NBER WORKING PAPER SERIES

CASH TRANSFERS FOR CHILD DEVELOPMENT: EXPERIMENTAL EVIDENCE FROM INDIA

Jeffrey Weaver
Sandip Sukhtankar
Paul Niehaus
Karthik Muralidharan

Working Paper 32093
http://www.nber.org/papers/w32093

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2024

We thank the Government of Jharkhand State Nutrition Mission for their support of this project, and are particularly grateful to Mukhmeet Bhatia, Dr. DK Saxena, and Ajay Verma. We thank Radu Ban, Prashant Bharadwaj, Hoyt Bleakley, and Tom Vogl and numerous seminar and conference participants for helpful comments and suggestions. Hannah Blackburn, Thomas Brailey, Soala Ekine, Burak Eskici, Stefan Faridani, Simoni Jain, Pankhuri Jha, Wilson King, Prerna Kundu, Xinyi Liu, Frances Lu, Avantika Prabhakar, Sabareesh Ramachandran, Akanksha Saletore, Mayank Sharma, and Prashansa Srivastava provided excellent research assistance and field support. This study was registered with the AEA Trial Registry (RCT ID: 0002899). It was approved by the Institutional Review Boards of the University of California, San Diego and the Institute for Financial Management and Research. We thank the Bill and Melinda Gates Foundation and the Cash Transfers for Child Health Initiative at the Abdul Latif Jameel Poverty Action Lab for funding this study. All views expressed are those of the authors and not any of the funders or institutions with which they are affiliated. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32093

Cash Transfers for Child Development: Experimental Evidence from India
Jeffrey Weaver, Sandip Sukhtankar, Paul Niehaus, and Karthik Muralidharan
NBER Working Paper No. 32093
January 2024
JEL No. D13,I12,I15,I38,O15,Q53

ABSTRACT

Despite significant economic growth, child development outcomes in India remain poor. Using a large-scale experiment in which randomly-selected mothers receive cash transfers for the first two years of their child's life, we examine the relationship between income and child development in the Indian state of Jharkhand. Treated mothers and children experienced large increases in nutritional intake, including increases in caloric consumption of 9.6-15.5%. However, child anthropometric indicators improved only in areas with low rates of open defecation. These results suggest that poor sanitation is a key explanatory factor for the poor translation of increases in income into child growth in India.

Jeffrey Weaver
University of Southern California
Department of Economics
3620 South Vermont Ave.
Kaprielian (KAP) Hall, 300
Los Angeles, CA 90089
jeffrey.b.weaver@gmail.com

Sandip Sukhtankar
Department of Economics
University of Virginia
Charlottesville, VA 22904
and NBER
srs8yk@virginia.edu

Paul Niehaus
Department of Economics
University of California, San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
pniehaus@ucsd.edu

Karthik Muralidharan
Department of Economics, 0508
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
kamurali@ucsd.edu

A randomized controlled trials registry entry is available at https://www.socialscienceregistry.org/trials/2899

Reducing child malnutrition is a critical global development goal that is both intrinsically important for the well-being of affected children, and instrumentally important for boosting long-term human capital and incomes of low- and middle-income countries. Reflecting this view, the UN Sustainable Development Goals seek to end child malnutrition by 2030. However, any strategy for achieving this goal must address India, home to 24.6% of all stunted children globally (World Health Organization, 2023). Thirty-six percent of Indian children under 5 are stunted as of 2019 (Ministry of Health and Family Welfare, 2022a), and rates of this and other measures of malnutrition have remained persistently higher than those in many *poorer* sub-Saharan African countries (Ramalingaswami et al., 1996; Klasen, 2008; Jayachandran and Pande, 2017; World Health Organization, 2023). The World Bank has referred to child malnutrition as India's "silent emergency, and recent projections suggest that India's stunting rate is likely to remain above 22% even in 2047."[1]

These stubbornly high rates pose a puzzle. India's economy has grown rapidly and quintupled real per-capita incomes since 1984. Yet income growth has only had modest effects on reduced stunting, which is surprising both intuitively and in light of recent evidence that, in other settings, exogenous income increases *have* yielded meaningful anthropometric gains (Manley et al., 2012; Bastagli et al., 2016; Baird et al., 2019). Further, cash transfers to the poor are increasingly being considered as a policy option for tackling malnutrition, making it practically important to understand the causal relationship between income growth and improved child development.

Several explanations have been proposed for the poor translation of gains in mean per-capita income into child growth in India. These include: (a) gains in mean income may not translate into income gains for the poor, among whom malnutrition is concentrated; (b) low income-elasticity of calorie consumption (Deaton and Drèze, 2009); (c) low nutritional value of consumed food as people may select for taste over nutrition; (d) high intra-household inequality in food allocation (Jayachandran and Pande, 2017); and (e) environmental factors such as poor sanitation (Coffey and Spears, 2017), which might limit the benefits of consuming more or better food. But quantifying the relative importance of these channels has not been feasible with existing data, as it requires both exogenous variation in income and detailed measurement of the intermediate steps in the causal chain from income to child development.

To that end, this paper reports results from a large-scale randomized evaluation of a maternal cash transfer program in the state of Jharkhand—to our knowledge, the first experimental evaluation of an unconditional cash transfer at comparable scale anywhere in India. The study was conducted in a sample of 480 public early-childhood care centers (*anganwadi* centers, or AWCs) representative of 32 million people. AWCs provide health, nutrition, and developmental services to children and mothers as part of India's Integrated Child Development Services (ICDS) program.

For our study, AWCs registered pregnant women in their catchment area as eligible for the program. Half of the AWCs were then randomly selected, and the mothers registered at these AWCs were issued Rs. 12,000 (USD $170) in monthly installments of Rs. 500 over the course of

[1] See https://www.worldbank.org/en/news/feature/2013/05/13/helping-india-combat-persistently-high-rates-of-malnutrition for the quote, and Muralidharan (2024) for projections based on the low historical time-series income elasticity of stunting reduction in India.

two years, starting from birth. The transfer amount was equal to $\sim$10% of mean monthly household consumption in this setting, and was set so that one year of transfers would be similar to the value of India's flagship maternal cash transfer program (Pradhan Mantri Matru Vandana Yojana). These transfers were accompanied by light-touch messaging encouraging mothers to use the money to purchase nutritious food for themselves and their children, as is typical of this type of program.

Our primary focus is on child development outcomes measured approximately one year after the program ended, when children were around three years old. To interpret these effects, we also examine several intermediate outcomes measured during the course of the intervention. These include unusually detailed consumption data in which we measure the exact weight or volume of each ingredient for each item of food consumed in the day prior to the survey and translate this into precise measures of nutrient intake for targeted household members.

We report three main sets of results. First, transfers significantly increased mother and child nutritional intake. Total household food expenditure increased by over 15%, and the intended beneficiaries (mothers and target children) consumed at least a proportionate share of this additional food. Their caloric intake increased by 9% and 14% in the first and second years, respectively. The quality of their food consumption also improved, as measured by dietary diversity, a nutrient intake index, and consumption of specific key nutrients such as protein, visible fat, and iron.

Second, despite these improvements in diet, standard measures of children's anthropometrics (height- and weight-for age z-scores, or HAZ and WAZ) did *not* increase significantly on average. Our estimates are precise enough to rule out improvements of more than 0.15 standard deviations for each outcome. We find that this puzzling lack of effect on anthropometrics can be explained by accounting for sanitation. Comparing children living in better versus worse sanitation environments (measured primarily by neighborhood rates of open defecation), there are positive and significant treatment effects on anthropometrics only for the former, with this difference itself significant ($p = 0.02$). Young children in the top quartile of our sanitation index (i.e. lowest rates of open defecation) saw significant HAZ and WAZ gains of 0.13 and 0.12 standard deviations, respectively. We also show that this result cannot be explained by differences in nutritional intake across sanitation environments or by interactions of the treatment with observable correlates of sanitation.

Finally, nutritional gains were accompanied by improvements in measures of children's functional development. Using the standard Ages and Stages questionnaire to measure child progress on age-appropriate developmental milestones (Bricker et al., 1999), we find that the treatment improved outcomes at age 3 by 0.12 standard deviations ($p < 0.01$), with gains in parental reports of cognitive, gross motor, and fine motor skills. While not based on direct measurement, this is at least suggestive of broad-based gains in child development.

Our results make several contributions. First, they corroborate and strengthen the argument that poor sanitation is a key explanatory factor for India's poor child health outcomes. Building on work that shows how the *level* of sanitation affects child health (e.g., Esrey (1996); Checkley et al. (2004); Fink et al. (2011); Rah et al. (2015); Spears (2020); Cameron et al. (2022)), we demonstrate an *interaction* between income and sanitation in the child health production function.[2,3] We show that even when incomes of the poor go up, and all intermediate steps in the causal chain from income to nutritional intake are met, anthropometric outcomes of young children did not improve when the sanitation environment was poor. These results suggest that poor sanitation may be a critical binding constraint to the translation of income growth into child growth in India, and highlight the centrality of improving sanitation to accelerate this translation going forward.

Second, the key mediating role of sanitation in translating income growth into child growth helps to situate our results in the global literature on cash transfers to pregnant and nursing mothers. Several recent studies have found meaningful positive effects of such transfers on child anthropometrics in other low- and middle-income country settings including Levere et al. (2016) in Nepal, Field and Maffioli (2021) in Myanmar, and Carneiro et al. (2021) in Nigeria. However, these are all settings with much lower rates of open defecation (averaging under 20%) than rural Jharkhand, where the median community has a 35% rate of open defecation, and the rate is 57% in the 75th percentile community (Ministry of Health and Family Welfare, 2022a). We find positive effects on child growth in the top quartile of the Jharkhand community-level sanitation distribution where the open defecation rate is 19% (comparable to the settings in the studies above), but not in places where it is worse. Our results are thus both consistent with the global literature, and extend it by demonstrating a key environmental factor that mediates the effect of cash transfers.

Third, the fact that we find gains in mean child cognitive and motor skills but not anthropometrics highlights the importance of considering a broader set of child development outcomes. Several studies outside of India have documented encouraging impacts of early-childhood interventions on these measures of child development (Paxson and Schady, 2010; Macours et al., 2012; Gilligan and Roy, 2013), while work in India has focused primarily on anthropometrics. Our findings on cognitive skills suggest that both merit attention, especially in light of the possibility that "brains" may be more important than "brawn" for children's futures over time (Pitt et al., 2012).

We also contribute some of the first evidence on the causal impacts of unconditional cash transfers in India. Such transfers remain limited in India, despite world-leading infrastructure for making and using electronic cash transfers (Sukhtankar, 2023). This is in part due to the entrenched idea that the poor would use these poorly (Khera, 2014), and a lack of well-powered experimental evidence to contradict this idea from the Indian context (Niehaus and Suri, 2024).[4] Our results suggest that fears of recipients wasting cash transfers are unfounded.

Finally, our results are directly relevant for policy. India's policy response to the problem of

[2]The technically most rigorous way of demonstrating this interaction would be to run a 2x2 experiment that randomly varies both sanitation and income. However, we show that this design would likely be infeasible in practice since it would be highly underpowered to detect interactions under plausible sample sizes (see appendix B). As a result, our approach of leveraging random variation in income and cross-sectional variation in sanitation is likely to be the most practically feasible approach for generating evidence on this interaction.

[3]The most related analysis is Geruso and Spears (2018), who find that sanitation interacts with how breastfeeding is related to child mortality. However, neither breastfeeding nor sanitation are experimentally varied, and mortality is a different margin of child health.

[4]Davala et al. (2015) experimentally show that cash transfers in India were spent well, but in a pilot conducted in only 20 clusters (8 treatment, and 12 control villages in the state of Madhya Pradesh).

stunting has focused on supplemental nutrition programs for children, with policy debates emerging on whether cash transfers may be more effective (Narayanan and Saha, 2020). However, our findings suggest that independent of the form of transfers to households, simultaneous investment is needed in public goods like sanitation. The policy importance of focusing on sanitation is magnified by the fact that poor households will under-invest in sanitation relative to socially optimal levels due to both spillovers and fixed costs. We discuss these issues further in the conclusion.

1 Context & intervention

1.1 The ICDS program

The prevalence of stunted and underweight children in India is among the highest in the world, and "has its origins almost entirely during the first two to three years of life" (World Bank, 2009). The Indian government created the Integrated Child Development Services (ICDS) program in 1975 to address the health, nutrition and developmental needs of children and mothers. ICDS services are provided through 1.35 million anganwadi centers (AWCs), where each AWC serves a catchment area of 500 to 1000 people and is typically staffed by one anganwadi worker (AWW) and one helper. The AWW oversees the wide range of services provided at the AWC, including supplementary nutrition programs, community health education, and immunization.

To address nutrition for pregnant women, lactating mothers, and children aged 6 months to 3 years, the status-quo approach in many states has been for AWCs to provide "take-home rations": ready-to-eat packets of flour fortified with vitamins and protein. These are substantial, providing over 15,000 calories and 480 grams of protein each month, often in areas where such nutritionally-fortified food is otherwise not easily accessible. However, doubts have been raised about the value of this program—based for example on journalistic accounts of households giving the food to their animals rather than eating it themselves (Malik, 2016). Reflecting these doubts, India's central government has introduced cash transfers as an additional instrument through its Pradhan Mantri Matru Vandana Yojana (PMMVY) scheme, which was launched in 2017 and provides transfers of Rs. 5,000 to first-time pregnant and lactating mothers.

How much cash transfers will impact child development is an open question, mirroring open questions about the translation of income growth into child development in India more generally (Jayachandran and Pande, 2017; Spears et al., 2022). The answer depends (among other things) on the extent to which households spend additional income on food, the nutritional content of said food, and which household members consume it. Further, the extent to which both income growth and income-transfer programs translate into child development may also depend on contextual and environmental factors mediating the nutrition-development relationship such as genetics and sanitation.

1.2 The cash transfer intervention

Our study is set in the north Indian state of Jharkhand, which was recently ranked as the second poorest state in the country on a multi-dimensional poverty index (NITI Aayog, 2021). Child development outcomes are poor: 39.6% of children are stunted, and 39.4% are underweight (Ministry of Health and Family Welfare, 2022b). Among women aged 15-49 years, 65.7% are anaemic, and 26.2% are classified as underweight on the Body Mass Index scale (Ibid).

Seeking to reduce these stubbornly high figures, the Government of Jharkhand (GoJH) decided to test the impact of cash transfers to new mothers. For a period of 3 months, AWWs in the 480 sampled AWCs informed pregnant women in their 1st and 2nd trimesters in their catchment areas that they were eligible to register to receive cash transfers. There were no further ex ante eligibility requirements and no ex post actions upon which transfers were conditional. To register, women were asked to fill out a one-page form and supply their bank account information and a photocopy of their personal identification card. Assistance in registering for bank accounts was provided to women who did not currently have one, including through organized registration "camps."

Eligible women were informed that half of AWCs would be selected, and that if their AWC was selected they would receive monthly cash transfers of Rs. 500 (approximately US$7) for one year.[5] The resulting amount (Rs. 6000 annually) was selected to be similar in total to related government cash transfer programs (PMMVY and Janani Suraksha Yojana). The randomization was done *after* registration to ensure that the universe of eligible women who had registered for the program was comparable across treatment and control groups.

GoJH began issuing transfers in the treatment group approximately four months after the close of registration. As a result, most beneficiaries began receiving transfers around the time of the child's birth, although some children were slightly older or younger.[6] Transfers were made into the registered woman's personal bank account and were typically withdrawn through in-person visits to bank branches (50%) or in-village common service centers (37%).

Transfers were accompanied by messaging encouraging beneficiaries to spend them on nutritious food for the mother and child. This messaging was delivered at the time of registration through flyers and verbally by the AWWs, as well as through monthly automated (IVR) calls to treated individuals. These calls played a recorded message that informed recipients that a transfer had just been sent to their accounts and also provided suggestions on nutritious types of food that beneficiaries could purchase with it that were customized to their child's current age.[7] Slightly fewer than half of beneficiaries received calls each month, reflecting poor cellular connectivity in many

[5] At the start of the project, the government was uncertain whether funding would be available for a second year, and so only informed potential beneficiaries about one year of transfers at the time of registration. After one year, the government confirmed funding for another year, and the treatment group was informed that their transfers would continue for a second year. However, given the large cash amounts involved, the set of registrants would have likely been the same if the second year had been initially publicized.

[6] Appendix Table A.8 test for and finds no heterogeneity in effects with respect to child age at first transfer.

[7] Both the flyers and verbal information from AWWs were also delivered to the control group, so the only messaging difference was through the automated calls. We later present evidence that the calls were relatively ineffective and do not independently explain the observed treatment effects.

areas. Overall, the transfers should be seen as unconditional but "framed," albeit with much weaker framing and messaging around behavioral change than in recent studies examining the interaction of cash transfers and messaging (Levere et al., 2016; Field and Maffioli, 2021); those studies had at least monthly in-person meetings that provide messaging. This approach to messaging reflects those studies' findings that information can matter for the effects of cash transfers, but also the constraint that a less resource-intensive approach is needed for feasible implementation at scale in settings of weak state capacity.

Implementing cash transfers in this context posed many challenges. Transfers were often systematically delayed by government fiscal processes, arriving more than one month late several times.[8] Such delays are common in Indian government transfer programs, so these ITT effects are the appropriate ones for policy evaluation here.

