

# **Baby's First Years: Summary, Pre-registered Hypotheses, Analysis Strategies and Paper Plans**

**July 3, 2020**

## **Project Summary**

One thousand infants born to mothers with incomes falling below the federal poverty threshold in four metropolitan areas in the United States are being assigned at random within metropolitan area to one of two cash gift conditions. The sites are: New York City, the greater New Orleans metropolitan area, the greater Omaha metropolitan area, and the Twin Cities. IRB and recruiting issues will likely lead to a distribution of the 1,000 mothers across sites of roughly 115 in one site (the Twin Cities) and 295 in each of the three other sites. The high cash gift treatment group mothers (40% of all mothers) will receive unconditioned cash payments of \$333 per month (\$4,000 per year) via debit care for 40 months. Mothers in the low cash gift comparator group (60% of all mothers) receive a nominal payment – \$20 per month, delivered in the same way and also for 40 months. The 40/60 randomization assignment is stratified by site but not by hospitals within each of the four sites.

Mothers are being recruited in maternity wards of the 12 participating hospitals shortly after giving birth and, after consenting, are administered a 30-minute baseline interview. They then are asked to consent to the cash gifts. The three follow-up waves of data collection conducted at child ages 1, 2 and 3 will provide information about family functioning as well as developmentally appropriate measures of children's cognitive and behavioral development. An additional feature of our ages 1-3 data collection plans is that we will randomly assign a designated interview date within a one-month interval centered on the child's birthday. This provides variation in the timing of outcome data with respect to participants' receipt of the cash gift that will enable us to learn more about the incremental value of a stable predictable monthly infusion of cash.

We will collect information about the mother and child in the home when the child is 12 and 24 months of age (with information collected via phone during the COVID-19 restrictions on in-person data collection). At age 3, mothers and children will be assessed and interviewed in research laboratories at each site. Conditional on participants' consent and our success in securing agreements with state and county agencies, we will also collect state and local administrative data regarding parental employment, utilization of public benefits such as Medicaid and Supplemental Nutrition Assistance Programs (SNAP), and any involvement in child protective services. We also have plans to randomly sample 80 of the participating families in two of the sites (the Twin Cities and New Orleans) to participate in an in-depth qualitative study, but do not elaborate on those plans in this document.

The compensation difference between families in the high and low cash gift groups will boost family incomes by \$3,760 per year, an amount shown in the economics and developmental psychology literatures to be associated with socially significant and policy relevant improvements in children's school achievement. (We have worked with state and local officials to ensure to the extent feasible that our cash gifts are not considered countable income for the purposes of determining benefit levels from social assistance programs.) After accounting for likely attrition, our total sample size of 800 at age 3 years, divided 40/60 between high and low

payment groups, provides sufficient statistical power to detect meaningful differences in cognitive, emotional and brain functioning, and key dimensions of family context (see below).

Cognitive and emotional development measures will be gathered at 12, 24, and 36 months of age. At the age-three lab visit we will administer validated, reliable and developmentally sensitive measures of language, memory, executive functioning and socioemotional skills. We will also collect direct measures of young children's brain development at ages 1 and 3. Measures and preregistered hypotheses about them as well as family-based measures are shown in the two tables at the end of this document.

The family process measures that we will gather are based on two theories of change surrounding the income supplements: that increased investment and reduced stress will facilitate children's healthy development. We will obtain data measuring both of these pathways annually. *Investment pathway*: Additional resources enable parents to buy goods and services for their families and children that support cognitive development. These include higher quality housing, nutrition and non-parental child care; more cognitively stimulating home environments and learning opportunities outside of the home; and, by reducing or restructuring work hours, more parental time spent with children. *Stress pathway*: A second pathway is that additional economic resources may reduce parents' own stress and improve their mental health. This may allow parents to devote more positive attention to their children, thus providing a more predictable family life, less conflicted relationships, and warmer and more responsive interactions.

## **Analysis Plan**

*Pre-registered Hypotheses*. We preregistered hypotheses with [clinicaltrials.gov](https://clinicaltrials.gov) within a month after recruitment began (May, 2018) and in September, 2018 preregistered hypotheses with the [Registry of Effectiveness Studies](https://www.effectivepsychology.com/) and the [AEA RCT Registry](https://www.aea-rct.org/). Appendix Tables 1 and 2 detail our original hypothesized impacts and which groups of measures will be subject to multiple testing adjustments. Appendix Tables 3 and 4 incorporate minor changes to the tables that were originally posted in our pre-registrations. These changes are mostly made to data collection at age 2, with a few changes to age 3 data collection. There were no changes to age 1.

*Hypothesis Testing and Power Analysis*. Our key aims are to evaluate the impacts of income supplementation on: validated, reliable, and developmentally-sensitive measures of cognitive, language, memory, self-regulation, and socio-emotional functioning at child ages 1 (a small subset), 2 (a larger subset) and 3 (almost all) – this is Aim 1 in our NICHD application; developmentally-sensitive electroencephalographic-based measures of brain functioning at child ages 1 and 3 (Aim 2); and family expenditures, food insecurity, housing and neighborhood quality, parent stress and parenting practices, and child care arrangements gathered at child ages 1, 2 and 3 (Aim 3).

All of our pre-registered hypotheses focus on full-sample impacts, although we will also estimate in exploratory analyses moderation of impacts by gender, race/ethnicity (African American, Latino, White), family structure at birth and depth of poverty at birth (income to needs  $\leq .5$  or not). Before conducting these main analyses, all measures will be examined for psychometric equivalence across race/ethnicity and whether Spanish or English is a primary language spoken at home and we will compare high and low cash gift groups within site on all baseline characteristics to confirm successful implementation of random assignment.

