Working paper

IGC
International Growth Centre

Inputs, monitoring, and crowd-out in India's school-based health interventions

James Berry
Saurabh Mehta
Priya Mukherjee
Hannah Ruebeck
Gauri Kartini Shastry

August 2018

When citing this paper, please use the title and the following reference number:
S-89206-INC-2

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

INPUTS, MONITORING, AND CROWD-OUT IN SCHOOL-BASED HEALTH INTERVENTIONS: EVIDENCE FROM INDIA'S MIDDAY MEALS PROGRAM

James Berry, Saurabh Mehta, Priya Mukherjee, Hannah Ruebeck, Gauri Kartini Shastry[1]

August 2018

ABSTRACT

Governments often rely on school infrastructure to implement programs targeting children, but whether and to what extent managerial capacity constraints affect the implementation of these programs are important, open questions. We consider these questions in the context of India's school meals program and iron and folic acid (IFA) supplementation program. Using a randomized controlled trial in a rural district in the state of Odisha, we evaluate the impact of two interventions on child health and on how these government-run programs are implemented. First, we distribute a micronutrient mix (MNM) to be added to the school meal to complement the existing IFA program. Second, we monitor school meals with increased frequency early in the intervention. While we find significant positive effects of distributing the MNM on micronutrient levels in the meals, we find no detectable effects on child health. Increased monitoring of school meals, on the other hand, does improve hemoglobin levels. Monitoring did not affect take-up of the MNM, but it did improve implementation of the government's IFA program. We also find significant negative spillovers of the MNM intervention on how well the IFA program was implemented, suggesting that effort by school officials was crowded out by the introduction of the new MNM program. We present additional evidence suggesting that these effects are driven by managerial capacity constraints.

JEL Codes: O15, I12, H40

Keywords: public service delivery, micronutrient fortification, nutrition programs, monitoring

---

[1] We thank The International Initiative for Impact Evaluation (3ie), the Douglas B. Marshall, Jr. Family Foundation, and the International Growth Center for funding this research. Saurabh Bhajibakare, Averi Chakravarty, Kriti Malhotra, Dilip Rabha, and Naveen Sunder provided excellent research assistance. Urmi Bhattacharya, Anna Custers, and Bastien Michel provided excellent research support and management from J-PAL South Asia.*James Berry: University of Delaware jimberry@udel.edu; Saurabh Mehta: Cornell University, smehta@cornell.edu; Priya Mukherjee: College of William & Mary, pmukherjee@wm.edu; Hannah Ruebeck: MIT; Corresponding author: Gauri Kartini Shastry, Wellesley College, 106 Central Street, Wellesley, MA 02481, gshastry@wellesley.edu.

I. Introduction

Governments around the world rely on school infrastructure to implement many programs aimed at improving child welfare, some of which are only indirectly related to education. In India, for example, the central government has mandated the provision of nutritious midday meals to every primary and upper primary school student since 1995.[2] Subsequently, the government (occasionally in partnership with non-governmental organizations) has considered additional ways of delivering micronutrients to children through schools.

Delivering nutrients through schools is attractive to policy makers for several reasons. First, young children are an important demographic to target. In India, 43 percent of children under the age of five are underweight (International Institute for Population Sciences 2007) and 70 percent are anemic. These deficiencies have substantial consequences for productivity, both at the individual level (see, e.g., Thomas et al. 2006), and for aggregate economic growth (see, e.g., Shastry and Weil 2003). Recent work suggests that early childhood health interventions may have long-lasting effects (Hoddinott et al. 2008; Hoynes, Schanzenbach and Almond 2016). Second, the school system is often the most comprehensive infrastructure available to reach children in remote areas. Third, health interventions can influence educational outcomes: they may act as an incentive for children to attend school (Vermeersch and Kremer 2005) or may remove health constraints to attending school (Miguel and Kremer 2004).

There are a number of limitations to school-based distribution, however. Schools may not have sufficient managerial capacity to implement additional programs, and these new programs may crowd out existing school activities (Vermeersch and Kremer 2005). Leakages or corruption at the school level may also hamper implementation. Even if programs are implemented well, sicker and younger children may not attend school or pre-school frequently enough to benefit.