For logistical reasons the government implemented the project in two phases. In five districts, registration began in October 2017 and transfers in March 2018. In the other three districts, these occurred approximately six months later. We pool both phases for analysis since procedures were the same and we do not observe significant differences in estimated treatment effects on primary outcomes. Figure A.1 provides a full timeline of project activities.

2 Experimental design & methods

Our design and methods follow a set of registered pre-analysis plans: one for data from year 1, one for data from year 2, and one that updated the analysis plans to account for how the COVID-19 pandemic shifted data collection.[9] This was part of a larger experiment that included an additional 480 AWCs allocated to two other treatment arms (receiving a year of transfers during either the first or second year of the child's life). This paper focuses on the two year treatment arm, while analysis of the other two arms is available in a separate report (Weaver et al., 2023).[10]

2.1 Experimental design

The study population was selected so as to be as close as possible to a representative sample of pregnant mothers in Jharkhand. We randomly selected 8 of the 24 districts of the state, although two districts were excluded from sampling since they already had a cash transfer program targeted at

[8]The April, May, June and July 2019 transfers were delayed until June 12, June 25th, August 1st and August 29th, respectively. November and December 2019 transfers were delayed until February 11 and February 20 of 2020. March and April 2020 transfers were delayed until April 21 and May 28 of 2020 as a result of COVID-19. In addition, approximately one-tenth of the sample experienced idiosyncratic delays beyond those described above at least once due to problems with their bank account or documentation. These issues likely would not have been resolved without the intervention of JPAL staff, highlighting the demands that even a relatively simple transfer can place on over-taxed state capacity.

[9]See https://www.socialscienceregistry.org/trials/2899

[10]During the periods of cash transfer receipt, the effect of the treatment on nutritional intake and other intermediate outcomes was similar in all three treatment arms. However, we only detect consistent effects on anthropometric and functional development outcomes in the two year treatment arm. This is likely because the longer treatment period generated a larger treatment effect that we are better powered to detect than any correspondingly smaller effects in the one year arms.

mothers. Within each district, AWCs are grouped into "sectors" of approximately 30 geographically proximate AWCs. Our sampling procedure first selected a random sample of sectors in each district, and then sampled six AWCs per sector for inclusion in the study. We sampled 10 sectors (60 AWCs) per district for a total of 480 AWCs in the experimental sample. The sample is broadly representative of Jharkhand: when these sectors are compared to the rest of the state on 26 variables from the 2011 census, there is only one statistically significant difference at the 5% level.

We randomly assigned half of the AWCs to receive two years of cash transfers and half to not receive any transfers, corresponding to 240 treatment and 240 control AWCs with approximately 1200 women in each group. The randomization was stratified by sector (with three AWCs each per sector into treatment and control groups), and further stratified within sectors based on the number of registrations in the AWC. Figure A.2 shows the geographical distribution of treatment and control AWCs. Compliance with the randomization was high, with no women in control AWCs and nearly all women in treatment AWCs receiving transfers. Administrative records state that over 99% of transfers were made successfully, and we see this reflected in survey data where treatment households were around 70% more likely than control households to report having withdrawn any money from their bank account in the month prior to the survey (Table A.3). Treatment households were also more likely to state that they withdrew money from their account for the specific purposes of daily expenses (typically food) and medical expenses, consistent with the goals of the policy.

Table A.1 tests for balance on a pre-registered set of characteristics, grouped into three broad categories: household-level characteristics, number of women who registered at the AWC, and village characteristics as measured by the 2011 census of India. There was no baseline survey since the target children had not yet been born, so we test household characteristics that are invariant over the study period. We focus on characteristics that may be relevant to the implementation of the program, such as poverty, sanitation environment, and difficulty of travel to pick up cash transfers from the bank. Across the 12 tests, none of the differences between treatment and control are statistically significant at the 5% level and the joint p-value is equal to 0.80. Table A.2 also finds no differential attrition across treatment and control groups in any round of the survey.

2.2 Data collection

We gathered multiple rounds of survey data over the three years after the start of transfers (Figure A.1 provides a timeline), with the data collection instrument changing across rounds to reflect the relevant stage of child development. We surveyed the mother who registered for the transfers, and collected anthropometric measurements of the mother, child, and other children in the household.

In the first year, we conducted a survey after 11 months of transfers (endline 1). The household component of these surveys measured spending on food; nutritional intake for the mother, child, and household; child morbidity and mortality; and maternal health knowledge, stress and depression, and empowerment. To measure nutritional intake, enumerators asked respondents about each meal eaten the previous day. For every dish prepared at home, the enumerator recorded the ingredients used and the exact weight or volume of each ingredient (see Appendix C for details).[11] By combining these intensive measurements with data on the nutritional content of each ingredient from the Indian Food Consumption Database (Longvah et al., 2017), we derive precise measures of calories, protein, and micronutrients (e.g., iron) consumed. The enumerator also measured the quantity of the final dish consumed by the mother and child, meaning that we can measure nutritional intake for those two household members as well as per capita consumption of the rest of the household (by dividing the remaining consumption by the remaining household members).

In the second year, we conducted an endline survey after 23 months of transfers to measure the same outcomes as in year 1 as well as the cognitive/motor development of the child and weight-for-age of one randomly selected sibling under age 10 (if any). The outbreak of the COVID-19 pandemic in March 2020 caused the suspension of field operations after 613 surveys had been completed across four districts (21% of the sample). We then collected data over the phone for households that had not been reached in-person. Instead of the more intensive method based on measuring ingredients, the phone survey instrument asked whether the household had consumed particular food items in the previous day; based on this, we can measure maternal and child dietary diversity (whether they had consumed foods from different categories) in the phone data, but not calories or micro-nutrients consumed. We also were not able to measure anthropometrics, maternal stress, or depression over the phone.[12]

From September to November 2021, we conducted a final round of in-person data collection in all eight study districts during a period of low COVID-19 prevalence in Jharkhand (less than 0.1 daily cases per 100,000). Since this was approximately 1.5 years after the last transfer in five of the districts and a year after the last transfer in three districts, these data measure the longer run effect of receiving the cash transfers. This survey measured height and weight for the target child, mother, and up to three randomly selected siblings under 10 years of age,[13] the neighborhood sanitary environment (e.g., prevalence of open defecation), and child cognitive and motor skills.

We use these data to investigate each link of the causal chain between income increases and child development. We examine impacts on pre-specified measures of household spending, quantity and quality of food consumption, and intrahousehold food consumption during the period when households were still receiving transfers (using the year 1 field survey and year 2 phone and field surveys). In interpreting these results, we also discuss measures of child morbidity, maternal outcomes, beliefs and health-seeking behavior, and interaction with government services.

Ultimately, we are interested in seeing whether changes along these links in the causal chain manifest in longer run physical and cognitive development of the child at age 3. For anthropometric measures of physical development, we measured height using stadiometers and weight using SECA-876 scales, the standard scale in the nutrition literature. We focus on child weight-for-age (WAZ) and height-for-age (HAZ), standardized into z-scores using growth charts from the World Health

[11] For food consumed away from home, we use estimates on the average nutritional content of those items.

[12] Table A.4 examines how likelihood of contact over the phone varies with respondent characteristics (as measured in the year 1 survey). 62% of households were contacted for at least one year 2 phone survey.

[13] This random sample accounted for 95.9% of all siblings in this age range.

Organization. These anthropometric measures are important measures of child development and have been shown to predict later economic productivity (Currie and Vogl, 2013).

We measure child cognitive and motor skills development in year 3 using questions from the Ages and Stages Questionnaire (ASQ) (Doyle, 2020). This questionnaire asks mothers a series of questions about their child's ability to perform tasks across five categories – communication, gross motor, fine motor, problem solving, and personal-social (e.g., "can your child count to 10?", "can your child unbutton one or more button on their own?"). Each question is assigned a maximum of ten points – with ten points for a response of "yes," five for "sometimes", and zero for "no/never" – and these are summed for our measure of non-anthropometric child development.

For clarity, Table A.5 summarizes the outcomes of interest, and when they were collected. Our main analysis of these outcomes focuses on data from year 1 and year 3, as the much smaller sample in the year 2 field data significantly reduces statistical power. Fortunately, there is a very high correlation between anthropometric outcomes in year 2 and year 3 (e.g., a correlation of 0.845 for weight-for-age among children measured in both rounds), so year 3 is likely a good approximation of what happened in year 2 as well as of interest in its own right. We also report results with respect to three pre-registered dimensions of heterogeneity—child birth order and gender (Behrman, 1988; Jayachandran and Kuziemko, 2011; Barcellos et al., 2014), which we specified we would report regardless of the results, and sanitation environment (Coffey and Spears, 2017), which we specified we would report if it meaningfully affected our interpretation of the other results.[14]

2.3 Analytical methods

We estimate the effects of receiving transfers for two years using the following specification:

$$y^t_{has} = \beta_0 + \beta treatment_a + \phi_s + \epsilon_{has} \tag{1}$$

where h indexes households, a indexes AWCs, s indexes sectors. In this specification y^t is an outcome measured in year t, *treatment* is a dummy variable indicating whether the AWC was assigned to the treatment arm, and ϕ_s represents sector-level fixed effects.[15] Given high rates of compliance with the experimental assignment, we focus on ITT estimates. Standard errors are clustered at the AWC level, the level at which treatment was assigned.

[14] See tables A.6, A.7, A.8, A.9, A.10, and A.11.

[15] While additional control variables are not necessary given random assignment, the PAP proposed including controls selected using the post-double-selection approach of Belloni et al. (2014). In practice, enough observations are missing for these characteristics that gains would be more than offset by losses in sample size and representativeness. We therefore prefer results estimated without controls, but results are similar either way (Table A.12).

3 Results

3.1 Effects on food consumption and nutrition

We first examine the extent to which an exogenous increase in income translates into spending on food and other categories. We apply the inverse hyperbolic sine (IHS) transformation to spending amounts, as there are some observations of zero (e.g. the household didn't go to the market over the recall period). For food expenditure, column 1 of Table 1 finds substantial and significant increases in both year 1 (15 IHS points) and year 2 (21 IHS points), which roughly correspond to 15% and 21% increases. We do not see corresponding increases in reported "sin good" expenditure in either year (column 4), although the number of observations is lower in year 2 since we did not measure this in the phone survey. For non-food expenditure (column 7), which we observe only in year 2, the estimated increase (19 IHS points) is similar to that for food expenditure (21 IHS points), and we cannot reject equality between the two ($p = 0.83$). The data are consistent, in other words, with homothetic preferences within the range of expenditures induced by the experiment (at this broad level of disaggregation).[16]

This result is important in light of widely-discussed expenditure trends documented for example by Deaton and Drèze (2009), who show that real per capita food expenditure in India was essentially flat from the 1987-88 round to the 2004-5 round of the NSS even as overall expenditure increased substantially. As they point out, this puzzling fact must surely be part of the explanation for the (equally puzzling) fact that caloric consumption fell over the same period. However, without well-identified estimates of the causal effect of income on food expenditure, it has been difficult to separate out preference-based explanations from other time-varying factors, such as a reduction in caloric requirements–which in turn could reflect factors like a reduction in physical labor intensive jobs or a reduction in disease burden. Our results provide the first experimental evidence that we are aware of that, all else equal, poor Indian households have a meaningfully positive elasticity of spending on food with respect to income—albeit with the important caveats that the income was accompanied by messaging encouraging this and transfers were made to the mother's bank account, which may increase the likelihood of usage for maternal and child nutrition (Field et al., 2021).

At the same time, not *all* of the incremental spending was on food. This may be acceptable or even desirable according to standard notions of welfare, but a policy-maker narrowly focused on nutrition might see it as a "tax" on the intervention. Quantitatively, a proportionate increase in both food and non-food spending (as we see in year 2) implies that that 36% of increased spending was on food (i.e. the control mean share). In year 1, we do not observe non-food spending[17] so we can gauge the marginal propensity to spend on food only from impacts on the *levels* (rather than

[16]Chen and Roth (2023) note challenges in interpreting outcomes transformed using IHS. We follow their recommendation and also report results along both extensive and intensive margins in the remaining table columns (Table 1, columns 2, 3, 5, 6, 8 and 9). Effects are mostly on the intensive margin for food spending, consistent with 97% of households reporting non-zero food spending in each round. Results are unchanged for sin good spending, and we lack power to decompose effects for non-food spending, which exhibits more dispersion and noise than food spending.

[17]This is due to an error in the year 1 survey form that caused this section to be skipped, which was not caught until the survey was completed.

IHS) of spending, which we estimate relatively imprecisely due to the significant dispersion in that measure. That said, this yields a reasonably similar point estimate of 25%, with a 95% confidence interval of $[-21\%, 72\%]$ (Table 1). Using either estimate, it is clear that a meaningful share of the transfers were not spent on food, but non-food spending might also contribute to the overall goal of child development. For example, despite being no more likely to fall ill (Table 3, column 6), treated households were substantially more likely to visit a formal medical provider (Table 3, column 5) and state that they withdrew money from their bank account for medical expenses (appendix table A.3), indicating that their children may be better cared for in cases of illness.

Table 2 shows that the increased food expenditure translated into large improvements in nutrition, both overall and for the targeted household members (mothers and infants).[18] Daily caloric intake for targeted members increased by 180 calories (9%) in year 1 and 392 calories (14%) in year 2 (Table 2, column 1), which is around a tenth of average caloric consumption in the control group and approximately equal to the difference between the 40th and 50th percentile of the control caloric consumption distribution. The increases for non-targeted members were smaller but similar, at 131 calories (8%) in year 1 and 96 (6%) in year 2 (Table 2, column 2); this difference is statistically significant in year 2 ($p = 0.01$) but not year 1 ($p = 0.45$).[19]

Dietary quality also increased for targeted household members. For both mothers and children, we use a standard dietary diversity score from 0-7 measuring the number of distinct dietary groups from which they consumed food (Ruel, 2003; Arimond and Ruel, 2004). These diversity scores increased for both mothers and targeted children in both years. We also document significant increases in an index of mothers' consumption of key nutrients in both years (Table 2, column 6), where this index is an average of the fraction of their recommended daily consumption of each nutrient (e.g. protein, visible fat, and iron) that they consume.[20] The increases are economically large: for example, in year 2, there is a 0.21SD increase in child dietary diversity, a 0.21SD increase in maternal dietary diversity score, and a 0.33SD increase in maternal nutrient score.[21]

Overall, transfers caused mothers and young children to consume more nutritious food. This is notable given recent concerns about intra-household allocation norms in Indian households (e.g., Jayachandran and Pande (2017)).[22] At the same time, a meaningful share of the transfers was—as

[18]In year 1, we can measure nutritional outcomes for all households. In year 2, we can only measure caloric intake and the micro-nutrient intake for the 596 households for which an in-person survey was completed. For the other outcomes, we supplement the in-person data with phone data on whether the individual had consumed particular items over the last day, which allows measurement of dietary diversity and minimum meal frequency.

[19]Targeted children do not see gains in terms of minimum meal frequency in either year, but gains in total consumption at each meal appear to compensate for this. The latter analysis likely also suffers from ceiling effects given the control group base rate has increased to 91% by year 2 as families shift to solid foods.

[20]Table A.13 presents effects on individual index components since some studies show that nutrients such as iron and protein matters more for cognitive outcomes than others (Roberts et al., 2022; Ip et al., 2017).

[21]These quality improvements could potentially reflect a mix of the cash transfers themselves and of the accompanying behavioral change messaging. To test the importance of the messaging, we examine whether there is heterogeneity in the treatment effect with respect to whether the mother listed a mobile number when registering for the program (which is balanced across treatment and control), as those who did not list a number did not receive the IVR messages. Table A.17 finds that across a variety of nutrition outcomes, treatment effects are no stronger when the respondent received the IVR messaging, indicating this messaging is unlikely to explain the treatment effects.

[22]Though see Spears et al. (2022) for a re-analysis of the data from this paper reaching different conclusions.

one would expect—spent on non-food items, or on food for other household members. These facts underscore the point that cash transfers may be a relatively blunt policy instrument for addressing specific nutritional deficiencies (and conversely, that a comprehensive analysis of the welfare benefits of the transfers would need to account for the value of non-food spending).

We also examine channels through which the treatment could affect downstream child development aside from the direct income effect of the transfer. One concern of the government was whether receiving cash transfers would reduce the interest of the household in other services at the AWC. On the other hand, transfers could crowd *in* utilization of services. Although there were no mechanical reasons to expect this—mothers were not required to make additional AWC visits to collect transfers, for example—we might see it if, for example, treated mothers visited the AWC in order to check on whether transfers had been sent. In practice, we see some evidence of crowd-*in*, where the mean total number of services received in year 1 increased by 0.23, or 4.6% of the control mean (Table 3, Column 2). The specific services for which we see significant increases were obtaining iron or calcium tablets, obtaining nutrition information, and obtaining vaccine shots (Table A.14). In year 2, the estimate is statistically insignificant, but we are underpowered given the much smaller sample size; the standard errors are so large that we could not reject a point estimate of the magnitude observed in year 1. From the point of view of local policy-makers, who generally wanted to encourage engagement with the ICDS system, these are positive results.

Columns (1), (3) and (4) of Table 3 examine nutritional knowledge, maternal empowerment and depression. Nutritional knowledge is modestly higher in the treatment group,[23] while there are some improvements in maternal empowerment in year 1 (and statistically insignificant decreases in maternal depression). This set of outcomes could plausibly affect child anthropometrics by inducing more spending on nutrition or greater allocation of nutritional resources to children, but we would not necessarily expect them to independently affect anthropometrics.[24] On net, these results imply that any measured impacts on child development are upper bounds on the pure income effect of transfers themselves, but the bounds are likely tight since effects on other outcomes are modest.