Our basic empirical approach will use the survey and neuroscience data to compare the pooled cross-city \$333/month and \$20/month groups on a wide range of family process and child outcome measures. Because of random assignment, the low cash gift group average enables us to identify the average outcomes corresponding to the counterfactual state that would have occurred for individuals in the high cash gift group if they had not been offered the additional \$313/month income supplement. Therefore differences in outcomes for the high compared with the low group (after random assignment) can be interpreted as estimates of causal treatment effects of the \$313/month higher income (regardless of whether treatment-group participants actually expend all of the funds.) These are commonly known as intent-to-treat effects.

*Estimation strategy.* We illustrate our approach to estimation in a simple regression framework. The “Intent-To-Treat effect” (ITT) is captured by the estimate of the coefficient  $\pi_1$  in a regression of some child or family process outcome (Y) on a dichotomous indicator for assignment (Z) to the high payment group as in (1).

$$(1) \quad Y = Z\pi_1 + X\beta_1 + \varepsilon_1$$

Consistent with experiences from a 30-family pilot study we conducted in 2014, we anticipate extremely low rates of “non-compliance” with the offer of cash gifts paid via the debit cards.

We will adjust standard errors using robust variance estimation techniques (Cameron et al. 2008). We will estimate (1) without and then with baseline demographic child and family characteristics (X) to improve the precision of our estimates by accounting for residual variation. These baseline measures, all gathered prior to random assignment, will first be checked for adequate variation and sufficient independence from other baseline measures. They include: dummy variables for three of the four sites; mother’s age, completed schooling, household income, net worth, general health, mental health, race and Hispanic ethnicity, marital status, number of adult in the mother’s household, number of other children born to the mother, whether the mother smoked or drank alcohol during pregnancy and whether the father is currently living with the mother; and child’s sex, birth weight, gestational age at birth and birth order.

We will apply our regression estimation strategy to the assessment-based measures of cognitive, language, memory, self-regulation, and socio-emotional functioning at child ages 2 and 3, and the EEG measures of brain activity at ages 1 and 3 and ERP measures of brain activity at age 3 (see Appendix Tables 1 and 3). To investigate family process impacts, we will apply our estimation strategy to measures of stress physiology, family expenditures, food insecurity, housing and neighborhood quality, mothers’ executive function, parent stress and parenting practices, and child care arrangements gathered at child ages 1, 2 and 3, as shown in Appendix Tables 2 and 4 and described in the section on paper plans.

*Attrition.* The greatest threat to internal validity is potential bias from sample attrition overall, within site, and differential attrition rates by treatment status overall and within site. We will carefully track response rates by site, by treatment status across sites, and then treatment status within site. Any early signs of differential attrition will be expediently addressed through small, strategic adjustments in survey follow-up efforts, including use of financial incentives, or more tailored strategies such as using on-the-ground reconnaissance techniques to locate individuals. Based on the successes in our pilot study, our investigators’ prior experience with the Survey Research Center, and because of the continued contact with all participants the debit card ensures, we anticipate high response rates in later data collection (80+% at 36 months) with little to no differential attrition.

If necessary, we will consider a two-stage sampling procedure at the final stages of our data collection efforts during each wave in order to minimize attrition-related biases. The procedure calls for randomly subsampling from the remaining difficult-to-reach nonrespondents and concentrating resources and efforts to locate them. Analysis weights will be developed to adjust for the possible two-stage survey response sampling. This weighting approach has been successfully implemented in comparable studies. In addition to case-based nonresponse we also anticipate the usual (i.e., infrequent but not nonexistent) item-based nonresponse owing to refusals, interview breakoffs, etc.

We will also conduct sensitivity checks to evaluate whether missing data might be biasing estimates. Most sample attrition that is systematically related to our outcomes of interest (Y) would presumably also be related to the distribution of baseline characteristics (X), and so bias due to sample attrition would be evident if our estimates are sensitive to conditioning on baseline characteristics. Some attrition may be due to time-varying (or unobserved) characteristics and we can approach this problem in two ways. First, we will examine the sensitivity of our results to worst-case bounds, which enable us to bracket the true effects of our treatment without imposing any assumptions about the unobserved outcomes of participants (Manski, 1989; Manski, 1990; Manski, 1995). A second approach to addressing the problem of missing data will be to use multiple imputation strategies with all available data, (including all survey and administrative data on outcomes and predictor variables). Multiple imputation is an appropriate method if, conditional on all observed information, data are missing at random. Finally, because we expect relatively high rates (~80%) of baseline consent to collect administrative data, we will be able to compare survey respondents and survey non-respondents on formal earnings and receipt of income from social programs.

*Interpretation of parameters.* The coefficients obtained in our regression models will be used to quantify the causal effects of the \$313/month difference in income supplementation on age-1 and 3 child brain circuitry, cognitive development and socioemotional functioning. We will use the same methods to generate causal impact estimates for the family processes in each of the conceptual pathways. Examining the possible explanatory mechanisms in this way uses a series of separate regression equations to estimate program effects on possible treatment mediators, rather than estimating a structural-equation mediation model, and has been effectively used to infer possible mediation in comparable studies. This approach is preferred because it preserves the experimental variation in income generated by random assignment. The underlying insight is that randomization occurred with respect to receipt of the cash gifts and not on the basis of the proposed pathway mediators. With the potential for multiple mediators, a causal interpretation cannot be given to mediational models without very strong, often implausible, assumptions that there are no unobserved confounds of the association between the mediator and outcome. Still, the pattern of impacts can yield important insight as to which processes are likely to be present and absent and set the stage for future analyses.

