D***	3	(empty)
4	Child A	4	(empty)	4	Child A
5	Child B	5	(empty)	5	~~*Child B*~~ (empty) **Child C***
				6	

Notes: New household members are **bolded** with an asterisk and existing members who leave the household are ~~struck out~~ and *italicized*.

Table 7.c Stylized example of household roster's relative position and member list when a child at Age-1 becomes a legal adult by Age-2 survey.

Preloaded Age-2		Age-3 Survey Adult		Age-3 Survey Child	
1	Adult A	1	Adult A	1	**Child B**
2	Adult B	2	Adult B	2	(empty)
3	Adult C	3	Adult C	3	(empty)
4	**Child A**	4	**Adult D (Child A)**	4	(empty)
5	Child B	5	(empty)	5	(empty)

Notes: The child at Baseline who becomes a legal adult by the Age-2 Survey is bolded.