3.2 Effects on anthropometrics

Given their increased nutritional intake, one might hope to see children's anthropometrics improve as well. However, even though the estimated treatment effect coefficients were positive, they are mostly not significantly different from zero on average. Columns 1 and 3 of Table 4 report insignif-

[23]Increased engagement with the AWC likely explains this effect. IVR messages sent only to the treatment group were the only other messaging difference with the control group, but effects on nutritional knowledge are similar among treatment individuals who did and did not receive the IVR messaging (Table A.17).

[24]To place bounds on this effect on nutrition, we first estimate the cross-sectional relationships between knowledge/empowerment/depression and each nutrition outcome in Table 2 to get an upper bound on the *causal* effect of these services on the outcome. We then multiply each of those by the estimated effect of the treatment on the outcome to get an upper bound on how much of the effect of the treatment on nutritional intake is coming from a channel aside from income. We also do this for the AWC services discussed in the previous paragraph. Across the six nutritional outcomes in Table 3, these alternative channels explain an average of 5.7% of the overall treatment effect, suggesting that the income effect of transfers (or other unobservable channels) drives most of the observed effects.

icant effects on HAZ and WAZ in years 1 and 3.[25] This basic picture does not change if we focus instead on thresholds for being stunted or underweight (Columns 2 and 4); there is a reduction in the probability of being underweight in Year 1, but the result does not persist to Year 3.

How do these results fit into the landscape of related work? Three recent studies which evaluate broadly comparable interventions of cash transfers targeted to mothers and children do find significant mean effects on child anthropometrics: Ahmed et al. (2019) in Bangladesh, Carneiro et al. (2021) in Nigeria, and Field and Maffioli (2021) in Myanmar.[26] To benchmark our year 1, 2, and 3 HAZ and WAZ estimates against theirs, we normalize estimated effects by the total amount of money transferred (in PPP adjusted dollars) as of the time the outcome was measured (though these amounts happen to be fairly similar across studies). With just one exception (the impact after one year in Carneiro et al. (2021)), the resulting point estimates all lie within a 95% confidence interval of our year 3 estimate, and similarly our estimate lies within the 95% confidence intervals for all other estimates (Figure A.3). In this statistical sense, that implies there is nothing unusual about our estimates. Yet the confidence intervals are wide, implying that the data are also consistent with the conjecture that there are economic or biological factors that could make these settings meaningfully different.

To better assess this idea we turn next to examining one factor in particular that differs substantially across these study locations and could play a mediating role—namely, sanitation. Recent work has highlighted the role of poor sanitation in child development. Medical research shows that a poor sanitation environment can result in diminished physical development through factors such as malabsorption of nutrients due to intestinal disease, loss of nutrients due to diarrhea, and energy expended in fighting disease (e.g. Checkley et al. (2008); Petri et al. (2008); Lin et al. (2013)). Spears (2020) shows that sanitation environment explains a large fraction of international differences in child stunting, and that the higher prevalence of open defecation in India can account for much or all of the excess stunting in India relative to Africa. Among others, Hammer and Spears (2016) show that an intervention conducted to improve sanitation in villages in Maharashtra had large effects on child HAZ scores, consistent with work examining these links in India and some (but not all) other contexts (e.g. Bleakley (2007); Cameron et al. (2019, 2022)).

Consistent with this literature, there is a tight link between poor sanitation and child development outcomes in our sample. As a summary measure of sanitation environment, we construct a principal components index from variables such as whether the household's neighbors use a toilet.[27] The

[25]Table A.20 finds similar results in year 2, but we are underpowered to detect effects due to the smaller sample.

[26]We focus on these studies, which target children in the first few years of life when cash transfers are thought to have particularly strong potential, rather than other studies that target slightly older children (e.g., Fernald et al. (2008); Paxson and Schady (2010); Macours et al. (2012)) or outcomes at birth (e.g., Amarante et al. (2016)).

[27]The full set of variables are the presence of either (i) observable feces, (ii) wastewater or (iii) an open sewage ditch around the house; and the fraction of the household's neighbors who (iv) own or (v) use a toilet. Since there was no baseline survey, these are measured after the start of the treatment. However, the sanitation index is not related to treatment status (column (5) of table A.2), so it is still econometrically valid to analyze heterogeneity with respect to it. We prefer the principal components approach to other types of indices (e.g. equal weighting) because these variables reflect the underlying sanitation environment to different extents, and the PCA weights at least partially account for this. Results are nearly identical when we additionally include measures from the household itself, such as whether they use a toilet.

index is closely linked to neighborhood-level open defecation – the R^2 of a regression of the index on neighbor *usage* of toilets is 0.7 and results are similar if we use open defecation directly rather than the index (Table A.21) – where the neighborhood is likely the right level for thinking about effects of sanitation (Geruso and Spears, 2018). For ease of interpretation, we construct a percentile-based index of the sanitation distribution, where a value of zero is equal to the best sanitation environment (lowest incidence of open defecation) and one is the worst. In the control group, moving from the best to worst index value is associated with a 0.33σ reduction in WAZ and a 0.45σ reduction in HAZ, even after conditioning on household food expenditures.

The area we study rates very poorly in terms of sanitation environment. Figure 1 visualizes sanitation environment in Jharkhand in terms of open defecation, where the solid curve indicates the distribution of the open defecation rate across villages in Jharkhand as per the fifth round of the National Family Health Survey (2019-2020). In the median community, an estimated 35% of households practice open defecation. In other contexts with recent experimental work studying similar case transfer programs (colored lines), rates of open defecation are similar to or below even the 25th percentile community in Jharkhand. A poor sanitation environment may not only *directly* harm child development, but also mute the beneficial effects of income increases on child development. For example, frequent bouts of diarrheal disease transmitted via open defection could cause loss of the additional nutritional intake induced by transfers.

Table 5 tests whether the treatment interacts with the sanitation environment faced by the household on our primary outcomes. We estimate the effect of the treatment at multiple points in the sanitation distribution, ranging from the 10th percentile to the 90th percentile. For all four of the main anthropometric outcomes (weight-for-age, height-for-age, stunting, and being underweight), the effect of the cash transfers is statistically significant in the 10th percentile of the"poor sanitation" index, and three of the four are statistically significant at the 25th percentile. Furthermore, we can reject equality of the treatment effect across the sanitation distribution (joint $p = 0.01$ across all five outcomes), and observe that the linear interaction of the sanitation index with treatment is statistically significant at the 5% level for two of the outcomes, and of similar magnitude for the other two outcomes. The effect sizes are meaningful—a 0.13SD and 0.12SD increase in HAZ and WAZ for communities at the 25th percentile of the sanitation index—and comparable to experimental estimates for cash transfer programs of a similar value in other contexts.[28]

This heterogeneity also helps explain why we do not find significant *average* effects of the cash transfers on anthropometrics despite finding similar effects on nutritional intake to other recent experimental studies that do find anthropometric gains. Figure 1 shows that rates of open defecation were considerably lower in those studies (e.g., Field and Maffioli (2021) in Myanmar, Carneiro et al. (2021) in Nigeria, Levere et al. (2016) in the hill regions of Nepal, Fernald et al. (2008) in Mexico): this ranges from 13.8% in Myanmar to 24.0% in Nepal, equal to about the 30th percentile in our

[28] In the worst sanitation environments, the point estimates in panel A of Table 5 are mostly statistically insignificant but in the direction of *worse* outcomes in treatment areas. This may simply be statistical noise, but some scientific studies have found that increased intake of iron can worsen severity of malaria and worm-based diseases (Clark et al., 2014; Zhang et al., 2018; Held et al., 2006). We thank Hoyt Bleakley for suggesting this point.

sample.[29] Restricting to areas of our study that are comparable to those areas in terms of open defecation, the point estimates are quite similar, consistent with sanitation environment being an important mediator of the nutrition-anthropometric relationship.[30]

There are two main nuances involved in interpreting this heterogeneity. First, it could reflect differences across sanitation environments in the effect of treatment on nutritional intake itself. Areas with better sanitation might also be areas where recipients can more easily access their payments or where women have more influence in how money is used. We can test for this directly by examining whether effects on food expenditure, maternal or child consumption, or other related outcomes themselves depend significantly on the sanitation environment. Table A.22 shows that they do not: neither the joint p-value (panel A) nor tests of the linear interaction between treatment and the sanitation index (panel B) reject the null of no heterogeneity for these outcomes. Instead it must be that sanitation predicts how outcomes like these translate into anthropometric gains.

Next, sanitation could proxy for other characteristics that are correlated with it and affect the translation of income into growth. This would make no difference for the policy question of where to target cash transfers in order to promote child development, as that is simply a prediction problem. But it matters for inferences about what complementary interventions are most likely to amplify the effects of cash transfers. To explore this issue, we replace our sanitation index with its residual after regressing it on a set of seven observable variables selected using LASSO for their ability to predict the index.[31] Results using this residualized regressor (Table A.23) are similar to the original ones, and in some cases actually larger. It is possible that unobserved correlates of sanitation affect the income-growth translation, but this seems less plausible given how little things change when we remove the influence of observables. The results thus appear most consistent with the idea that the same nutritional intake does less for child growth in less sanitary places.

The ideal design for experimentally testing if interactions are significant would be a fully saturated experimental design with a cash transfer arm, a sanitation intervention arm, and an arm receiving both. However, this would need a very large sample in order to be statistically well-powered to detect interactions (Muralidharan et al., 2023). Calculations in Appendix B, based on our data as well as estimated coefficients from other studies of sanitation interventions, imply that a design with 80% power would require at least 6,000 clusters (see in particular Figure B.1). This is an order of magnitude larger than any similar field experiments with which we are familiar, and over 12 times larger than our current study. It seems unlikely that such a study will be feasible in the foreseeable future, and so our approach appears to be the most practicable way of demonstrating complementarities between income and sanitation in reducing child stunting.

[29]We calculate open defecation rates using the round of the Demographic and Health Survey for the country that was closest to the time of the study, subsetting to approximate the sub-national regions studied.

[30]This finding also argues against the role of genetics in the high rates of stunting among Indian children as posited by Panagariya (2013), consistent with work showing that the height of young children born in England to Indian migrants is similar to that of the children of native English (Alacevich and Tarozzi, 2017).

[31]These variables are selected from a larger list of fourteen variables using LASSO. The selected variables are: indicators for whether the household head is Muslim, Christian, below the poverty line, or a member of a scheduled caste or tribe, respondent education level, household size, and for child gender.

3.3 Additional measures of children's development

Most research on child development in India has focused on anthropometrics. We augment this by analyzing effects on an index of child development milestones. We use the parent-reported Ages and Stages Questionnaire (ASQ), which measures important milestones for gross motor, fine motor and cognitive development. The ASQ has been shown to have a high degree of correlation with clinical evaluations of child development via direct observations by trained psychologists.[32] The estimated treatment effect is a positive and statistically 0.12 standard deviations ($p < 0.01$) (Table A.16), which is equivalent to 19.5% of the difference between children in the poorest and wealthiest quartiles within the control group (as measured by an assets index).

Examining the underlying components of the index, the overall gain reflects increases across cognitive, gross motor and fine motor skills, such as the child's ability to say their own name and count to 10 (Table A.15, Panel B) as well as serve themselves food, hold a pen correctly, draw a basic figure, and fold paper (Table A.15, Panel A) (Schonhaut et al., 2013; Macours et al., 2012).[33] Based on these impacts across a broad range of functional capacities, it appears that gains in nutritional intake for mothers and children did indeed translate into accelerated child development.[34] Our findings are therefore consistent with experimental studies in other contexts that have found positive links between cash transfers in early childhood and cognitive development (Paxson and Schady, 2010; Macours et al., 2012; Gilligan and Roy, 2013; Barham et al., 2013). It may be that anthropometric gains for young children are more sensitive to health shocks from a poor sanitation environment than developmental milestones.

We also see a significant mean effect on some anthropometric outcomes for the *siblings* of targeted children (a secondary outcome in our pre-analysis plan). Sibling WAZ, increased by 0.11σ in year 3 ($p < 0.05$), and by a similar 0.13σ in year 2 (Table A.18).[35] Additionally, there is a 10 percentage point decline in the likelihood of the sibling being moderately underweight ($p < 0.05$) in year 2. We

[32]For children who are around 3 years (as in our study population at the time of ASQ administration), Schonhaut et al. (2013) find a correlation of 0.75 between ASQ and direct observation using the "gold-standard" Bayley Scales of Infant and Toddler Development. Attanasio et al. (2016) find slightly lower, but still moderate to high correlations for each sub-components of these indices in a developing country context.

[33]One concern with these results is experimenter demand effects, where treatment parents overstate their children's performance on these tasks relative to control. This seems unlikely for a few reasons. First, we observe treatment effects even on outcomes that could be directly observed by the surveyors, such as the child ability to say their name or serve themselves. We also do not see a uniform inflation across questions or more inflation for less objective questions or ones that are harder for the surveyor to observe (e.g., naming primary colors): we can reject equivalence of estimates across outcomes in table A.15, and observe effects on highly objective questions such as ability to say their name or count to 5. Second, the Y3 measures are from over a year after the transfers had ceased, so the incentives to under-report child development to get transfers are the same in treatment and control. Finally, experimenter demand effects require the subjects having an understanding of researcher incentives that seems unlikely in this context. The treatment group is not told that outcomes are measured in a control group and compared to theirs to evaluate the program. Thus experimenter demand effects seem unlikely here, even if they are plausible in settings where the control and treatment groups are more visible to one another.

[34]We cannot reject the null of no impact on our index of second-year child development (Table A.20). However, this index has fewer questions and was collected from a smaller sample primarily over the phone, so that analysis is likely underpowered. The year 2 and year 3 questions are almost completely different; each is tailored to try to capture age-appropriate measures of functioning, and so they have only one question in common.

[35]The year 2 WAZ estimate is statistically insignificant, likely because the standard errors are twice the size of those for year 3 (0.10 as compared to 0.05) due to the much smaller sample size in year 2.

only collected sibling height in year 3, and estimated effects on HAZ are positive (at 0.07σ), but not precise enough to rule out zero or that they are significantly different from the WAZ effects. These findings are consistent with the fact that food consumption and nutritional value of foods improved for other household members in addition to the mother and targeted child. It also underscores the point that, regardless of labelling, households have substantial freedom in how to spend the extra income received from transfer programs.[36]

4 Conclusion

The slow translation of income growth into reductions in child stunting in India has been a puzzle for both research and policy. It has also contributed to skepticism regarding the effectiveness of income transfers to poor households as a policy instrument for improving child development (Khera, 2014). Reasons for skepticism include concerns that expansion of household budget sets may not necessarily translate into more spending on food, that households cannot access or would not select nutritious food, or that intra-household dynamics of resource allocation may limit the benefits for mothers and young children.

Our results suggest that such concerns may be overstated, as the experimentally induced increases in income in our setting successfully improved nutritional intake of mothers and children. One caveat is that the elasticity of nutritional intake with respect to income we estimate in our study may be somewhat higher than that of other types of income increases due to the transfer going into the bank account of the female head of household and the associated light-touch messaging. However, both features are common in targeted income transfer programs. Most of them prefer to transfer funds to women, and programs targeting child welfare are often accompanied by some framing or messaging. Thus, our estimates are directly relevant to policy debates on the food consumption and nutrition impacts of maternal income transfer programs, and are also informative about the likely effects of income growth in general.[37]

We find that these nutritional improvements also led to large gains in cognitive and motor skills of targeted children. To the extent that non-anthropometric forms of child development may be as or more predictive of later economic productivity in an economy that rewards cognitive functioning more than physical strength (Pitt et al., 2012), these are encouraging results for policies focused on income growth or cash transfers. Further, we do not observe interactions with sanitation for these outcomes, suggesting that income growth in India may have produced gains in child human development beyond those measured in anthropometrics. However, additional research is needed to investigate the long-run persistence of gains observed at age 3, and their significance for adult economic outcomes.

[36] We do not find evidence of heterogeneity with respect to sanitation for sibling anthropometric outcomes (Table A.19). This is what one might expect given evidence to date on age differentials in the effects of poor sanitation: the loss of nutrients from diarrhea, which is strongly positively correlated with the prevalence of open defecation (Lin et al., 2013), has been shown to be more costly for the physical growth of younger children (Nasrin et al., 2023).

[37] We also find that the treatment effect of messaging was limited in this study (Table A.17), increasing the likelihood that our estimates generalize to other forms of income growth.

In contrast, we observe that the nutritional gains only translate into effects on the more commonly-used anthropometric measures of child development when sanitation is relatively good. Previous work has shown a negative relationship between *levels* of sanitation and child anthropometrics (Hammer and Spears, 2016; Gertler et al., 2015), but this paper extends that result by demonstrating that income and sanitation are *complements* in production of child physical development.

These results are directly relevant for policy in India, where efforts to reduce child stunting have focused on programs to improve feeding of children in poor households. Policy debates in recent years have revolved around the *form* that such targeted redistribution should take—especially around cash versus in-kind programs. Our results indicate that even if child nutritional intake is significantly improved (through either cash or kind transfers), the impacts on stunting may be limited in settings of poor sanitation. They suggest that a more effective approach would be to identify the areas with the highest levels of stunting (which are correlated with both poverty and open defecation), and prioritize programs to improve *both* sanitation and feeding of poor children in a coordinated way. Paying attention to coordination may be especially important because sanitation and child feeding programs are implemented by different government departments.