*Statistical power.* The compensation difference between families in the high and low cash gift groups amounts to \$313 per month and \$12,520 over the course of the 40 months. This amount is in the range of income increases associated with child impacts of around .20 sd in studies of welfare experiments and the EITC (Duncan, Morris & Rodrigues, 2011; Morris, Duncan, Clark-Kauffman, 2005; Dahl & Lochner, 2012). After accounting for likely 20% attrition, and in the absence of adjustments for sample clustering within hospitals or increased precision owing to the inclusion of baseline covariates in our impact estimates, the sample size of

800 at age 3, divided 40%/60% between high and low payment groups, provides 80% statistical power to detect a .219 sd impact at  $p < .05$  in a two-tailed test on cognitive functioning and family processes. The use of baseline covariates in estimation models will improve this power, while the use of bootstrap standard errors will decrease it. Based on exploratory analyses of age-3 cognitive outcomes in the Fragile Families study, we expect that these two offsetting factors will have little net impact on the size of our estimated standard errors.

*Multiple comparisons.* One strength of our proposal is the collection of survey, neuroscience lab and administrative data on a wide range of outcomes and explanatory pathways. However, the probability of rejecting a true null hypothesis for at least one outcome is greater than the significance level used for each test. We will address the possibility of false positives while minimizing the reduction in statistical power to detect meaningful effects. Best-practice methods differ across disciplines so we will draw from multiple approaches with the goal of ensuring that results from one approach are consistent with results from others (Romano & Wolfe, 2005; Porter, 2018; Benjamini, 2010; Holm, 1979, Westfall & Young, 1993; Schochet, 2008). Where possible we have aggregated measures used to test our pre-registered hypotheses into indexes. In the case of related measures that cannot be aggregated into a single index, we will estimate the statistical significance of the entire family (“familywise error rate”) using step-down resampling methods in Westfall and Young (1993; Westfall, Tobias, Wolfinger, 2011). Pre-registered clusters of measures are identified with grey bars in appendix tables.

*Data release.* We will release data and documentation for our study to the research community at the end of each data collection wave once data are cleaned and coded, to enable independent researchers to pursue replication, mediation, moderation as well as other related analytic questions.

## References

- Benjamini, Y. (2010). Simultaneous and selective inference: Current successes and future challenges. *Biometrical Journal*, 52(6), 708–721. <https://doi.org/10.1002/bimj.200900299>
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414-427. <https://doi.org/10.1162/rest.90.3.414>
- Dahl, G. B., & Lochner, L. (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, 102(5), 1927–1956. <https://doi.org/10.1257/aer.102.5.1927>
- Duncan, G. J., Morris, P. A., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental Psychology*, 47(5), 1263-1279. <http://dx.doi.org/10.1037/a0023875>

- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70. <https://www.jstor.org/stable/4615733>
- Manski, C. F. (1989). Anatomy of the Selection Problem. *The Journal of Human Resources*, 24(3), 343–360. <https://doi.org/10.2307/145818>
- Manski, C. F. (1990). Nonparametric Bounds on Treatment Effects. *The American Economic Review*, 80(2), 319–323. <https://www.jstor.org/stable/2006592>
- Manski, C. F. (1995). Learning about social programs from experiments with random assignment of treatments. Institute for Research on Poverty Discussion Papers 1061-95, University of Wisconsin Institute for Research on Poverty.
- Morris, P. A., Duncan, G. J., & Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform: the sensitivity of transitions in development to policy change. *Developmental Psychology*, 41(6), 919–932. <https://doi.org/10.1037/0012-1649.41.6.919>
- Porter, K. E. (2018). Statistical Power in Evaluations That Investigate Effects on Multiple Outcomes: A Guide for Researchers. *Journal of Research on Educational Effectiveness*, 11(2), 267–295. <https://doi.org/10.1080/19345747.2017.1342887>
- Romano, J. P. & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237-1282. <https://doi.org/10.1111/j.1468-0262.2005.00615.x>
- Schochet, P. Z. (2008). *Guidelines for multiple testing in impact evaluations of educational interventions. Final report*. Princeton, NJ: Mathematica Policy Research, Inc. Retrieved from <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED502199>
- Westfall, P. H., Tobias, R. D., & Wolfinger, R. D. (2011). *Multiple comparisons and multiple tests using SAS, second edition*. Cary, NC: The SAS Institute.
- Westfall, P. H. & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. Hoboken, New Jersey: John Wiley & Sons.



Appendix Table 1: Original Child Focused Pre-registered Hypotheses

Measure description	Measure source	Psychometrics	Wave preregistered Primary Outcome	Wave preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Language Development</b>					
Language Milestones	Squires et al., 2009	sensitivity .86 specificity .85		1,2,3	Measured using ASQ- Communication Subscale
Child Vocalization	Xu et al., 2009	internal consistency .69-.86 test-retest reliability .76		2	Measured using LENA processing software
Verbal Comprehension	Wechsler, 2012	internal consistency .95 test-retest reliability .86-.92		3	Measured by WPPSI-IV- Vocabulary subtest
Language Processing	Golinkoff et al., 2017		3		Measured by Quick Interactive Language Screener (QUILS)- Language Processing Subscale
Communicative Development	Fenson, 2002	internal consistency .85		2	Measured by MacArthur Communicative Development Inventories
<b>Executive Function &amp; Self-Regulation</b>					
Self-Regulation	Smith-Donald et al., 2007	internal consistency of assessor report (not full assessment) .82-.93	3		Measured by the Preschool Self-Regulation Assessment - PSRA (parent report and child assessment)
Executive Function	Carlson, 2017 OR Wechsler, 2012 OR Willoughby et al., 2011	MEFS: validity .92 test-retest .93 WPPSI-IV: internal consistency .95 test-retest reliability .86-.92	3		Minnesota Executive Function Scale OR WPPSI-IV Working Memory OR EF Touch Executive Functioning
<b>Socio-Emotional Processing</b>					
Social-Emotional Problems	Briggs-Gowan et al., 2004	internal consistency .65-.79 test-retest reliability .87	3	1,2	Measured by the Brief Infant-Toddler Social and Emotional Assessment (BITSEA)- Problem Scale
Behavior / Emotional Problems	Achenbach et al., 2000	parent report reliability .80	3	2	Measured by the Child Behavior Checklist
Social-Emotional Behavior	Roggman et al., 2013; Griffen & Friedman, 2007			1,2	Measured using NICHD SECCYD parent-child-interaction task coding scheme, with child codes Positive Mood, Negative Mood, Activity Level, Sustained Attention, Positive Engagement
<b>IQ</b>					
Intelligence	Wechsler, 2012	internal consistency .95 test-retest reliability .86-.92	3		Measured by WPPSI-IV
<b>Brain Function</b>					
Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004	n/a	3	1	Measured by EEG
Language-Related Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2005	n/a		3	Measured by electroencephalogram
<b>Health: BMI</b>					
Body Mass Index (BMI)	Kuczumski, 2000	n/a	3		Measured by CDC scales
<b>Health: Physiological Stress</b>					
Physiological Stress	Ursache et al., 2017; Meyer et al., 2014; Davenport et al., 2006	n/a		2	Measured by hair cortisol
<b>Health: Sleep</b>					
Sleep problems	Yu et al., 2012	reliability .9	3	1,2	Measured by PROMIS Sleep Disturbance- Short Form adapted from ECHO; Additive index of the following items: 1. trouble falling asleep (0: never; 1: 1-2 nights; 2: 3-6 nights; 3: every night) 2. sleeping through night (reverse coded) 3. problem with sleep