The case for policy focus on sanitation to accelerate the translation of income growth into reductions in child stunting is magnified by the fact that poor households are likely to under-invest in sanitation relative to socially optimal levels as their income grows. Income growth will lead households to consume both additional food and improve sanitation. However, the latter will grow slower than socially optimal both because individually rational households may not account for the positive spillovers across households from reduced open defecation, and because of the considerable fixed costs of constructing toilets of a high-enough quality to effectively sequester fecal matter.[38] Both spillovers and fixed costs (or indivisibility of capital goods) are classic ideas in development economics for explaining under-investment in items that have positive social net present value (Murphy et al., 1989). Our results emphasize the need for policies to boost sanitation and use of toilets in India, as some recent government initiatives have attempted. However, they also highlight that continued and sustained efforts to reduce open defecation further are likely needed to accelerate the translation of India's income gains into gains in child development.

[38]The most cost-effective option in these settings is a twin-pit toilet, which costs at least Rs. 12,000, over two months of household consumption in our setting.

References

Ahmed, Akhter, John F Hoddinott, and Shalini Roy, "Food transfers, cash transfers, behavior change communication and child nutrition: Evidence from Bangladesh," Technical Report, International Food Policy Research Institute 2019.

Alacevich, Caterina and Alessandro Tarozzi, "Child height and intergenerational transmission of health: evidence from ethnic Indians in England," *Economics & Human Biology*, 2017, *25*, 65–84.

Amarante, Verónica, Marco Manacorda, Edward Miguel, and Andrea Vigorito, "Do cash transfers improve birth outcomes? Evidence from matched vital statistics, program, and social security data," *American Economic Journal: Economic Policy*, 2016, *8* (2), 1–43.

Arimond, Mary and Marie T Ruel, "Dietary diversity is associated with child nutritional status: evidence from 11 demographic and health surveys," *The Journal of nutrition*, 2004, *134* (10), 2579–2585.

Attanasio, Orazio, Marta Rubio-Codina, M Caridad Araujo, Pablo Muñoz, and Sally Grantham-McGregor, "Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies," *PloS one*, 2016, *11* (8), e0160962.

Baird, Sarah, Craig McIntosh, and Berk Özler, "When the money runs out: Do cash transfers have sustained effects on human capital accumulation?," *Journal of Development Economics*, 2019, *140*, 169–185.

Barcellos, Silvia Helena, Leandro S Carvalho, and Adriana Lleras-Muney, "Child gender and parental investments in India: Are boys and girls treated differently?," *American Economic Journal: Applied Economics*, 2014, *6* (1), 157–89.

Barham, Tania, Karen Macours, and John A Maluccio, "Boys' cognitive skill formation and physical growth: Long-term experimental evidence on critical ages for early childhood interventions," *American Economic Review*, 2013, *103* (3), 467–71.

Bastagli, Francesca, Jessica Hagen-Zanker, Luke Harman, Valentina Barca, Georgina Sturge, Tanja Schmidt, and Luca Pellerano, "Cash transfers: what does the evidence say," *A rigorous review of programme impact and the role of design and implementation features. London: ODI*, 2016, *1* (7), 1.

Behrman, Jere R, "Nutrition, health, birth order and seasonality: Intrahousehold allocation among children in rural India," *Journal of Development Economics*, 1988, *28* (1), 43–62.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on Treatment Effects after Selection among High-Dimensional Controls," *Review of Economic Studies*, 11 2014, *81* (2), 608–650.

Bleakley, Hoyt, "Disease and development: evidence from hookworm eradication in the American South," *Quarterly Journal of Economics*, 2007, *122* (1), 73–117.

Bricker, Diane, Jane Squires, Linda Mounts, LaWanda Potter, Robert Nickel, Elizabeth Twombly, and Jane Farrell, "Ages and stages questionnaire," *Baltimore, MD: Paul H. Brookes*, 1999.

Cameron, Lisa, Paul Gertler, Manisha Shah, Maria Laura Alzua, Sebastian Martinez, and Sumeet Patil, "The dirty business of eliminating open defecation: The effect of village sanitation on child height from field experiments in four countries," *Journal of Development Economics*, 2022, *159*, 102990.

_ , Susan Olivia, and Manisha Shah, "Scaling up sanitation: Evidence from an RCT in Indonesia," *Journal of Development Economics*, 2019, *138*, 1–16.

Carneiro, Pedro, Lucy Kraftman, Giacomo Mason, Lucie Moore, Imran Rasul, and Molly Scott, "The impacts of a multifaceted prenatal intervention on human capital accumulation in early life," *American Economic Review*, 2021, *111* (8), 2506–49.

Checkley, William, Gillian Buckley, Robert H Gilman, Ana MO Assis, Richard L Guerrant, Saul S Morris, Kåre Mølbak, Palle Valentiner-Branth, Claudio F Lanata, Robert E Black et al., "Multi-country analysis of the effects of diarrhoea on childhood stunting," *International Journal of Epidemiology*, 2008, *37* (4), 816–830.

_ , Robert H Gilman, Robert E Black, Leonardo D Epstein, Lilia Cabrera, Charles R Sterling, and Lawrence H Moulton, "Effect of water and sanitation on childhood health in a poor Peruvian peri-urban community," *The Lancet*, 2004, *363* (9403), 112–118.

Chen, Jiafeng and Jonathan Roth, "Log-like? Identified ATEs defined with zero-valued outcomes are (arbitrarily) scale-dependent," Technical Report, arXiv 2023.

Clark, Martha A, Morgan M Goheen, Anthony Fulford, Andrew M Prentice, Marwa A Elnagheeb, Jaymin Patel, Nancy Fisher, Steve M Taylor, Raj S Kasthuri, and Carla Cerami, "Host iron status and iron supplementation mediate susceptibility to erythrocytic stage Plasmodium falciparum," *Nature communications*, 2014, *5* (1), 4446.

Coffey, Diane and Dean Spears, *Where India goes: abandoned toilets, stunted development and the costs of caste*, Harper Collins, 2017.

Currie, Janet and Tom Vogl, "Early-life health and adult circumstance in developing countries," *Annual Review of Economics*, 2013, *5* (1), 1–36.

Davala, Sarath, Renana Jhabvala, Kapoor-Mehta Sowmya, and Guy Standing, *Basic Income: A Transformative Policy for India*, Bloomsbury Academic, 2015.

Deaton, Angus and Jean Drèze, "Food and Nutrition in India: Facts and Interpretations," *Economic and Political Weekly*, 2009, *44* (7), 42–65.

Doyle, Orla, "The first 2,000 days and child skills," *Journal of Political Economy*, 2020, *128* (6), 2067–2122.

Esrey, Steven A, "Water, waste, and well-being: a multicountry study," *American journal of epidemiology*, 1996, *143* (6), 608–623.

Fernald, Lia CH, Paul J Gertler, and Lynnette M Neufeld, "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades," *The Lancet*, 2008, *371* (9615), 828–837.

Field, Erica M and Elisa M Maffioli, "Are behavioral change interventions needed to make cash transfer programs work for children? Experimental evidence from Myanmar," Technical Report, National Bureau of Economic Research 2021.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "On her own account: How strengthening women's financial control impacts labor supply and gender norms," *American Economic Review*, 2021, *111* (7), 2342–75.

Fink, Günther, Isabel Günther, and Kenneth Hill, "The effect of water and sanitation on child health: evidence from the demographic and health surveys 1986–2007," *International journal of epidemiology*, 2011, *40* (5), 1196–1204.

Ganimian, Alejandro J, Karthik Muralidharan, and Christopher R Walters, "Augmenting state capacity for child development: Experimental evidence from India," Technical Report, National Bureau of Economic Research 2021.

Gertler, Paul, Manisha Shah, Maria Laura Alzua, Lisa Cameron, Sebastian Martinez, and Sumeet Patil, "How does health promotion work? Evidence from the dirty business of eliminating open defecation," Technical Report, National Bureau of Economic Research 2015.

Geruso, Michael and Dean Spears, "Neighborhood sanitation and infant mortality," *American Economic Journal: Applied Economics*, 2018, *10* (2), 125–62.

Gilligan, Daniel O and Shalini Roy, "Resources, stimulation, and cognition: How transfer programs and preschool shape cognitive development in Uganda," Technical Report, International Food Policy Research Institute 2013.

Glennerster, Rachel, Claire Walsh, and Lucia Diaz-Martin, "A Practical Guide to Measuring Women's and Girl's Empowerment in Impact Evaluations," Technical Report, Abdul Latif Jameel Poverty Action Lab 2018.

Hammer, Jeffrey and Dean Spears, "Village sanitation and child health: effects and external validity in a randomized field experiment in rural India," *Journal of Health Economics*, 2016, *48*, 135–148.

Held, Melissa R, Richard D Bungiro, Lisa M Harrison, Iqbal Hamza, and Michael Cappello, "Dietary iron content mediates hookworm pathogenesis in vivo," *Infection and immunity*, 2006, *74* (1), 289–295.

Ip, Patrick, Frederick Ka Wing Ho, Nirmala Rao, Jin Sun, Mary Eming Young, Chun Bong Chow, Winnie Tso, and Kam Lun Hon, "Impact of nutritional supplements on cognitive development of children in developing countries: A meta-analysis," *Scientific reports*, 2017, *7* (1), 10611.

Jayachandran, Seema and Ilyana Kuziemko, "Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India," *Quarterly Journal of Economics*, 2011, *126* (3), 1485–1538.

_ **and Rohini Pande**, "Why are Indian children so short? The role of birth order and son preference," *American Economic Review*, 2017, *107* (9), 2600–2629.

Khera, Reetika, "Cash vs. in-kind transfers: Indian data meets theory," *Food Policy*, 2014, *46*, 116–128.

Klasen, Stephan, "Poverty, undernutrition, and child mortality: Some inter-regional puzzles and their implicationsfor research and policy," *The Journal of Economic Inequality*, 2008, *6*, 89–115.

Levere, Michael, Gayatri Acharya, and Prashant Bharadwaj, "The role of information and cash transfers on early childhood development: evidence from Nepal," Technical Report, National Bureau of Economic Research 2016.

Lin, Audrie, Benjamin F Arnold, Sadia Afreen, Rie Goto, Tarique Mohammad Nurul Huda, Rashidul Haque, Rubhana Raqib, Leanne Unicomb, Tahmeed Ahmed, John M Colford Jr et al., "Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh," *American Journal of Tropical Medicine and Hygiene*, 2013, *89* (1), 130–137.

Longvah, T, R Ananthan, K Bhaskarachary, and Venkaiah K, "Indian Food Composition Tables," Technical Report, National Institute of Nutrition (Indian Council of Medical Research) 2017.

Macours, Karen, Norbert Schady, and Renos Vakis, "Cash transfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment," *American Economic Journal: Applied Economics*, 2012, *4* (2), 247–73.

Malik, Faisal, "Take-home rations used to feed cattle," *Hindustan Times*, July 2016.

Manley, James, Seth Gitter, and Vanya Slavchevska, "How effective are cash transfer programmes at improving nutritional status? A rapid evidence assessment of programmes' effects on anthropometric outcomes," *Social Science Research Unit, Institute of Education, University of London*, 2012.

Ministry of Health and Family Welfare, "National Family Health Survey (NFHS-5) 2019-2021," Technical Report 2022.

_ , "National Family Health Survey (NFHS-5) 2019-2021: Jharkhand Fact Sheet," Technical Report 2022.

Muralidharan, Karthik, *Accelerating India's Development: A State-led Roadmap for Effective Governance*, Penguin Viking, 2024.

_ , Mauricio Romero, and Kaspar Wüthrich, "Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments," *The Review of Economics and Statistics*, 03 2023, pp. 1–44.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "Industrialization and the big push," *Journal of political economy*, 1989, *97* (5), 1003–1026.

Narayanan, Sudha and Shree Saha, "Take home rations (THR) and cash transfers for maternal and child nutrition: A synthesis of evidence in India," *Indira Gandhi Institute of Development Research, Mumbai Working Papers*, 2020, (2020-039).

Nasrin, Dilruba, Yuanyuan Liang, Helen Powell, Ines Gonzalez Casanova, Samba O Sow, M Jahangir Hossain, Richard Omore, Doh Sanogo, Boubou Tamboura, Syed MA Zaman et al., "Moderate-to-severe diarrhea and stunting among children younger than 5 years: findings from the Vaccine Impact on Diarrhea in Africa (VIDA) study," *Clinical Infectious Diseases*, 2023, *76* (Supplement_1), S41–S48.

National Institute of Nutrition, "Dietary Guidelines for Indians - A Manual," Technical Report, National Institute of Nutrition (Indian Council of Medical Research) 2011.

Niehaus, Paul and Tavneet Suri, "Cash transfers," *In preparation for the Handbook of Social Protection*, 2024.

NITI Aayog, "Multi-dimensional Poverty Index," *NITI report*, 2021.

Panagariya, Arvind, "Does India really suffer from worse child malnutrition than sub-Saharan Africa?," *Economic and Political Weekly*, 2013, pp. 98–111.

Patil, Sumeet R., Benjamin F. Arnold, Alicia L. Salvatore, Bertha Briceno, Sandipan Ganguly, John M. Colford Jr, and Paul J. Gertler, "The Effect of India's Total Sanitation Campaign on Defecation Behaviors and Child Health in Rural Madhya Pradesh: A Cluster Randomized Controlled Trial," *PLOS Medicine*, 2014, *11*, 1–17.

Paxson, Christina and Norbert Schady, "Does money matter? The effects of cash transfers on child development in rural Ecuador," *Economic Development and Cultural Change*, 2010, *59* (1), 187–229.

Petri, William A, Mark Miller, Henry J Binder, Myron M Levine, Rebecca Dillingham, Richard L Guerrant et al., "Enteric infections, diarrhea, and their impact on function and development," *Journal of Clinical Investigation*, 2008, *118* (4), 1277–1290.

Pitt, Mark M, Mark R Rosenzweig, and Mohammad Nazmul Hassan, "Human capital investment and the gender division of labor in a brawn-based economy," *American economic review*, 2012, *102* (7), 3531–3560.

Rah, Jee Hyun, Aidan A Cronin, Bhupendra Badgaiyan, Victor M Aguayo, Suzanne Coates, and Sarah Ahmed, "Household sanitation and personal hygiene practices are associated with child stunting in rural India: a cross-sectional analysis of surveys," *BMJ open*, 2015, *5* (2), e005180.

Ramalingaswami, Vulimiri, Urban Jonsson, and Jon Rohde, "The asian enigma," *Progress of nations*, 1996, pp. 10–17.

Roberts, Marina, Terezie Tolar-Peterson, Abby Reynolds, Caitlin Wall, Nicole Reeder, and Gina Rico Mendez, "The effects of nutritional interventions on the cognitive development of preschool-age children: A systematic review," *Nutrients*, 2022, *14* (3), 532.

Ruel, Marie T, "Operationalizing dietary diversity: a review of measurement issues and research priorities," *The Journal of nutrition*, 2003, *133* (11), 3911S–3926S.

Schonhaut, Luisa, Iván Armijo, Marianne Schönstedt, Jorge Alvarez, and Miguel Cordero, "Validity of the ages and stages questionnaires in term and preterm infants," *Pediatrics*, 2013, *131* (5), e1468–e1474.

Spears, Dean, "Exposure to open defecation can account for the Indian enigma of child height," *Journal of Development Economics*, 2020, *146*, 102277.

_ , Diane Coffey, and Jere R Behrman, "Endogenous inclusion in the Demographic and Health Survey anthropometric sample: Implications for studying height within households," *Journal of Development Economics*, 2022, *155*, 102783.

Sukhtankar, Sandip, "Digital Technology and Social Protection: Promises and Pitfalls," Technical Report, University of Virginia 2023.

Vidmar, Suzanna, Tim Cole, and Huiqi Pan, "Standardizing anthropometric measures in children and adolescents with functions for egen: Update," *The Stata Journal*, 2013, *13*, 366–378.

Weaver, Jeffrey, Sandip Sukhtankar, Paul Niehaus, and Karthik Muralidharan, "Jharkhand ICDS Pre-Analysis Plan Report," `https://www.socialscienceregistry.org/trials/2899/` 2023.

WHO, "WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age," Technical Report, World Health Organization 2006.

_ , "Indicators for assessing infant and young child feeding practices. Part 1: Definitions," Technical Report, World Health Organization 2008.

World Bank, "High Economic Growth has not Translated into Superior Nutritional Status for India," Technical Report, The World Bank 2009.

World Health Organization, "Levels and trends in child malnutrition: UNICEF," 2023.

Zhang, De-Liang, Jian Wu, Binal N Shah, Katja C Greutélaers, Manik C Ghosh, Hayden Ollivierre, Xin zhuan Su, Philip E Thuma, George Bedu-Addo, Frank P Mockenhaupt et al., "Erythrocytic ferroportin reduces intracellular iron accumulation, hemolysis, and malaria risk," *Science*, 2018, *359* (6383), 1520–1523.