**Appendix Table 1: Original Child Focused Pre-registered Hypotheses**

Measure description	Measure source	Psychometrics	Wave preregistered Primary Outcome	Wave preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Health: Other Indicators</b>					
<b>Overall Health, Medical Care, Diagnosis of Condition or Disability</b>	Halim 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
<b>School Achievement &amp; Behavior</b>					
<b>School test scores for target children and siblings</b>	Administrative data	n/a	School age (target child)	School age (siblings)	
<b>Student behavioral data for target children and siblings</b>	Administrative data	n/a		School age (target child and siblings)	

**Appendix Table 2: Original Maternal and Family Focused Pre-registered Hypotheses**

Measure description	Measure/Item source	Psychometrics	Wave preregistered Primary Outcome	Wave preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Household Economic Hardship</b>					
<b>Index of economic stress</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Additive index of dichotomized 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; 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; 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). 9. needed medical or dental care and did not get it? (0: No; 1: Yes) *changes to "in the past 12 months" for surveys at ages 24 and 36 months
<b>Household Poverty rate</b>	US Census Bureau			1,2,3	Measured using the Census Bureau's poverty thresholds by size of family and number of children
<b>Index of food insufficiency</b>	Economic Research Service, USDA, 2012			1, 2, 3	Additive index of 5 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 meals (0: No; 1: Yes) 4. Eat less than should (0:No; 1: Yes) 5. Hungry (0:No; 1: Yes)
<b>Social Services Receipt</b>					
<b>Number of Benefits received by mother</b>	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 10. LIHEAP / heat/AC assistance
<b>Mother's Labor Market and Education Participation</b>					
<b>Time to labor market reentry from birth</b>	Current Population Survey			1, 2, 3	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? *changes to "in the past 12 months" for surveys at ages 24 and 36 months
<b>Time to full-time labor market reentry from birth</b>	Current Population Survey			1, 2, 3	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? *changes to "in the past 12 months" for surveys at ages 24 and 36 months
<b>Mother's education and training attainment</b>	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 24 and 36 months **will be adding new items for surveys at age 24 and 36 months
<b>Child-Focused Expenditures</b>					
<b>Index of child-focused expenditures (since birth)</b>	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)?
<b>Index of child-focused expenditures (in past 30 days)</b>	Lugo-Gil, Yoshikowa, 2006			1, 2, 3	Continuous dollar amount. Past 30 days, total \$ amount spent on... 9. books 10. toys 11. clothes 12. diapers 13. videos* *products will be adjusted for child age at the age 24 and 36 month data collections
<b>Cost of paid child care</b>	National Study of Early Care and Education			1, 2, 3	Out of pocket spending on child care last week. 1. altogether, about how much money did you spend out-of-pocket on all of [CHILDNAME]'s child care arrangements last week?
<b>Use of center-based care</b>	National Study of Early Care and Education			1, 2, 3	1. Has child spent any time in childcare or day care? (Y/N)
<b>Housing and Neighborhoods</b>					
<b>Index of perceptions of neighborhood safety</b>	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)
<b>Index of housing quality</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	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)
<b>Homelessness</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Additive index of two dichotomized items (higher score=more homelessness): 1. Since child's birth, have you been homeless? * (0: Yes; 1: No) 2. Since birth, have you been in a group shelter? * (0: Yes; 1: No) *changes to "in the past 12 months" for surveys at ages 24 and 36 months
<b>Excessive Residential mobility</b>	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 24 and 36 months
<b>Neighborhood poverty</b>	Census			1, 2, 3	# of residents below poverty line in census tract divided by total number of residents in census tract
<b>Family and Maternal Perceived Stress</b>					

Appendix Table 2: Original Maternal and Family Focused Pre-registered Hypotheses

Measure description	Measure/Item source	Psychometrics	Wave preregistered Primary Outcome	Wave preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Perceived stress:</b> Perceived Stress Scale (PSS)	Cohen et al., 1994, 1983	alpha: .86		1, 2, 3	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
<b>Parenting stress:</b> Aggravation in Parenting Scale	PSID-Child Development Supplement	alpha: .71		1, 2, 3	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)
<b>Maternal Happiness and Optimism</b>					
<b>Global happiness</b>	The General Social Survey from NORC			1, 2, 3	One-item with 3-point response scale "Taken altogether, how happy are you these adys?" (0: not happy; 1: pretty happy; 2: very happy)
<b>Maternal Agency:</b> HOPE scale	Snyder et al., 1991	alpha: .86 test-retest: .81		1, 2, 3	Additive index of 9 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 5. ways to get what's important 6. solves problems 7. past has prepared me for future 8. pretty successful in life 9. meets goals set for oneself
<b>Maternal Physiological Stress</b>					
<b>Maternal hair cortisol:</b> from sample of hair that is at least 15mg and ~3cm long	Ursache et al., 2017		2	1	Measured using a sample of hair that is >=15mg 1 weight and ~3cm long; analyzed with sensitive and specific enzyme-linked immunosorbent assay; assay readout converted to pg cortisol per mg dry hair weight
<b>Maternal Mental Resources</b>					
<b>Maternal cognitive resources:</b> Flanker Inhibitory Control and Attention Test	Zelazo et al., 2013	test-retest: .92		2	Additive score of two outcome vectors (accuracy and response time)
<b>Maternal Mental Health</b>					
<b>Index of maternal depression:</b> PHQ-8	Kroenke & Spitzer, 2002			1, 2, 3	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
<b>Index of maternal anxiety:</b> Beck Anxiety Inventory	Steer & Beck, 1997	alpha: .92 test-retest: .75		1, 2, 3	Additive index of 21 common anxiety symptom items (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome)
<b>Maternal Substance abuse</b>					
<b>Alcohol and cigarette use</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 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?
<b>Opioid use</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Number of times per week used opioids?
<b>Chaos in Home</b>					
<b>Index of chaos in the home:</b> Home Environment Chaos Scale	Evans et al., 2005	alpha: .77 test-retest: .93		1, 2, 3	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 w/o 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)
<b>Maternal Relationships</b>					

**Appendix Table 2: Original Maternal and Family Focused Pre-registered Hypotheses**

Measure description	Measure/Item source	Psychometrics	Wave preregistered Primary Outcome	Wave preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Physical Abuse</b>	Fragile Families and Child Wellbeing Study			1,2,3	1. Ever abused? (1: yes; 0: no)
<b>Frequency of Arguing</b>	Fragile Families and Child Wellbeing Study			1,2,3	1. How often argue about things that are important to you? (1: never; 2: almost never; 2: sometimes; 3: fairly often; 4: very often)
<b>Relationship quality</b>	Fragile Families and Child Wellbeing Study			1, 2, 3	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)
<b>Maternal Physical Health</b>					
<b>Global health</b>	Idler & Benyamini, 1997			1, 2, 3	One item with 5-point response scale "overall, how would you describe your health..." (0: excellent-5: poor)
<b>Sleep</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 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)
<b>Mother's BMI</b>	CDC scales			3	Calculated by dividing weight by stature
<b>Parent-Child Interaction Quality</b>					
<b>Adult word count: LENA</b>	Xu et al (2009), LENA foundation			1, 2	
<b>Conversational turns: LENA</b>	Xu et al (2009), LENA foundation			1, 2	
<b>Index of mother's positive parenting behaviors</b>	Roggman, et al., 2013; Griffen & Friedman, 2007; Belsky, et al., 2007	inter-rater reliability varies by domain: .69-.80; alpha: .78	2	1	Measured using PICCOLO coding of parenting behaviors from three sub-scales (affection, responsiveness, encouragement and teaching) with responses ranging from 0: absent, 1: barely, 2: clearly
<b>Maternal Epigenetic Age</b>					
<b>Epigenetic age</b>	Fiorito et al., 2017			2	Measured by the Horvath Method
<b>Maternal DNA Methylation</b>					
<b>DNA methylation</b>	Hughes, et al., 2018; Cao-Lei et al., 2014			2	
<b>Frequency of Parent Child Activity</b>					
<b>Self-Report of Parent-child activities</b>	Rodriguez & Tamis-LeMonda, 2011			1, 2	Additive index of 4 items with response scale (lower score=higher frequency of activities): 1. read books (0: everyday; 1: a few times/week; 2: a few times/month; 4: rarely or never) 2. tell stories 3. play together 4. play groups
<b>Maternal Discipline</b>					
<b>Spanking discipline strategy: one-item, yes or no question</b>	Reichman et al., 2001			1, 2	1. In past month, have you spanked child due to misbehavior (1: yes; 2:no)