Table 1: Household Spending

	Food Spending			Sin Spending			Non-Food Spending		
	IHS	Extensive	Intensive	IHS	Extensive	Intensive	IHS	Extensive	Intensive
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Year 1 Outcomes									
Treatment	0.15^{***}	0.005	0.11^{***}	−0.08	−0.01	−0.01			
	(0.05)	(0.004)	(0.04)	(0.14)	(0.02)	(0.07)			
Control Mean	8.33	0.99	7.75	4.1	0.63	5.85			
Observations	2,360	2,360	2,334	2,337	2,337	1,469			
Panel B: Year 2 Outcomes									
Treatment	0.21^{***}	0.01^{**}	0.09^{**}	0.11	0.03	−0.10	0.19^{*}	0.02	0.04
	(0.07)	(0.01)	(0.04)	(0.29)	(0.04)	(0.14)	(0.11)	(0.01)	(0.06)
Control Mean	8.41	0.97	7.99	4.16	0.63	5.94	9.02	0.96	8.67
Observations	1,869	1,869	1,824	582	582	369	1,377	1,377	1,336

The unit of analysis is the household. IHS results pertain to an outcome variable transformed via inverse hyperbolic sine. Extensive margin results concern whether households recorded *any* of the associated spending in the previous month. Intensive margin results concern the subset of households for whom expenditure is positive, and pertain to an outcome variable transformed via natural log. Outcomes in Year 2 (panel B) correspond to a pooled sample of field and phone observations. In panel B, columns (1) to (3) combine field surveys and all of the year 2 phone surveys, columns (4) to (6) use only field survey data since sin good spending was not collected over the phone, and columns (7) to (9) pool data from the field survey and second round of phone surveys as this information was not collected in the first round of phone surveys. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table 2: Quantity & Quality of Food Intake

	Household		Child		Mother	
	Caloric consumption (mother and child)	Caloric consumption (per capita)	Dietary diversity score	Minimum meal frequency	Dietary diversity score	Nutrient index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year 1 Outcomes						
Treatment	180.43***	131.34***	0.16***	0.03	0.08***	0.09***
	(36.29)	(42.38)	(0.05)	(0.02)	(0.02)	(0.02)
Control Mean	1873.54	1678.39	1.59	0.61	2.78	0.19
Observations	2,360	2,321	2,349	2,344	2,349	2,349
Panel B: Year 2 Outcomes						
Treatment	392.53***	95.56	0.24***	0.01	0.22***	0.17***
	(93.33)	(79.44)	(0.06)	(0.02)	(0.05)	(0.04)
Control Mean	2532.01	1499.86	3.4	0.91	3.35	0.22
Observations	596	564	1,421	960	1,421	596

Unit of analysis is indicated in the first row of the table. Dietary diversity score (0-7) is defined as the number of the following dietary groups that were consumed by the respondent: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables. Nutrient index is computed by estimating the percent of the recommended daily value of calories, protein, visible fat, calcium, iron, thiamine, riboflavin, niacin, pyridoxin, and dietary folate consumed and averaging. Observation counts differ across year 2 outcomes because some outcomes could not be collected via phone survey, which field work transitioned to after the onset of COVID-19. For the year 2 outcomes in panel B, Columns (1), (2), and (6) use field survey data, as it is necessary to measure the quantity of individual ingredients to create these measures. Columns (3), (4), and (5) pool data from the field and multiple rounds of phone surveys (see table A.5 for a visualization of each of the outcomes of interest and the survey round in which they were collected). Column (4) pertains to the subset of the sampled children who are under 23 months — i.e. those for whom minimum meal frequency is considered a clinically meaningful measure. Standard errors are in parentheses and are clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table 3: Additional Outcomes

	Mother				Child	
	Nutritional knowledge index	Anganwadi services received	Empowerment index	Depression index	Probability of visiting formal medical provider	Total illnesses in the past three months
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year 1 Outcomes						
Treatment	0.17**	0.15**	0.13***	−0.08	0.03*	
	(0.07)	(0.07)	(0.05)	(0.07)	(0.02)	
Control Mean	3.42	3.87	2.31	2.58	0.77	
Observations	2,349	2,300	2,349	2,347	2,242	
Panel B: Year 2 Outcomes						
Treatment	0.29***	−0.002	0.06	−0.24	0.10***	0.02
	(0.11)	(0.17)	(0.07)	(0.16)	(0.03)	(0.05)
Control Mean	3.64	3.35	2.59	3.19	0.5	1.88
Observations	596	586	1,049	596	1,271	1,271

The unit of analysis is indicated in the first row of the table. The nutritional knowledge (0-6), empowerment (0-5), and depression (0-5) indices correspond to the number of questions answered by the respondent that suggest a high degree of the corresponding characteristic. Anganwadi services received corresponds to the number of AWC services (0-9) that the respondent received. For results on each AWC service, see Table A.14. Observation counts differ across Y2 outcomes because some outcomes could not be collected via phone survey, which field work transitioned to after the onset of COVID-19. Others were only collected during one round of phone surveys since the phone instrument had to be kept short to reduce respondent fatigue. Nutritional knowledge (column 1), anganwadi services (column 2), and the depression index (column 4) were collected in the field survey; the empowerment index (column 3) was collected in the field survey and second round of phone data collection; and child illness information (columns 5 and 6) was collected in the field survey and first round of phone data collection (see table A.5 for a visualization of each of the outcomes of interest and the survey round in which they were collected). Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table 4: Anthropometric Outcomes

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Year 1 Outcomes				
Treatment	0.005	−0.002	0.002	-0.04^{**}
	(0.05)	(0.02)	(0.04)	(0.02)
Control Mean	-1.47	0.31	-1.67	0.4
Observations	2,355	2,355	2,355	2,355
Panel B: Year 3 Outcomes				
Treatment	0.06	0.003	0.04	-0.02
	(0.05)	(0.02)	(0.04)	(0.02)
Control Mean	-0.88	0.19	-1.2	0.23
Observations	2,124	2,124	2,165	2,165

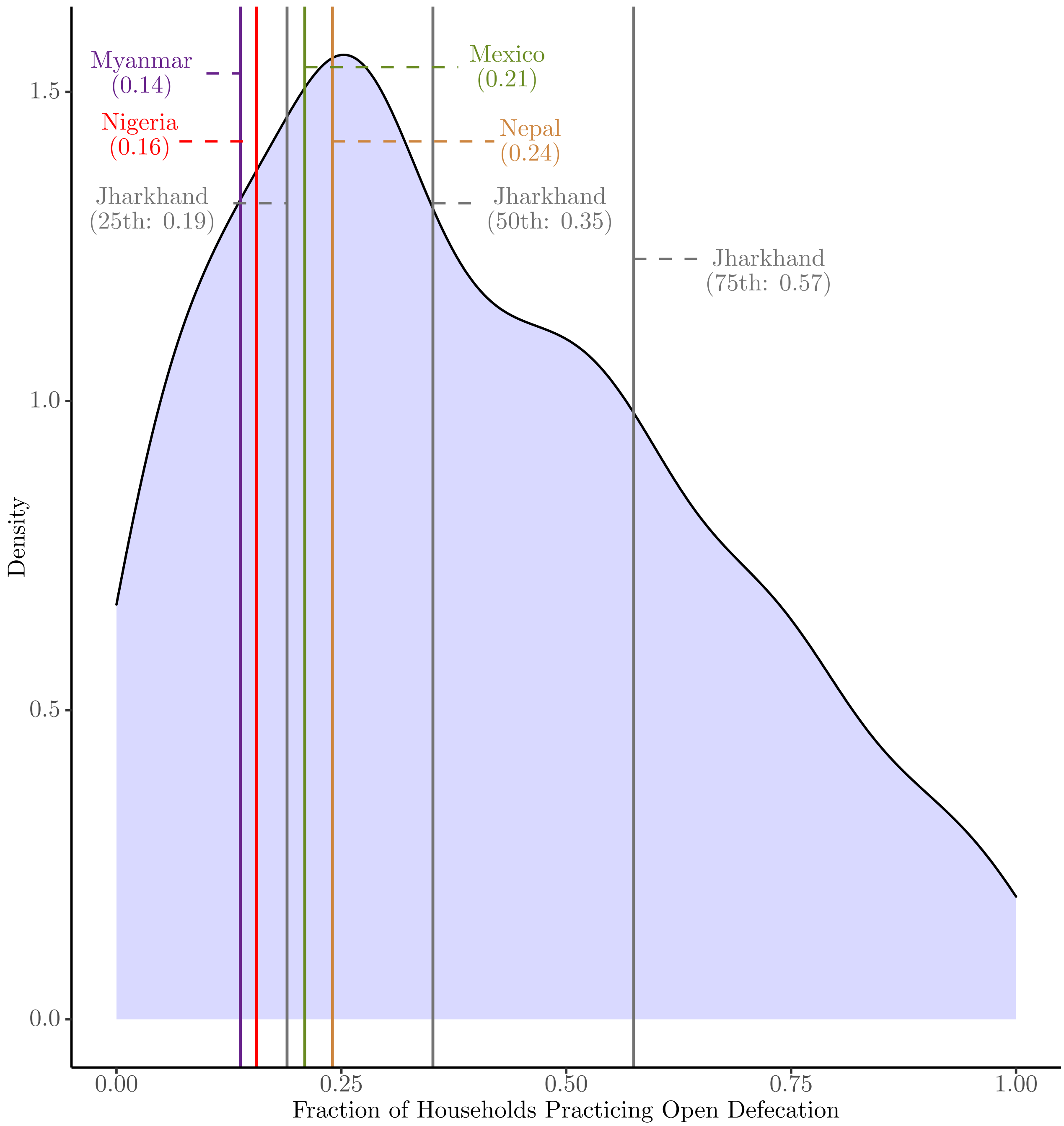
The unit of analysis is the child. HAZ and WAZ denote the child's height-for-age and weight-for-age z-scores, respectively. Children with WAZ and HAZ of less than -2 are classified as moderately stunted and underweight respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table 5: Interactions with Neighborhood Sanitation Environment

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Implied Effects by Percentile				
10%	0.16*	-0.04	0.15**	-0.04
25%	0.13*	-0.03	0.12**	-0.04
50%	0.08	0	0.05	-0.02
75%	0.02	0.02	-0.01	0
90%	-0.06	0.06*	-0.11	0.02
Test of Equality: p-value	0.21	0.04	0.05	0.21
Joint p-value	0.02			
Panel B: Underlying Parameter Estimates				
Treatment	0.19*	-0.06*	0.18**	-0.05*
	(0.10)	(0.03)	(0.08)	(0.03)
Poor Sanitation Index	-0.28**	-0.02	-0.12	0.02
	(0.14)	(0.05)	(0.12)	(0.05)
Treatment × Poor Sanitation Index	-0.26	0.12**	-0.32**	0.08
	(0.19)	(0.06)	(0.14)	(0.06)
Control Mean	-0.88	0.19	-1.2	0.23
Observations	2124	2124	2165	2165

The unit of analysis is the child. Outcomes are from Year 3. HAZ and WAZ denote the child's height-for-age and weight-for-age z-score. Children with HAZ and weight-for-height z-scores of less than -2 are classified as stunted and wasted, respectively. As a summary measure of sanitation environment, we constructed a principal components index from 8 variables such as whether the household or their neighbors use a toilet. The poor sanitation index used in this regression is equal to the household's percentile rank on that principal components index, where a value of zero is equal to the best sanitation environment (lowest incidence of open defecation) and one is the worst. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Figure 1: Distribution of Fraction of Households Practicing Open Defecation in the DHS



This figure presents the distribution of the fraction of households per community in rural Jharkhand practicing open defecation per the Demographic and Health Surveys-VII (2019-2021) program. Vertical gray lines correspond to the quantiles of the DHS community-wise fraction of households practicing open defecation in Jharkhand. The other vertical lines correspond to the mean community-wise fraction of households practicing open defecation in other countries in which studies of conditional cash transfers have taken place: Mexico (Fernald et al., 2008) (use 1987 DHS), Nepal (Levere et al., 2016) (use DHS 2016), Myanmar (Field and Maffioli, 2021) (use DHS 2015-2016), and Nigeria (Carneiro et al., 2021) (use DHS 2018). In each case, DHS data is sub-setted to best approximate the samples used and year of the respective studies. For example, (Fernald et al., 2008) uses data from 1998 in Mexico, so we use the most temporally proximate DHS wave (1987), which may overstate the rate of open defecation in 1998 (and certainly in the present).

A Appendix: Additional Exhibits

Table A.1: Balance Tests

Outcome	Control (1)	Treatment (2)	p-value (3)
Time Invariant Household Characteristics			
Respondent's Education (Years)	7.05	7.19	0.15
Scheduled Caste or Tribe	0.59	0.59	0.95
Birth Order	2.19	2.13	0.1
AWC-level Characteristics			
Complete Registrations	7.66	7.48	0.61
Village-level Characteristics (2011 Census)			
% Households Living in Poor Condition Houses	0.62	0.64	0.57
% Households with a Toilet	0.08	0.08	0.56
Area of Village (Hectares)	478.33	403.77	0.3
Distance from All-Weather Road (km)	0.58	0.48	0.19
Distance from Nearest Bank (km)	1.90	2.04	0.35
Distance from Regular Market / Mandi (km)	1.88	1.88	0.81
% Scheduled Caste or Tribe (2011)	0.42	0.39	0.18
Village Population	1938.75	1863.12	0.56
Number of AWCs	240	240	
Joint F-test			0.8

This table presents balance tests for key outcomes across the treatment and control group. Columns 1-2 present means for each outcome by treatment group. Column 3 shows p-values for balance tests of each outcome across treatment groups. The final row corresponds to a joint F-test conducted using a seemingly unrelated regression framework. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.2: Survey Completion

	Year 1	Year 2		Year 3	
	Endline (1)	Endline (2)	Phone (3)	Endline (4)	Insanitation (5)
Treatment	0.00	−0.01	−0.00	0.01	−0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Control Mean	0.78	0.20	0.62	0.72	0.37
Observations	2956	2956	2956	2956	2166

Columns (1) to (4) of this table presents estimates for the relationship between treatment and probability of survey completion across each of this study's survey waves. Column (5) regresses our sanitation index on the treatment indicator to see whether sanitation is related to treatment status. Regressions include sector and AWC fixed effects. $^{*}p < .10,^{**}p < .05,^{***}p < .01$

Table A.3: Bank Account Usage

		Purpose of withdrawal				
	Withdrew in last 30 days	Daily expenses	Medical expenses	Non-food consumption	Loan payments	House
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.14***	0.10***	0.09**	−0.04	−0.07**	−0.04***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.01)
Control Mean	0.2	0.47	0.43	0.21	0.12	0.07
Observations	830	822	822	822	822	822

The unit of analysis is the respondent. Purpose of withdrawal refers to whether their most recent withdrawal was used for the corresponding class of expenditure. "Daily expenses" most commonly refers to food or other small purchases for the household. Non-food consumption refers to non-food items that could be considered consumption rather than investment (e.g., clothes). House refers to spending on their house. Note that since households are making more withdrawals, the magnitude of the negative coefficient for loan payments and house payments is consistent with a smaller proportion of withdrawals being for those categories rather than a reduction in the absolute number of withdrawals for those purposes. Standard errors are in parentheses and are clustered at the AWC level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.4: Likelihood of Year 2 Phone Contact by Year 1 Characteristics

	Reached Over Phone
Respondent Education (Years)	0.004
	(0.003)
Scheduled Caste or Tribe	−0.04
	(0.03)
Birth Order	0.01
	(0.01)
Completed Registration	0.001
	(0.004)
Husband Education (Years)	0.01***
	(0.004)
Rural	−0.05
	(0.08)
Hindu	−0.004
	(0.05)
Christian	0.16
	(0.10)
Aadivaasi	−0.03
	(0.06)
Child Female	0.05**
	(0.02)
Age at First Transfer (Months)	0.01
	(0.01)
Below the Poverty Line	−0.05
	(0.03)
Household Size	0.01
	(0.004)
Bank Distance (Kilometers)	−0.002
	(0.001)
Uses Toilet	0.05*
	(0.03)
Observations	1,550

This table examines how likelihood of completing a phone survey during year 2 is related to respondent characteristics (as measured during earlier surveys). Standard errors are in parentheses and clustered at the AWC-level. This sample only includes households with whom we did not conduct a year 2 field survey. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.5: Timing and Measurement of Primary Outcome Variables

	Year 1		Year 2				Year 3
	Phase 1	Phase 2	Phase 1		Phase 2		Phase 1 & 2
	Field	Field	Field	Phone	Phone		Field
	Jan '19	Nov '19	Feb '20	May '20	Sep '20	Nov '20	Nov '21
Anthropometrics & Cognition							
Weight-for-Age (WAZ)	✓	✓	✓				✓
Height-for-Age (HAZ)	✓	✓	✓				✓
Sibling Weight-for-Age (HAZ)			✓				✓
Sibling Height-for-Age (HAZ)							✓
Cognitive & Motor Skills			✓		✓		✓
Expenditures							
Food Spending	✓	✓	✓	✓	✓	✓	
Sin Spending	✓	✓	✓				
Non-Food Spending			✓	✓		✓	
Nutrition							
Caloric Consumption	✓	✓	✓				
Child's Dietary Diversity	✓	✓	✓	✓	✓		✓
Mother's Dietary Diversity	✓	✓	✓	✓	✓		✓
Child's Minimum Meal Frequency	✓	✓	✓	✓	✓		
Mother's Nutrient Index	✓	✓	✓				
Social & Behavioral							
Visitation of Formal Medical Provider	✓	✓	✓	✓		✓	
Illnesses in Past Three Months	✓	✓	✓	✓		✓	
Anganwadi Services	✓	✓	✓		✓		
Mother's Nutritional Knowledge	✓	✓	✓				
Empowerment Index	✓	✓	✓		✓		
Depression Index	✓	✓	✓				

The project was rolled out in two phases, where the treatment began in five districts in March 2018 (labeled "Phase 1") and in another three districts in November 2018 ("Phase 2"). As a result, data collection took place at different times in each phase. For phase 1 districts, there was an in-person survey at the end of year 1 in January of 2019 (column 1) and an in-person survey at the end of year 2 beginning in February of 2020 (column 3). The year 2 survey was halted midway due to COVID-19 and data collection was completed over the phone in May of 2020 (column 4). For phase 2 districts, there was an in-person survey at the end of year 1 in November of 2019 (column 2). Due to COVID-19, the year 2 surveys in phase 2 were conducted over the phone. To reduce respondent fatigue, this was split into two phone surveys: one in September of 2020 (column 5) and one in November of 2020 (column 6). Finally, in both phase 1 and phase 2 districts, a final in-person survey was conducted in November of 2021 (column 7).

Table A.6: Child Development Outcomes: Heterogeneity by Gender

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Year 1 Outcomes				
Treatment	0.04	0.004	0.05	−0.04
	(0.07)	(0.03)	(0.06)	(0.03)
Treatment × Female	−0.05	−0.02	−0.09	−0.01
	(0.10)	(0.04)	(0.09)	(0.04)
Control Mean	−1.47	0.31	−1.67	0.4
Observations	2,355	2,355	2,355	2,355
Panel B: Year 3 Outcomes				
Treatment	0.11	0.02	0.06	−0.02
	(0.08)	(0.02)	(0.06)	(0.02)
Treatment × Female	−0.09	−0.03	−0.05	−0.004
	(0.11)	(0.03)	(0.09)	(0.03)
Control Mean	−0.88	0.19	−1.2	0.23
Observations	2,124	2,124	2,165	2,165
Panel C: Year 3 Outcomes (Sibling)				
Treatment	0.05	−0.01	0.13	−0.01
	(0.09)	(0.02)	(0.08)	(0.03)
Treatment × Female	0.04	−0.01	−0.05	0.02
	(0.14)	(0.04)	(0.12)	(0.04)
Control Mean	−0.9	0.17	−1.49	0.29
Observations	2,046	2,046	2,138	2,138

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on child development outcomes related to the gender of the targeted child. The unit of analysis is the child. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Regressions estimated are fully saturated, with gender dummies excluded for brevity. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.7: Child Development Outcomes: Heterogeneity by Birth Order

	HAZ	Stunted	WAZ	Underweight
	(1)	(2)	(3)	(4)
Panel A: Year 1 Outcomes				
Treatment	−0.04	0.01	−0.05	−0.03
	(0.06)	(0.03)	(0.06)	(0.02)
Treatment × Birth Order	0.04	−0.01	0.04	−0.01
	(0.04)	(0.02)	(0.03)	(0.01)
Control Mean	−1.47	0.31	−1.67	0.4
Observations	2,355	2,355	2,355	2,355
Panel B: Year 3 Outcomes				
Treatment	−0.03	0.01	−0.01	−0.01
	(0.08)	(0.02)	(0.06)	(0.02)
Treatment × Birth Order	0.07	−0.005	0.04	−0.01
	(0.05)	(0.02)	(0.04)	(0.02)
Control Mean	−0.88	0.19	−1.2	0.23
Observations	2,065	2,065	2,105	2,105
Panel C: Year 3 Outcomes (Sibling)				
Treatment	0.03	−0.01	0.04	0.001
	(0.12)	(0.04)	(0.10)	(0.04)
Treatment × Birth Order	0.02	−0.003	0.04	−0.004
	(0.06)	(0.02)	(0.05)	(0.02)
Control Mean	−0.9	0.17	−1.49	0.29
Observations	1,996	1,996	2,086	2,086

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on child development outcomes related to child birth order. The unit of analysis is the child. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Regressions estimated are fully saturated, with birth order excluded for brevity. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table A.8: Child Development Outcomes: Heterogeneity by Age at First Transfer (Months)

	HAZ	Stunted	WAZ	Underweight
	(1)	(2)	(3)	(4)
Panel A: Year 1 Outcomes				
Treatment	−0.01	0.01	−0.02	−0.03
	(0.05)	(0.02)	(0.04)	(0.02)
Treatment × Age	0.004	−0.01	0.03	−0.01
	(0.03)	(0.01)	(0.02)	(0.01)
Control Mean	−1.47	0.31	−1.67	0.4
Observations	2,355	2,355	2,355	2,355
Panel B: Year 3 Outcomes				
Treatment	0.06	−0.01	0.02	−0.02
	(0.07)	(0.02)	(0.05)	(0.02)
Treatment × Age	−0.002	0.004	0.004	0.001
	(0.01)	(0.003)	(0.01)	(0.003)
Control Mean	−0.88	0.19	−1.2	0.23
Observations	2,065	2,065	2,105	2,105

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on child development outcomes related to child age at the time of the first transfer. The unit of analysis is the child. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Regressions estimated are fully saturated, with age at first transfer excluded for brevity. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.9: Nutritional Outcomes: Heterogeneity by Gender

	Household		Child		Mother	
	Caloric consumption (mother and child)	Caloric consumption (per capita)	Dietary diversity score	Minimum meal frequency	Dietary diversity score	Nutrient index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year 1 Outcomes						
Treatment	189.52***	109.53	0.15*	0.01	0.05	0.11***
	(47.75)	(66.59)	(0.08)	(0.03)	(0.03)	(0.03)
Treatment × Female	−17.38	46.57	0.03	0.04	0.06	−0.04
	(73.32)	(96.99)	(0.11)	(0.04)	(0.05)	(0.05)
Control Mean	1873.54	1678.39	1.59	0.61	2.78	0.19
Observations	2,360	2,321	2,349	2,344	2,349	2,349
Panel B: Year 2 Outcomes						
Treatment	343.89**	20.47	0.27***	0.03	0.24***	0.20***
	(137.56)	(106.30)	(0.08)	(0.02)	(0.07)	(0.06)
Treatment × Female	87.66	155.67	−0.06	−0.03	−0.04	−0.06
	(251.32)	(176.99)	(0.12)	(0.04)	(0.10)	(0.10)
Control Mean	2532.01	1499.86	3.4	0.91	3.35	0.22
Observations	596	564	1,421	960	1,421	596

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on nutritional outcomes related to child gender. The unit of analysis is either the mother or child, and this is indicated in the first row of the table. Regressions estimated are fully saturated, with gender dummies excluded for brevity. Dietary diversity score (0-7) is defined as the number of the following dietary groups that were consumed by the respondent: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables. Nutrient index is computed by estimating the percent of the recommended daily value of calories, protein, visible fat, calcium, iron, thiamine, riboflavin, niacin, pyridoxin, and dietary folate consumed and averaging. For the year 2 outcomes in panel B, Columns (1), (2), and (6) use field survey data, as it is necessary to measure the quantity of individual ingredients to create these measures. Columns (3), (4), and (5) pool data from the field and multiple rounds of phone surveys (see table A.5 for a visualization of each of the outcomes of interest and the survey round in which they were collected). Column (4) pertains to the subset of the sampled children who are under 23 months — i.e. those for whom minimum meal frequency is considered a clinically meaningful measure. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.10: Nutritional Outcomes: Heterogeneity by Birth Order

	Household		Child		Mother	
	Caloric consumption (mother and child)	Caloric consumption (per capita)	Dietary diversity score	Minimum meal frequency	Dietary diversity score	Nutrient index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year 1 Outcomes						
Treatment	236.16***	115.33*	0.19***	0.06**	0.07**	0.16***
	(52.32)	(65.51)	(0.07)	(0.03)	(0.03)	(0.04)
Treatment × Birth Order	−44.69	6.75	−0.02	−0.02	0.01	−0.06**
	(29.35)	(36.20)	(0.04)	(0.02)	(0.02)	(0.02)
Control Mean	1873.54	1678.39	1.59	0.61	2.78	0.19
Observations	2,358	2,319	2,347	2,344	2,347	2,347
Panel B: Year 2 Outcomes						
Treatment	329.06***	109.63	0.17**	0.03	0.18***	0.14**
	(122.58)	(127.25)	(0.08)	(0.03)	(0.07)	(0.06)
Treatment × Birth Order	52.35	−12.33	0.06	−0.01	0.03	0.02
	(66.80)	(64.04)	(0.04)	(0.02)	(0.04)	(0.03)
Control Mean	2532.01	1499.86	3.4	0.91	3.35	0.22
Observations	596	564	1,421	960	1,421	596

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on nutritional outcomes related to birth order. The unit of analysis is either the mother or child, and this is indicated in the first row of the table. Regressions estimated are fully saturated, with birth order excluded for brevity. Dietary diversity score (0-7) is defined as the number of the following dietary groups that were consumed by the respondent: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables. Nutrient index is computed by estimating the percent of the recommended daily value of calories, protein, visible fat, calcium, iron, thiamine, riboflavin, niacin, pyridoxin, and dietary folate consumed and averaging. For the year 2 outcomes in panel B, Columns (1), (2), and (6) use field survey data, as it is necessary to measure the quantity of individual ingredients to create these measures. Columns (3), (4), and (5) pool data from the field and multiple rounds of phone surveys (see table A.5 for a visualization of each of the outcomes of interest and the survey round in which they were collected). Column (4) pertains to the subset of the sampled children who are under 23 months — i.e. those for whom minimum meal frequency is considered a clinically meaningful measure. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.11: Nutritional Outcomes: Heterogeneity by Age at First Transfer (Months)

	Household		Child		Mother	
	Caloric consumption (mother and child)	Caloric consumption (per capita)	Dietary diversity score	Minimum meal frequency	Dietary diversity score	Nutrient index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year 1 Outcomes						
Treatment	171.64***	112.06**	0.15***	0.03	0.09***	0.09***
	(37.72)	(46.47)	(0.05)	(0.02)	(0.03)	(0.03)
Treatment × Age	13.44	23.35	0.02	0.0003	−0.01	0.002
	(17.47)	(22.17)	(0.03)	(0.01)	(0.01)	(0.01)
Control Mean	1873.54	1678.39	1.59	0.61	2.78	0.19
Observations	2,358	2,319	2,347	2,344	2,347	2,347
Panel B: Year 2 Outcomes						
Treatment	−588.13	−788.50	0.29***	0.04*	0.30***	−0.21
	(515.29)	(754.24)	(0.08)	(0.02)	(0.08)	(0.25)
Treatment × Age	78.32*	69.44	−0.01	−0.01	−0.01*	0.03
	(41.14)	(59.07)	(0.01)	(0.01)	(0.01)	(0.02)
Control Mean	2532.01	1499.86	3.4	0.91	3.35	0.22
Observations	596	564	1,421	960	1,421	596

This table follows the pre-analysis plan and tests for heterogeneity in treatment effects on nutritional outcomes related to the age of the child at the time of the first transfer. The unit of analysis is either the mother or child, and this is indicated in the first row of the table. Regressions estimated are fully saturated, with age at first transfer excluded for brevity. Dietary diversity score (0-7) is defined as the number of the following dietary groups that were consumed by the respondent: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables. Nutrient index is computed by estimating the percent of the recommended daily value of calories, protein, visible fat, calcium, iron, thiamine, riboflavin, niacin, pyridoxin, and dietary folate consumed and averaging. For the year 2 outcomes in panel B, Columns (1), (2), and (6) use field survey data, as it is necessary to measure the quantity of individual ingredients to create these measures. Columns (3), (4), and (5) pool data from the field and multiple rounds of phone surveys (see table A.5 for a visualization of each of the outcomes of interest and the survey round in which they were collected). Column (4) pertains to the subset of the sampled children who are under 23 months — i.e. those for whom minimum meal frequency is considered a clinically meaningful measure. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table A.12: Anthropometric Outcomes with LASSO Controls

	HAZ	Stunted	WAZ	Underweight
	(1)	(2)	(3)	(4)
Panel A: Year 1 Outcomes				
Treatment	0.001	−0.001	−0.01	−0.03*
	(0.04)	(0.02)	(0.04)	(0.02)
Control Mean	−1.47	0.31	−1.67	0.4
Observations	2,269	2,041	2,041	2,269
Panel B: Year 3 Outcomes				
Treatment	0.06	0.002	0.04	−0.02
	(0.05)	(0.02)	(0.04)	(0.02)
Control Mean	−0.88	0.19	−1.2	0.23
Observations	1,937	1,965	1,974	2,012

This table adds controls selected via LASSO to Table 4. This reduces the number of observations relative to Table 4 due to missing values for some of the selected variables. The unit of analysis is the child. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Regressions estimated are fully saturated, with age at first transfer excluded for brevity. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.13: Nutrients Index Components (Macro-Nutrients)

	Iron	Protein	Visible Fat	Calcium
	(1)	(2)	(3)	(4)
Panel A: Year 1 Outcomes				
Treatment	0.05***	0.04***	0.02***	0.004
	(0.01)	(0.01)	(0.01)	(0.01)
Control Mean	0.41	0.67	0.17	0.18
Observations	2,360	2,360	2,360	2,360
Panel B: Year 2 Outcomes				
Treatment	0.07***	0.07***	0.04***	0.003
	(0.02)	(0.02)	(0.01)	(0.02)
Control Mean	0.38	0.68	0.17	0.24
Observations	603	603	603	603

This table reports the effect of the treatment on individual macro-nutrients. The unit of analysis is the mother. Outcome variables correspond to the percent of daily value consumed for each macro-nutrient used in the construction of the nutrients index (see Table 2). Observations in Year 2 correspond to only the subset of respondents who were sampled in the field. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.14: Take-Up of AWC Services

	Preschool	Hot cooked meals	Deworming	Growth measurement	Nutrition information	Iron/calcium tablets	Vaccination	Take-home rations	Government schemes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Year 1 Outcomes									
Treatment	−0.01	−0.003	0.06***	0.02	0.09***	0.04*	−0.002	0.001	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
Control Mean	0.3	0.33	0.42	0.7	0.5	0.59	0.73	0.88	0.62
Observations	2,329	2,320	2,325	2,329	2,326	2,327	2,329	2,330	2,323
Panel B: Year 2 Outcomes									
Treatment	0.03	0.03	−0.003	0.07**	0.09***	−0.0005	−0.03	0.03	−0.004
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
Control Mean	0.37	0.44	0.57	0.57	0.31	0.36	0.26	0.85	0.3
Observations	586	586	1,066	1,066	1,066	1,066	1,066	1,066	1,066

This table presents estimates for the proportion of respondents that take up various AWC services each year. "Deworming" refers to whether the mother took her child to receive deworming medication. "Government schemes" refers to visiting one's AWC to sign up for other government services or programs. "Measurement" refers to whether the child had their height and weight measured at the AWC. "Hot cooked meals," "iron/calcium tablets", and "nutrition information" refer to whether the mother (and her child) received these items or services, the latter two during pregnancy and lactation. "Preschool" refers to whether the child received pre-school services administered by the AWC. "Take-home rations" refers to whether the mother received these from her AWC. Finally, "vaccination" refers to whether the child visited the AWC to receive immunizations for her child or herself during pregnancy. In panel B, we did not ask about pre-schools or hot cooked meals during the year 2 phone survey since AWCs did not offer them at that time due to the covid-19 pandemic. As a result, the observation counts are lower for the first two columns. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.15: Child Development Index Components

	Child can say their own name	Child can say the names of others	Child can identify their village	Child can tell the day of the week	Child can count to 5	Child can count to 10	Child can count to 20	Child can name primary colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Cognitive Outcomes								
Treatment	0.32**	0.08	0.25	−0.02	0.42**	0.30*	0.14	0.14
	(0.15)	(0.14)	(0.20)	(0.07)	(0.19)	(0.17)	(0.09)	(0.19)
Control Mean	7.6	8.54	5.43	0.24	5.58	2.57	0.64	3.62
Observations	2,164	2,164	2,164	2,164	2,164	2,164	2,164	2,163

	Child can walk up stairs	Child can kick a ball	Child can catch a ball	Child can serve themselves food	Child can brush their teeth
	(9)	(10)	(11)	(12)	(13)
Panel B: Gross Motor Skills					
Treatment	0.11	0.17	−0.11	0.60***	0.15
	(0.12)	(0.13)	(0.18)	(0.18)	(0.11)
Control Mean	8.65	8.56	2.77	6.47	9.09
Observations	2,164	2,164	2,162	2,163	2,164

	Child can unbutton their clothes	Child can hold a pen correctly	Child can draw a basic figure	Child can fold paper
	(14)	(15)	(16)	(17)
Panel C: Fine Motor Skills				
Treatment	0.09	0.32*	0.12**	0.37**
	(0.18)	(0.18)	(0.06)	(0.19)
Control Mean	7.46	6.53	0.17	2.02
Observations	2,162	2,163	2,164	2,163

The unit of the analysis is the child, and outcomes are measured in the year 3 survey. Outcomes are indices taking value 0 if the child does not ever, 5 if the child sometimes does, and 10 if the child does the listed action. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.16: Child Development Sub-Indices

	Child development index	Gross motor skills index	Cognitive index	Fine motor skills index
	(1)	(2)	(3)	(4)
Treatment	0.12***	0.09**	0.09**	0.10**
	(0.04)	(0.04)	(0.04)	(0.04)
Control Mean	−0.06	−0.05	−0.04	−0.05
Observations	2,164	2,162	2,163	2,161

The unit of analysis is the child, and outcomes are measured in the year 3 survey. Indices are demeaned and expressed in standard deviation units. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.17: Heterogeneity by Reporting of Mobile Phone Number Conditional on Assets