Measure description	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Language Development</b>					
Language Milestones	Squires et al., 2009	sensitivity .86 specificity .85		1, 3	Measured using ASQ- Communication Subscale
Child Vocalization^	Xu et al., 2009	internal consistency .69-.86 test-retest reliability .76		2	Measured using LENA processing software
Verbal Comprehension	Wechsler, 2012	internal consistency .95 test-retest reliability .86-.92		3	Measured by WPPSI-IV- Vocabulary subtest
Language Processing	Golinkoff et al., 2017		3		Measured by Quick Interactive Language Screener (QUILS)- Language Processing Subscale
Communicative Development	Fenson, 2002	internal consistency .85		2	Measured by short-form versions of the MacArthur Communicative Development Inventories
<b>Executive Function &amp; Self-Regulation</b>					
Self-Regulation	Smith-Donald et al., 2007	internal consistency of assessor report (not full assessment) .82-.93	3		Measured by the Preschool Self-Regulation Assessment - PSRA (parent report and child assessment)
Executive Function	Carlson, 2017 OR Wechsler, 2012 OR Willoughby et al., 2011	MEFS: validity .92 test-retest .93 WPPSI-IV: internal consistency .95 test-retest reliability .86-.92	3		Minnesota Executive Function Scale OR WPPSI-IV Working Memory OR EF Touch Executive Functioning
<b>Socio-Emotional Processing</b>					
Social-Emotional Problems	Briggs-Gowan et al., 2004	internal consistency .65-.79 test-retest reliability .87	3	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		Measured by the Child Behavior Checklist
Social-Emotional Behavior^	Roggman et al., 2013; Griffen & Friedman, 2007			1, 2	Measured using NICHD SECCYD parent-child-interaction task coding scheme, with child codes Positive Mood, Negative Mood, Activity Level, Sustained Attention, Positive Engagement
<b>IQ</b>					
Intelligence	Wechsler, 2012	internal consistency .95 test-retest reliability .86-.92	3		Measured by WPPSI-IV
<b>Brain Function</b>					
Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004	n/a	3	1	Measured by electroencephalogram
Language-Related Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2005	n/a		3	Measured by electroencephalogram
<b>Health: BMI</b>					
Body Mass Index (BMI)	Kuczmarski, 2000	n/a	3		Measured by CDC scales
<b>Health: Physiological Stress</b>					
Physiological Stress	Ursache et al., 2017; Meyer et al., 2014; Davenport et al., 2006	n/a		3	Measured by hair cortisol
<b>Health: Sleep</b>					
Sleep problems	Yu et al., 2012	reliability .9	3	1, 2	Measured by PROMIS Sleep Disturbance- Short Form adapted from ECHO; Additive index of the following items: 1. trouble falling asleep (0: never; 1: 1-2 nights; 2: 3-6 nights; 3: every night) 2. sleeping through night (reverse coded) 3. problem with sleep
<b>Health: Other Indicators</b>					
Overall Health, Medical Care, Diagnosis of Condition or Disability	Halim 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
<b>Child Epigenetic Age^</b>					
Epigenetic age	Fiorito et al., 2017	n/a		2	Measured by the Horvath Method
<b>Child DNA Methylation^</b>					
DNA methylation	Hughes et al., 2018; Cao-Lei et al., 2014	n/a		2	
<b>Child Nutrition</b>					
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
<b>School Achievement &amp; Behavior</b>					
School test scores for target children and siblings	Administrative data	n/a	School age (target child)	School age (siblings)	

**Appendix Table 3: Child Focused Pre-registered Hypotheses**

Measure description	Measure source	Psychometrics	Age preregistered <u>Primary</u> Outcome	Age preregistered <u>Secondary</u> Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Student behavioral data for target children and siblings</b>	Administrative data	n/a		School age (target child and siblings)	
<p><i>Notes.</i> 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.</p> <p>^ indicates outcomes that are contingent on in-person data collection at age 2.</p>					

Measure description	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Household Economic Hardship</b>					
<b>Index of economic stress</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	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; 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; 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)* 9. needed medical or dental care and did not get it? (0=no; 1=yes) *changes to "in the past 12 months" for surveys at ages 2 and 3
<b>Household Poverty rate</b>	US Census Bureau			1, 2, 3	Measured using the Census Bureau's poverty thresholds by size of family and number of children
<b>Index of food insufficiency</b>	Economic Research Service, USDA, 2012			1, 2, 3	Additive index of 5 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 meals (0: No; 1: Yes) 4. Eat less than should (0:No; 1: Yes) 5. Hungry* (0:No; 1: Yes)
<b>Social Services Receipt</b>					
<b>Number of Benefits received by mother</b>	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 10. LIHEAP/ heat/AC assistance
<b>Mother's Labor Market and Education Participation</b>					
<b>Time to labor market reentry from birth</b>	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?
<b>Time to full-time labor market reentry from birth</b>	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?
<b>Mother's education and training attainment</b>	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
<b>Child-Focused Expenditures</b>					
<b>Index of child-focused expenditures (since birth)</b>	Lugo-Gil, Yoshikawa, 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)?
<b>Index of child-focused expenditures (in past 30 days)</b>	Lugo-Gil, Yoshikawa, 2006			1, 2, 3	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 age 2 *products will be adjusted for child age at age 3 data collection
<b>Cost of paid child care</b>	National Study of Early Care and Education			1, 2, 3	Out of pocket spending on child care last week. 1. altogether, about how much money did you spend out-of-pocket on all of [CHILDDNAMEFF]'s child care arrangements last week?
<b>Use of center-based care</b>	National Study of Early Care and Education			1	1. Has child spent any time in childcare or day care? (Y/N)
				2, 3	1. Has child spent 5 or more hours in a child care or day care center last week? (Y/N)
<b>Housing and Neighborhoods</b>					
<b>Index of perceptions of neighborhood safety</b>	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)
<b>Index of housing quality</b>	MTO; Kling, Liebman, Katz, 2007			1, 3	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 8 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. overall condition (3: excellent, 2: good 1: fair, 0: poor)
<b>Homelessness</b>	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Additive index of two dichotomized items (higher score=more homelessness): 1. Since child's birth, have you been homeless?* (0: Yes; 1: No) 2. Since birth, have you been in a group shelter?* (0: Yes; 1: No) *changes to "in the past 12 months" for surveys at ages 2 and 3
<b>Excessive Residential mobility</b>	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
<b>Neighborhood poverty</b>	Census			1, 2, 3	# of residents below poverty line in census tract divided by total number of residents in census tract
<b>Family and Maternal Perceived Stress</b>					