	Nutritional knowledge index	Dietary diversity (child)	Dietary diversity (mother)	Nutrient index
	(1)	(2)	(3)	(4)
Panel A: Year 1 Outcomes				
Treatment	0.19***	0.22***	0.09***	0.11***
	(0.07)	(0.05)	(0.03)	(0.03)
Treatment × Mobile Number | Assets	−0.13	0.20*	−0.05	0.04
	(0.13)	(0.12)	(0.05)	(0.06)
Control Mean	3.42	1.59	2.78	0.19
Observations	2,158	2,158	2,158	2,158
Panel B: Year 2 Outcomes				
Treatment	0.33***	0.27***	0.24***	0.18***
	(0.12)	(0.06)	(0.05)	(0.05)
Treatment × Mobile Number | Assets	−0.21	0.14	0.04	0.01
	(0.26)	(0.12)	(0.10)	(0.11)
Control Mean	3.57	3.4	3.35	0.22
Observations	559	1,327	1,327	553

This table investigates whether the treatment has heterogeneous effects based on whether the mother registers a mobile phone number, meaning that that she is able to receive IVR calls with messaging on nutrition. The unit of analysis is the child or mother, and is indicated in the first row of the table. Mobile number refers to whether the respondent reported a mobile phone number in initial surveys. The interaction variable corresponds to residuals from a regression of mobile phone on an asset index. Regressions estimated are fully saturated, with dummy the interaction variable excluded for brevity. Standard errors are in parentheses and clustered at the AWC-level. In panel B, columns (1) and (4) use only field survey data, while columns (2) and (3) combine field and phone survey data. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.18: Sibling Anthropometrics

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Year 2 Outcomes				
Treatment			0.13	−0.10**
			(0.10)	(0.04)
Control Mean			−1.45	0.32
Observations			440	440
Panel B: Year 3 Outcomes				
Treatment	0.07	−0.01	0.11**	−0.002
	(0.06)	(0.02)	(0.05)	(0.02)
Control Mean	−0.9	0.17	−1.49	0.29
Observations	2,046	2,046	2,138	2,138

The unit of analysis is the child. HAZ and WAZ denote the sibling's height-for-age and weight-for-age z-scores, respectively. Children with HAZ and WAZ of less than -2 are classified as moderately stunted and underweight respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.19: Sibling Anthropometrics and Interactions with Neighborhood Sanitation Environment

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Implied Effects by Percentile				
10%	-0.07	0.05*	0.12	0.03
25%	-0.03	0.03	0.12*	0.02
50%	0.06	-0.01	0.11**	0
75%	0.14**	-0.04**	0.11*	-0.02
90%	0.26**	-0.09***	0.1	-0.05
Test of Equality: p-value	0.06	0.01	0.9	0.17
Joint p-value	0.01			
Panel B: Underlying Parameter Estimates				
Treatment	-0.12	0.07**	0.12	0.04
	(0.12)	(0.03)	(0.12)	(0.04)
Poor Sanitation Index	-0.26	0.07	-0.13	0.02
	(0.18)	(0.05)	(0.14)	(0.06)
Treatment × Poor Sanitation Index	0.40*	-0.17***	-0.02	-0.10
	(0.24)	(0.07)	(0.20)	(0.07)
Control Mean	-0.9	0.17	-1.49	0.29
Observations	2046	2046	2138	2138

This table tests for treatment effects on sibling anthropometrics. In the year 2 endline survey, we randomly selected one sibling under the age of 10 and measured their weight and height. In the year 3 endline survey, we randomly selected up to three of the focal child's siblings under the age of 10 and measured their height/weight. HAZ and WAZ denote the sibling's height-for-age and weight-for-age z-scores, respectively. Children with HAZ and WAZ of less than -2 are classified as moderately stunted and underweight respectively. As a summary measure of sanitation environment, we constructed a principal components index from 8 variables such as whether the household or their neighbors use a toilet. The poor sanitation index used in this regression is equal to the household's percentile rank on that principal components index, where a value of zero is equal to the best sanitation environment (lowest incidence of open defecation) and one is the worst. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.20: Child Development Outcomes (Year 2)

	HAZ	Stunted	WAZ	Underweight
	(1)	(2)	(3)	(4)
Treatment	−0.10	0.03	0.02	−0.02
	(0.09)	(0.03)	(0.08)	(0.04)
Control Mean	−1.92	0.48	−1.91	0.46
Observations	607	607	607	607

This table examines outcomes collected in year 2, when our field survey was interrupted by the COVID-19 pandemic. As a result of this disruption, sample sizes and power are significantly lower than the year 1 or year 3 analysis. The unit of analysis is the child. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.21: Interactions with Sanitation Environment (Open Defecation)

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Implied Effects by Percentile				
10%	0.16*	-0.04	0.15**	-0.04
25%	0.13*	-0.03	0.12**	-0.04
50%	0.08	0	0.05	-0.02
75%	0.02	0.02	-0.01	0
90%	-0.06	0.06*	-0.11	0.02
Test of Equality: p-value	0.21	0.04	0.05	0.21
Joint p-value	0.02			
Panel B: Underlying Parameter Estimates				
Treatment	0.19*	-0.06*	0.18**	-0.05*
	(0.10)	(0.03)	(0.08)	(0.03)
Poor Sanitation Index	-0.28**	-0.02	-0.12	0.02
	(0.14)	(0.05)	(0.12)	(0.05)
Treatment × Poor Sanitation Index	-0.26	0.12**	-0.32**	0.08
	(0.19)	(0.06)	(0.14)	(0.06)
Control Mean	-0.88	0.19	-1.2	0.23
Observations	2124	2124	2165	2165

The unit of analysis is the child. This table differs from table 5 in looking at an index of only variables directly related to open defecation rather than all sanitation variables. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.22: Interactions with Sanitation Environment: Nutrition, Expenditure, & Empowerment (Year 1)

	Total food expenditure (1)	Caloric consumption (mother and child) (2)	Caloric consumption (per capita) (3)	Child's dietary diversity score (4)	Nutritional knowledge index (5)	Anganwadi services received (6)	Empowerment index (7)
Panel A: Implied Effects by Percentile							
10%	0.14*	217.57***	235.07***	0.32***	0.07	0.24**	0.15*
25%	0.15**	210.79***	204.37***	0.29***	0.09	0.21**	0.14*
50%	0.15***	193.89***	127.89***	0.2***	0.15**	0.13*	0.1*
75%	0.16**	183.55***	81.05	0.15**	0.19**	0.08	0.07
90%	0.16	163.08**	-11.58	0.05	0.26**	-0.01	0.02
Test of Equality: p-value	0.9	0.63	0.09	0.1	0.33	0.24	0.41
Joint p-value	0.02						
Panel B: Underlying Parameter Estimates							
Treatment	0.14	226.48***	270.15***	0.36***	0.04	0.28*	0.17
	(0.09)	(70.66)	(93.06)	(0.11)	(0.13)	(0.15)	(0.11)
Poor Sanitation Index	-0.23*	174.80*	207.90	-0.25	-0.50***	0.14	0.28*
	(0.14)	(100.91)	(151.53)	(0.15)	(0.18)	(0.21)	(0.14)
Treatment × Poor Sanitation Index	0.02	-66.83	-302.48	-0.34	0.24	-0.31	-0.16
	(0.19)	(141.94)	(189.27)	(0.21)	(0.25)	(0.28)	(0.20)
Control Mean	8.35	1895.43	1678.17	1.59	3.43	3.93	2.33
Observations	2095	2095	2066	2086	2086	2037	2086

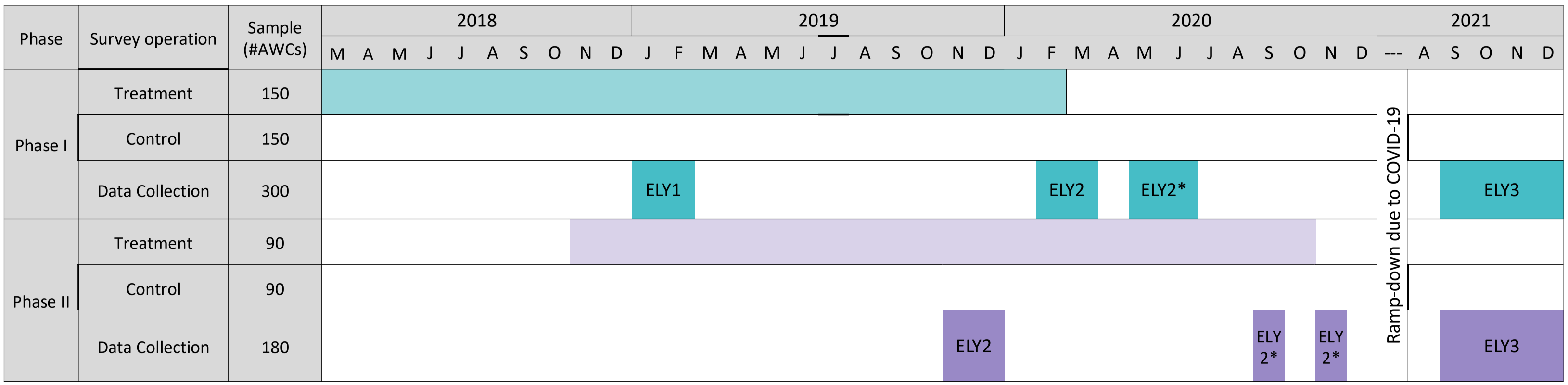
This table tests for heterogeneity in treatment effects on intermediate outcomes (e.g., nutrition, expenditure) related to the sanitation environment. The unit of analysis is either the mother, child, or household. Outcomes correspond to a pooled sample of observations from field surveys in P1 and P2 of Y1. As a summary measure of sanitation environment, we constructed a principal components index from 8 variables such as whether the household or their neighbors use a toilet. The poor sanitation index used in this regression is equal to the household's percentile rank on that principal components index, where a value of zero is equal to the best sanitation environment (lowest incidence of open defecation) and one is the worst. Total food expenditure is subjected to an inverse hyperbolic sine transformation. Variety score (0-20) measures variety across both food groups and sources of protein. Dietary diversity score (0-7) is defined as the number of the following dietary groups that were consumed by the respondent: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables. The nutritional knowledge (0-6) and empowerment (0-5) indices correspond to the number of questions answered by the respondent that suggest a high degree of the corresponding characteristic. Anganwadi services received corresponds to the number of the following that the respondent received from their local AWC: e-school, deworming, government schemes, growth measurement, hot cooked meals, iron/calcium tablets, nutritional information, take-home rations, and vaccination. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.23: Interactions with Sanitation Environment (Residualized)

	HAZ (1)	Stunted (2)	WAZ (3)	Underweight (4)
Panel A: Implied Effects by Percentile				
10%	0.26	-0.09*	0.25**	-0.07
25%	0.23*	-0.07*	0.21**	-0.06
50%	0.08	-0.01	0.06	-0.02
75%	-0.04	0.05*	-0.08	0.01
90%	-0.15	0.1**	-0.19	0.05
Test of Equality: p-value	0.19	0.04	0.07	0.23
Joint p-value	0.03			
Panel B: Underlying Parameter Estimates				
Treatment	0.39	-0.15*	0.39**	-0.11
	(0.25)	(0.08)	(0.20)	(0.08)
Poor Sanitation Index (Residuals)	0.08	-0.17	0.29	-0.07
	(0.31)	(0.11)	(0.27)	(0.11)
Treatment × Poor Sanitation Index (Residuals)	-0.63	0.28*	-0.67*	0.18
	(0.48)	(0.15)	(0.37)	(0.15)
Control Mean	-0.88	0.19	-1.2	0.23
Observations	2124	2124	2165	2165

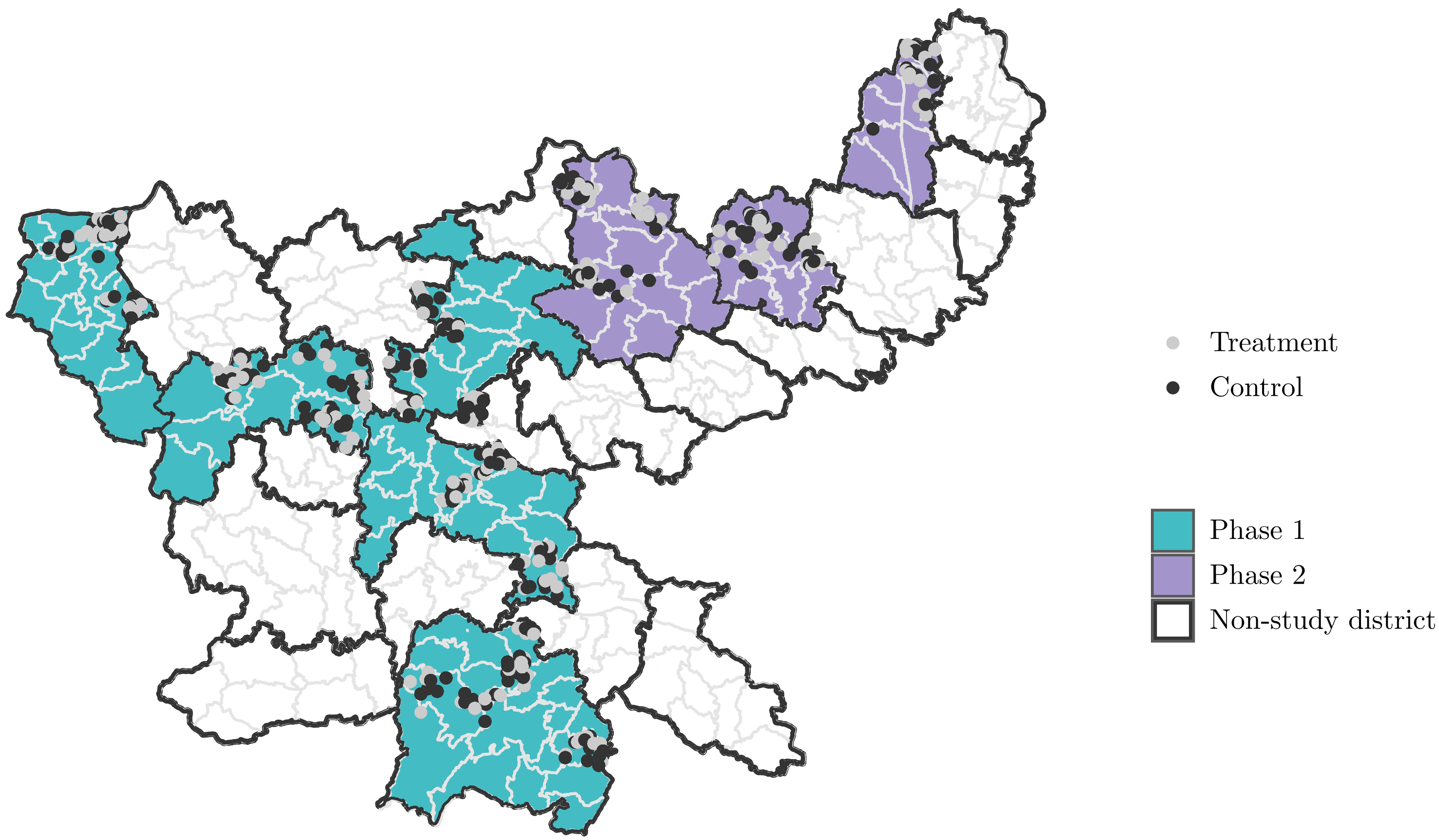
The unit of analysis is the child. This table takes the sanitation index from table 5, regresses it on variables related to sanitation, and saves the residuals (see section 4 for details). It then uses those residuals in the interaction term. WAZ and HAZ denote the child's weight-for-age and height-for-age z-scores, respectively. Children with WAZ and HAZ of less than minus two are classified as moderately underweight and stunted, respectively. Standard errors are in parentheses and clustered at the AWC-level. $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Figure A.1: Project Timeline



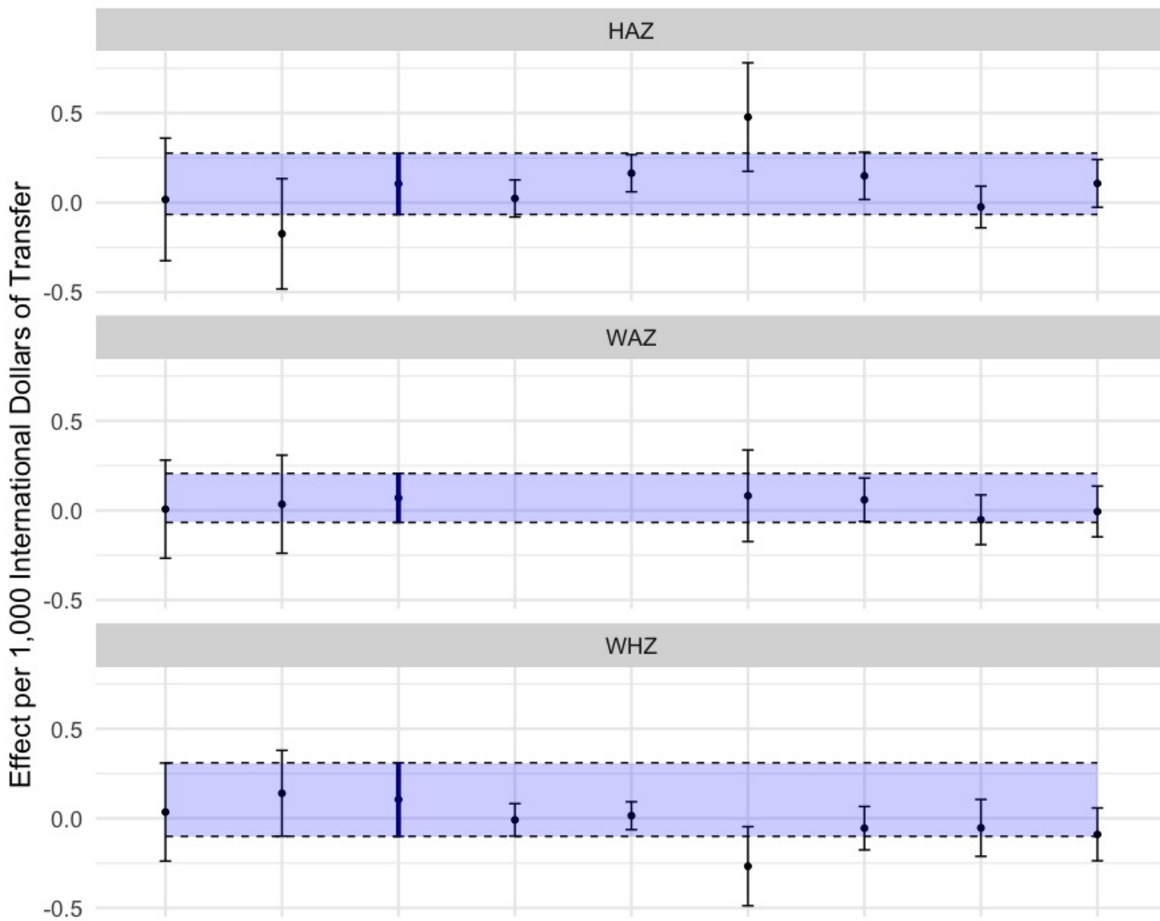
This figure shows how the intervention and data collection were carried out from 2018 to 2021. The project consisted of two separate "phases" – five initial districts (phase 1) and three additional districts (phase 2) in which implementation began approximately six months after the original set of districts. Data collection blocks marked with an asterisk (*) were conducted over the phone rather than in person. Months in which project activity ceased due to the COVID-19 pandemic have been omitted from this figure. In both phases, there was no baseline survey, and an in-person endline survey were conducted after one year of implementation. In phase I, an in-person, year two endline survey was halted midway through March 2020 due to COVID-19, and a phone survey was conducted from May to June of 2020 for the remainder of the sample. In phase II, a year two endline survey was conducted via phone, but split into two separate surveys due to concerns over survey length. In both phases, an in-person survey was conducted in September to December of 2021.

Figure A.2: Study Map



This figure shows the sub-districts of Jharkhand in which this study was conducted. Shaded blocks are those included in our study, with the two colors corresponding to two phases of treatment rollout (see Figure A.1). Dark lines denote district boundaries, and light lines denote Taluks, or sub-districts. Points represent the AWCs in our sample, with the color and shape of each corresponding to the different treatment groups the AWCs were assigned to.

Figure A.3: Effects of Unconditional Cash Transfers on Child Development Outcomes



Study	JHICDS	JHICDS	JHICDS	Ahmed	Ahmed	Carneiro	Carneiro	Field	Field
Country	India	India	India	Bangladesh	Bangladesh	Nigeria	Nigeria	Myanmar	Myanmar
Transfer Length	12 Months	24 Months	24 Months	24 Months	24 Months	12 Months	24 Months	24 Months	24 Months
Communication	Framing	Framing	Framing	NA	SBCC	Information	Information	NA	SBCC

This figure displays the estimated effects of unconditional cash transfers on Child Development Outcomes for several studies similar to Jharkhand ICDS. Cash transfer sizes are converted from local currency using World Bank PPP adjustments, calculated cumulatively (i.e. transfer size times total months of transfer), and used to divide point estimates and standard errors for the effects of treatment on HAZ, WAZ, and WHZ. The total length of time that transfers were received and information on any communication received are presented in the key. The studies considered in addition to ICDS are as follows: Ahmed et al. (2019), Carneiro et al. (2021), and Field and Maffioli (2021). The vertical range of the blue rectangles corresponds to the standardized confidence interval for Jharkhand ICDS Y3. WAZ is not reported as an outcome in Ahmed et al. (2019) and is thus excluded.

B Appendix: Simulated Power Calculations

Section 3.2 estimates whether there is an interaction between sanitation and income in the production of child anthropometrics. Although the cash transfer is exogenous, the sanitation variation is cross-sectional, requiring us to interpret the interaction in a nuanced way.

A stronger experimental design to test for interactions would also induce exogenous variation in sanitation, such as the sanitation promotion campaign studied in Patil et al. (2014). Such an experiment would have three treatment arms: one receiving cash transfers, one receiving sanitation promotion, and one receiving both. We would then test for complementarities between income and sanitation intervention with:

$$Y_{ia} = \alpha + \beta_1 * TREAT^1_{ia} + \beta_2 * TREAT^2_{ia} + \beta_3 * TREAT^1_{ia} * TREAT^2_{ia} \quad (2)$$

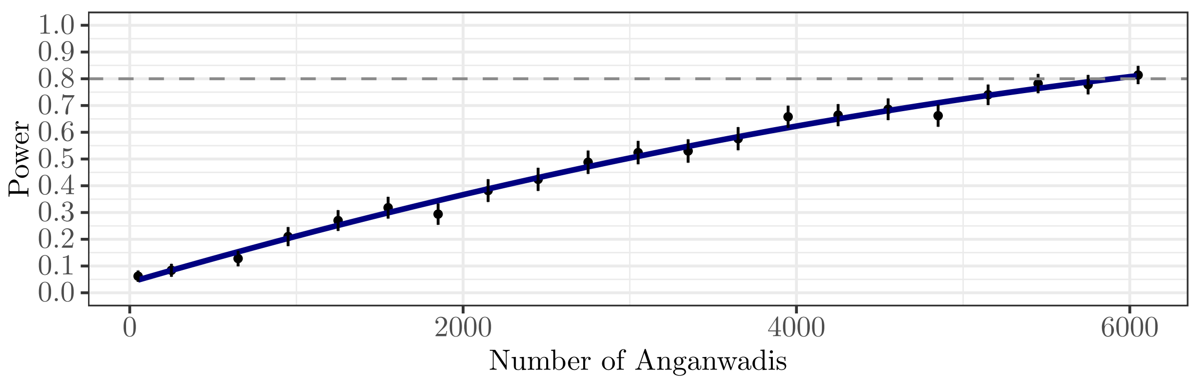
where $TREAT^1_{ia}$ and $TREAT^2_{ia}$ are dummy indicators for the cash and sanitation interventions, respectively. We are interested in evaluating the power of a test of $H_0 : \beta_1 + \beta_2 = \beta_3$ vs. $H_1 : \beta_1 + \beta_2 \neq \beta_3$, given a number of plausible design characteristics. However, detecting interactions has significantly greater power requirements than testing for level effects (Muralidharan et al., 2023), so we conduct power calculations to determine whether such an experiment would be feasible.

We simulate this data generating process, varying the number of Anganwadi centers that are evenly divided amongst four treatment arms (each with five households, our sample average). Household WAZ is generated by summing the control mean for WAZ as reported in household surveys with random draws from normal distributions with intra-cluster and inter-cluster variances corresponding to those observed in our sample. Treatment effects are simulated by assigning to each household a baseline level of toilet non-usage drawn from the empirical distribution of that variable as collected in household surveys and subtracting from this rate — in the relevant two treatment arms— the treatment effect as reported in Patil et al. (2014). We then estimate the treatment effect based on the estimates of treatment effects in our sample, add this to the simulated WAZ for the relevant treatment groups, and run the test described above.

We conduct this procedure for a variety of plausible Anganwadi counts, in each case running the simulation 500 times (see Figure B.1). Results suggest that in order to achieve 80% power, such an experiment would require coverage of approximately 6,000 Anganwadis.

This is an order of magnitude higher than most field experiments recently conducted in similar contexts (Field and Maffioli, 2021; Carneiro et al., 2021; Ganimian et al., 2021). It is also over 12 times larger than the number of anganwadis in our current study. In addition to the much larger budget that such a study would entail, it would also be logistically much more challenging to ensure implementation quality and design fidelity across multiple treatment arms at such large scales. Hence, budgetary and practical constraints would likely preclude a fully experimental test of the hypothesis of significant interactions between income and sanitation in the child health production function. As a result, the evidence we present – of a significant interaction between experimentally varied income transfers and cross-sectional variation in sanitation in the child health production function (along with additional analysis to confirm that this is not being driven by interactions of income with other observable correlates with sanitation) – probably provides the most practically feasible experimental evidence on the existence of this interaction.

Figure B.1: Simulated Power Curve for a Cash × Sanitation RCT



C Appendix: Variable Construction

This appendix presents detailed information on the construction of variables in this study, including technical material motivating indices or measures with clinical justification.

C.1 In-person measures of food consumption

In order to accurately gauge household consumption of nutrients, we used an intensive measurement protocol taken from the nutrition literature. We conducted a 24-hour recall history with the household in which they were instructed to list everything that they had eaten during the previous calendar day. The next steps were:

- Collect measurements of ingredients used in every food item. For each food item, respondents were asked to list every raw ingredient used to cook the item. We then measured the amount of that ingredient that they used by measuring the capacity of the utensils used to put in each ingredient.

  For example, suppose the respondent had used a particular ladle to measure out raw yellow daal for their dinner on the previous day. We would take the ladle they used to measure out the daal, fill it with water, and then pour that water into a beaker to measure ladle capacity in milliliters. We would then ask how many ladle servings they had put in (including partial servings) and convert that to an amount used of that ingredient. Using data from Longvah et al. (2017) on the nutritional content of raw ingredients used in Indian cooking, we then converted the raw ingredients to nutrient amounts (e.g. 100 grams of "black gram, daal" contains 23.1 grams of protein, 1.7 grams of fat, etc.).

- Collect measurements of the total amount of cooked food for every meal cooked yesterday. This means measuring the capacity of the vessels (in milliliters) in which food was cooked using the same methodology of pouring water into a beaker. We then record the level until which these vessels were filled after the item was cooked and converted that to a total volume cooked.

- Collect measurements of the share eaten of each cooked item by the mother and child. This meant asking the respondent to show how much they ate out of the cooked items in the unit of serving utensils (e.g. three spoonfuls of cooked yellow daal). We convert that to an amount in milliliters by using the same technique to measure the capacity of the utensils used to serve the food.

We combine these three measures to derive a precise measurement of the amount consumed by individual household members (mother and child). This is equal to the total share of cooked item consumed multiplied by the nutritional content of the ingredients for that item, summed across all the cooked items. Additionally, we collected measurements of any complementary feeding of the child (other than breastmilk) such as biscuits or snacks that were not produced at home. We convert those to nutrients using approximations of the nutritional content of those items. The full set of instructions is available at this link.

C.2 Nutritional Measures

- **Dietary diversity score**: Computed as the number of the following dietary groups consumed by the respondent or child: (1) grains, roots, and tubers, (2) legumes and nuts, (3) dairy products, (4) flesh foods, (5) eggs, (6) vitamin-A rich fruits and vegetables, (7) other fruits and vegetables; based on and calculated in accordance with the WHO's (2008) "Indicators for assessing infant and young child feeding practices"

- **Minimum meal frequency**: Binary indicator taking value 1 if the child meets WHO (2008) guidelines for the minimum number of times that children should consume solid, semi-solid, or soft foods and 0 otherwise; minimum frequency is twice per day for breastfed infants 6-8 months, three times per day for breastfed children 9-23 months, and four times per day for non-breastfed children 6-23 months
- **Nutrient index**: Taking the recommended daily value (as prescribed by the National Institute of Nutrition (2011) for calories, protein, visible fat, calcium, iron, thiamine, riboflavin, niacin, pyridoxine, and dietary folate, we calculated the fraction of this that respondents consumed and then averaged across all items. Foods were translated into nutrients using the Indian Food Consumption Tables as described above (Longvah et al., 2017).

C.3 Social & Behavioral Measures

- **Nutritional knowledge index**: Discrete variable taking values between 0 to 6, corresponding to the number of the following questions or statements that the mother answers or responds to "correctly", i.e. in line with clinical recommendations:
  - How much should you eat during pregnancy: more than normal, the same amount as normal, or less than normal?
  - How much should you eat while breastfeeding: more than normal, the same amount as normal, or less than normal?
  - Eating more during pregnancy affects child intelligence.
  - Eating more during pregnancy affects child height.
  - Eating more as a child affects child intelligence.
  - Eating more as a child affects child height.
- **Anganwadi services received**: Discrete variable taking values between 0 to 9 corresponding to the number of the following AWC services that the respondent received in the previous year: (1) deworming, (2) government schemes, (3) growth measurement, (4) hot cooked meals, (5) iron/calcium tablets, (6) nutrition information, (7) pre-school, (8) take-home rations, (9) vaccination
- **Empowerment index**: Discrete variable taking values between 0 to 5 corresponding to the number of the following questions adapted from J-PAL's "Practical Guide to Measuring Women's and Girls' Empowerment in Impact Evaluations" (Glennerster et al., 2018) that indicate respondent empowerment:
  - The last time you went to a relative or acquaintence's house inside the village, did you have to take permission from other members of your household?
  - The last time you went to the market without your village, did you have to take permission from other members of your household?
  - Do you have to ask someone for money if you want to purchase items from the market?
  - Imagine that you were home alone without your spouse or guardian and one of your children was very sick. Could you make the choice on your own to purchase medication to treat your child?

- Suppose you earned Rs 300 as part of a government program. Who would decide how to spend it?

- **Depression index**: Discrete variable taking integer values from 0-5, corresponding to the number of the following questions from the Patient Health Questionnaire-9 that are answered in a way indicating the presence of depressive disorders:
  - In the last 2 weeks, how often have you felt nervous or stressed?
  - Often there are multiple tasks that you have to do in a day like cooking, cleaning, taking care of your child, etc. In the last two did you feel that you couldn't manage all these tasks?
  - In the last 2 weeks, how often did you have trouble falling or staying asleep, or sleeping too much?
  - In the last 2 weeks, how often were you feeling tired or having little energy?
  - In the last 2 weeks, how often were you having trouble concentrating on things?

- **Probability of visiting a formal medical provider**: Binary indicator that takes value 1 if the child has visited a government doctor/hospital/clinic/PHC, private doctor/hospital/clinic, or an ANM/sub-centre and 0 otherwise

- **Total illnesses in the past three months**: Discrete index corresponding to the number of the following distinct illnesses or ailments experienced by the child in the previous three months: (1) cold/cough/fever, (2) diarrhea/vomiting/stomach infections, (3) malaria/jaundice/dengue/other vector-borne diseases, (4) measles/chickenpox, (5) pneumonia, (6) physical injuries/fractures, (7) other illnesses not listed here

C.4 Anthropometric Measures

- **HAZ**: Height-for-age z-score; computed using Stata's **zanthro** command (Vidmar et al., 2013) and the WHO Child Growth Charts (2006)

- **WAZ**: Weight-for-age z-score; computed using Stata's **zanthro** command (Vidmar et al., 2013) and the WHO Child Growth Charts (2006)

- **Moderately stunted**: Binary variable that takes value 1 if a child has HAZ $<$ -2 and 0 otherwise

- **Moderately wasted**: Binary variable that takes value 1 if a child has WHZ $<$ -2 and 0 otherwise

- **Severely stunted**: Binary variable that takes value 1 if a child has HAZ $<$ -3 and 0 otherwise

- **Severely wasted**: Binary variable that takes value 1 if a child has WHZ $<$ -3 and 0 otherwise

- **Child development index**: Index taking values 0-90 (Year 2) or 0-170 (Year 3) computed by summing scores associated with one of the following lists of questions asked to respondents, where 10 is granted for "Yes", 5 is granted for "Sometimes", and 0 is granted for "No"
  - **Year 2 Questions**
  - Does your child run, stopping herself and without bumping into things or falling over?

- Does your child climb on furniture?
- Can your child remove clothes on her own without your help?
- When your child wants something does she tell you by pointing to it?
- When you ask your child to, does she go into another room and find familiar objects or toys? For example you might ask your child to "Bring water" or "Go get your chappal (sandals)"
- Does your child say words other than "Mama" and "Papa?" For example words may include "Bakri (goat)" or "Gai (cow)" or "Kaan (ear)" or "Naak (nose)"
- Does your child say short sentences? Such as "Khaana do (Give me food)" or "Mama paani do (Mama, give me water)" or "Yeh kya hai? (What is this?)" or "Mera haath pakdo (Hold my hand)"
- Does your child scribble?
- Has your child started eating food on her own?
- **Year 3 Questions**
- When you ask "What is your name?" Does your child say her full name?
- Can your child tell the name of two or more family members or playmates?
- Can your child tell the correct name of the village/tola/block she stays in?
- Can your child tell which day of the week it is today?
- Can your child count to 5?
- Can your child count to 10?
- Can your child count to 20?
- Can your child name the primary colours (red, yellow, blue)?
- Does your child walk either up or down at least two steps of stairs by herself without holding onto the railing or wall?
- Without holding anything for support, does your child kick a ball by swinging her leg forward?
- Does your child catch a large ball with both hands?
- Does your child serve herself, taking food from one container to another using utensils? For example, does your child use a serving spoon to take rice?
- Does your child unbutton one or more buttons?
- Does your child use a pencil, crayon, or pen for writing or drawing and hold it properly like an adult between thumb and finger?
- Can your child draw a basic figure?
- Does your child brush her teeth by putting toothpaste on the toothbrush and brushing all her teeth without help?
- Can your child do paper folding?