Measure description	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Perceived stress</b>	Cohen et al., 1994, 1983	alpha: .86		1, 2, 3	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
<b>Parenting stress</b>	PSID-Child Development Supplement	alpha: .71		1, 2, 3	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)
<b>Maternal Happiness and Optimism</b>					
<b>Global happiness</b>	The General Social Survey from NORC			1, 2, 3	One-item with 3-point response scale "Taken altogether, how happy are you these adys?" (0: not happy; 1: pretty happy; 2: very happy)
<b>Maternal Agency</b>	Snyder et al., 1991	alpha: .86 test-retest: .81		1, 2, 3	HOPE Scale: additive index of 9 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 5. ways to get what's important 6. solves problems 7. past has prepared me for future 8. pretty successful in life 9. meets goals set for oneself
<b>Maternal Physiological Stress</b>					
<b>Maternal hair cortisol</b>	Ursache et al., 2017		3	1	Measured using a sample of hair that is >=15mg in weight and ~3cm long; analyzed with sensitive and specific enzyme-linked immunosorbent assay; assay readout converted to pg cortisol per mg dry hair weight
<b>Maternal Mental Resources<sup>^</sup></b>					
<b>Maternal cognitive resources</b>	Zelazo et al., 2013	test-retest: .92		2	Flanker Inhibitory Control and Attention Test: additive score of two outcome vectors (accuracy and response time)
<b>Maternal Mental Health</b>					
<b>Index of maternal depression</b>	Kroenke & Spitzer, 2002			1, 2, 3	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
<b>Index of maternal anxiety</b>	Steer & Beck, 1997 Spitzer et al., 2006	alpha: .92 test-retest: .75 alpha: .92 test-retest: .83		1, 3 2	Beck Anxiety Inventory: additive index of 21 common anxiety symptom items (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome) GAD-7: additive index of 7 items (0: not at all; 1: several days; 2: more than half the days; 3: nearly every day)
<b>Maternal Substance abuse<sup>x</sup></b>					
<b>Alcohol and cigarette use</b>	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?
<b>Opioid use</b>	MTO; Kling, Liebman, Katz, 2007			1, 3	Number of times per week used opioids?
<b>Chaos in Home</b>					
<b>Index of chaos in the home</b>	Evans et al., 2005	alpha: .77 test-retest: .93		1, 2, 3	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 w/o 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)
<b>Maternal Relationships<sup>*</sup></b>					
<b>Physical Abuse</b>	Fragile Families and Child Wellbeing Study			1,2,3	1. Ever abused? (1: yes; 0: no)
<b>Frequency of Arguing</b>	Fragile Families and Child Wellbeing Study			1,2,3	1. How often argue about things that are important to you? (1: never; 2: almost never; 2: sometimes; 3: fairly often; 4: very often)



Measure description	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Relevant Items (All measures between grey lines will be subject to multiple testing adjustments)
<b>Relationship quality</b>	MTO; Kling, Liebman, Katz, 2007			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 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).
<b>Maternal Physical Health</b>					
<b>Global health</b>	Idler & Benyamini, 1997			1, 2, 3	One item with 5-point response scale "overall, how would you describe your health..." (0: excellent-5: poor)
<b>Sleep</b>	MTO; Kling, Liebman, Katz, 2007			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)
<b>Mother's BMI</b>	CDC scales			3	Calculated by dividing weight by stature
<b>Parent-Child Interaction Quality<sup>^</sup></b>					
<b>Adult word count</b>	Xu et al (2009), LENA foundation			1, 2	Measured using LENA processing software
<b>Conversational turns</b>	Xu et al (2009), LENA foundation			1, 2	Measured using LENA processing software
<b>Index of mother's positive parenting behaviors</b>	Roggman, et al., 2013; Griffen & Friedman, 2007; Belsky, et al., 2007	inter-rater reliability varies by domain: .69-.80; alpha: .78	2	1	Measured using PICCOLO coding of parenting behaviors from three sub-scales (affection, responsiveness, encouragement and teaching) with responses ranging from 0: absent, 1: barely, 2: clearly
<b>Maternal Epigenetic Age<sup>^</sup></b>					
<b>Epigenetic age</b>	Fiorito et al., 2017			2	Measured by the Horvath Method
<b>Maternal DNA Methylation<sup>^</sup></b>					
<b>DNA methylation</b>	Hughes, et al., 2018; Cao-Lei et al., 2014			2	
<b>Frequency of Parent Child Activity</b>					
<b>Self-Report of Parent-child activities</b>	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	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 5. play pretend games
<b>Maternal Discipline<sup>x</sup></b>					
<b>Spanking discipline strategy</b>	Reichman et al., 2001			1, 2	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.  
<sup>+</sup> indicates that items were omitted or programmed incorrectly in the age survey administered to mothers and cannot be used to calculate outcomes. These include item 5 from the index of food insufficiency ("hungry"), and item 11 from the relationship quality index ("partner threatened or hurt your child/children?").  
<sup>^</sup> indicates outcomes that are contingent on in-person data collection at age 2.  
<sup>x</sup> indicates outcomes that were not administered at age 1 once in-person interviews switched to phone interviews due to COVID-19.

**ASQ**

Squires, J., Bricker, D. D., & Twombly, E. (2009). *Ages & stages questionnaires*. Baltimore, MD: Paul H. Brookes.

**LENA**

Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in young children's natural home environment. LENA Foundation.

**WPPSI-IV**

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence—fourth edition*. San Antonio, TX: The Psychological Corporation.

**QUILS**

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

**MACARTHUR**

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.

**PSRA**

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22 (2), 173-187.

**MINNESOTA EXECUTIVE FUNCTION SCALE**

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.

### EF TOUCH

Willoughby, M., & Blair, C. (2011). Test–retest reliability of a new executive function battery for use in early childhood. *Child Neuropsychology*, 17 (6), 564–579. doi:10.1080/09297049.2011.554390

Willoughby, M. T., Blair, C. B., & Family Life Project Investigators. (2015). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment*.

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & Family Life Project Investigators. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, 22(2), 306–317.

### BITSEA

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.

### CBCL

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.

### PICCOLO

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.

Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V., Christiansen, K., & Anderson, S. (2013). Parenting interactions with children: Checklist of observations linked to outcomes: PICCOLO user's guide. Baltimore, MD: Brookes.

### NICHD SECCYD Mother-Child Interaction Task (adapted script) and coding of Child Scales

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.

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.

**EEG/ERP**

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 Marshall, P. J., Fox, N. A., & Group, B. C. (2004). A comparison of the electroencephalogram

**WELL-BABY VISITS, ILLNESS, DIAGNOSIS AND MEDICATIONS, INJURIES, ER VISITS**

MetroBaby <https://steinhardt.nyu.edu/crcde/projects/childhood>

**BMI**

Kuczmariski, R. J. (2000). CDC growth charts; United States.

**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.

**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.

Scoring manual:

[http://www.healthmeasures.net/administrator/components/com\\_instruments/uploads/PROMIS%20Sleep%20Disturbance%20Scoring%20Manual.pdf](http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20Sleep%20Disturbance%20Scoring%20Manual.pdf)

**Child Health: Well Baby Visits, Illness, Diagnosis and Medications, Injuries, ER Visits**

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.

### **Child Epigenetic Age**

Fiorito, G., Polidoro, S., Dugue, P-A., Kivimaki, M., Ponzi, E., Matullo, G., Guarrera, S., Assumma, M.B., Georgiadis, P., Kyrtopoulos, S.A., Krogh, V., Palli, D., Panico, S., Sacerdota, C., Tumino, R., Chadeau-Hyam, M., Stringhini, S., Severi, G., Hodge, A.M., Giles, G.G., Marioni, R., Karlsson Linner, R., O'Halloran, A.M., Kenny, R.A., Layte, R., Baglietto, L., Robinson, O., McCrory, C., Milne, R.L., Vineis, P. (2017). Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Nature, 7*(16266), doi: 10.1038/s41598-017-16391-5.

### **Child DNA Methylation**

Hughes, A., Smart, M., Gorrie-Stone, T., Hannon, E., Mill, J., Bao, Y., Burrage, J., Schalkwyk, L., Kumari, M. (2018). Socioeconomic position and DNA methylation age acceleration across the life course. *American Journal of Epidemiology, 187*(11), doi: 10.1093/aje/kwy155.

Cao-Lei, L., Massart, R., Suderman, M.J., Machnes, Z., Elgbeili, G., Laplante, D.P., Szyf, M., King, S. (2014). DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project ice storm. *PLOS ONE*, <https://doi.org/10.1371/journal.pone.0107653>.

### **Child nutrition**

Los Angeles County WIC Survey. (2017). Retrieval from: <http://lawicdata.org/wp->

Measure description	Bibliography	
Preregistered measures	Source 1	Source 2
<b>Household Economic Hardship</b>		
<b>Index of economic stress</b>	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. <i>Econometrica</i> , 75(1), 83-119.	
<b>Index of food insufficiency</b>	<a href="https://www.ers.usda.gov/media/8282/short2012.pdf">https://www.ers.usda.gov/media/8282/short2012.pdf</a>	
<b>Household poverty rate</b>	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.	
<b>Social Services Receipt</b>		
<b>Number of Benefits received by mother</b>	Study PIs	
<b>Mother's Labor Market and Education Participation</b>		
<b>Time to labor market reentry from birth</b>	Current Population Survey, retrieved from:	
<b>Time to full-time labor market reentry from birth</b>	<a href="https://www.census.gov/programs-surveys/cps/technical-documentation/questionnaires.html">https://www.census.gov/programs-surveys/cps/technical-documentation/questionnaires.html</a>	
<b>Mother's education and training attainment</b>		
<b>Child-Focused Expenditures</b>		
<b>Index of child-focused expenditures</b>	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.	
<b>Child-focused expenditures</b>		
<b>Cost of paid child care</b>	National Study of Early Care and Education	
<b>Use of center-based care</b>		
<b>Housing and Neighborhoods</b>		
<b>Index of perceptions of neighborhood safety</b>		
<b>Index of housing quality</b>	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. <i>Econometrica</i> , 75(1), 83-119.	
<b>Residential mobility</b>		
<b>Homelessness</b>		
<b>Neighborhood poverty</b>		
<b>Family and Maternal Perceived Stress</b>		

- 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** [https://psidonline.isr.umich.edu/cds/cdsi\\_usergd.pdf](https://psidonline.isr.umich.edu/cds/cdsi_usergd.pdf)
- 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 validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, 60(4), 570-585.
- Maternal Epigenetic Age**
- Epigenetic age** Fiorito, G., Polidoro, S., Dugue, P-A., Kivimaki, M., Ponzi, E., Matullo, G., Guarra, S., Assumma, M.B., Georgiadis, P., Kyrtopoulos, S.A., Krogh, V., Palli, D., Panico, S., Sacerdota, C., Tumino, R., Chadeau-Hyam, M., Stringhini, S., Severi, G., Hodge, A.M., Giles, G.G., Marioni, R., Karlsson Linner, R., O'Halloran, A.M., Kenny, R.A., Layte, R., Baglietto, L., Robinson, O., McCrory, C., Milne, R.L., Vineis, P. (2017). Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Nature*, 7(16266), doi: 10.1038/s41598-017-16391-5.
- Maternal DNA Methylation**
- DNA methylation** Hughes, A., Smart, M., Gorrie-Stone, T., Hannon, E., Mill, J., Bao, Y., Burrage, J., Schalkwyk, L., Kumari, M. (2018). Socioeconomic position and DNA methylation age acceleration across the life course. *American Journal of Epidemiology*, 187(11), doi: 10.1093/aje/kwy155.  
Cao-Lei, L., Massart, R., Suderman, M.J., Machnes, Z., Elgbeili, G., Laplante, D.P., Szyf, M., King, S. (2014). DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project ice storm. *PLOS ONE*, <https://doi.org/10.1371/journal.pone.0107653>.
- 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** Zelazo, P.D., Anderson, J.E., Richler, J., Wallner-Allen, K., Beaumont, J.L., & Weintraub, S. (2013). II NIH Toolbox Cognitive Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, 78(4), 16-33.

### 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** Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.

**Mother's BMI** Kuczmariski, 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

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](https://fragilefamilies.princeton.edu/sites/fragilefamilies/files/year_3_guide.pdf#page=84)

### Frequency of Arguing 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., Innocenti, 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.

**Maternal Discipline****Spanking discipline strategy**

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