We examine these caveats to school-based distribution in the context of India's school meals program and a recently implemented iron and folic acid (IFA) supplementation program, which provided students with weekly iron tablets and biannual deworming tablets. We employ a randomized controlled trial in a rural district in the Indian state of Odisha to evaluate the impact of two additional interventions on child health, the primary outcome of interest, and on how well

[2] In India, primary school consists of grades 1 through 5, and upper primary school consists of grades 6 through 8.

school personnel implemented the government-run programs. We first investigate whether providing schools with a micronutrient mix (MNM) to be added to school meals improves child health. Next, we study the impact of monitoring school meals on program implementation and child health, by varying the number of unannounced monitoring visits (i.e., monitoring intensity) during mealtimes. The random assignment of the MNM and the high intensity monitoring treatments allows us to identify the causal impact of each treatment as well as any interaction effects. Finally, we investigate potential spillovers—particularly with respect to implementation—between the government's IFA program and the randomly assigned MNM program.[3]

We find no effect of the MNM treatment on hemoglobin levels or on anthropometric measures of child health – estimated coefficients are small, statistically insignificant and inconsistent in sign. Increased monitoring of school meals, on the other hand, did improve hemoglobin levels, when implemented with or without the MNM treatment. Take-up of the MNM program does not explain these results – we find consistent evidence of take-up from school self-reports, enumerator observation and laboratory tests of meal samples collected at the unannounced visits, and take-up of the MNM does not depend on the number of monitoring visits. Instead, we find that increased monitoring improved implementation of the government's IFA program – students in the high intensity treatment arm were *more* likely to report receiving IFA tablets in school regularly – and we find *negative* spillovers of the MNM intervention on how well the IFA program was implemented – students in the MNM treatment arm were *less* likely to report receiving IFA tablets in school regularly. Using detailed survey data on the implementation of the midday meals program, we show that these effects are driven by schools with lower managerial capacity.

Turning to heterogeneity with respect to initial anemia status, we find that the improvement in hemoglobin levels resulting from the monitoring intervention is driven by children who are mildly anemic, not moderately anemic. This result is surprising, since the nutritional literature on iron supplementation consistently finds that those who are most anemic are most likely to respond to treatment (see, e.g., Abrams et al. 2008, Tee et al. 1999). The coarse data we have on child attendance suggests that this result may be partly explained by more consistent attendance for less anemic children, relative to more anemic children. This result also indicates that while distributing

---

[3] The IFA program was first implemented prior to our study's interventions. In a companion paper, we use quasi-random variation in the government's implementation of the IFA program in the year *before* our intervention, to evaluate the effect of the iron fortification program itself (Berry et al. 2018).

nutrition through the school infrastructure can help keep relatively healthy children healthy, it may not be an effective way to reach the sickest children.

This paper relates to a number of literatures. Within our policy context, a few recent papers have studied the quality of service delivery in India's midday meals program. Closely related is a recent study using administrative data (Debnath and Sekhri 2016), which suggests that a mobile-based monitoring technology might reduce malfeasance and shirking in the school meals program in the Indian state of Bihar. Several other papers examine the health effects of the midday meals scheme and modifications to the program. Chakraborty and Jayaraman (2016) investigate the effects of the Indian midday meals program on learning outcomes by using its staggered implementation across states and time, finding that exposure to the program increased children's math and reading test scores. Krämer, Kumar and Vollmer (2018) evaluate the provision of double-fortified salt — salt fortified with iron and iodine — to schools to be used in midday meals in Bihar. The authors find that the provision of fortified salt increased hemoglobin levels but did not significantly impact test scores.

Our paper also relates to a broad literature on the delivery of publicly provided goods and services in developing countries. A number of studies evaluate multiple strategies to improve service delivery in a variety of sectors (see Finan, Olken, and Pande 2015, for a review). Our paper relates specifically to prior work on monitoring service providers: we find that monitoring, even without explicit stakes,[4] leads to significant improvements in program implementation.

A fairly unexplored area in the literature on service delivery is the potential for crowd-out of effort and negative spillovers on the quality of public goods. Our study is, to our knowledge, one of the first to measure the impact on existing programs when new programs are added within existing public infrastructure. We document that the introduction of the new MNM program crowded out the implementation of the pre-existing IFA program. A novel contribution of our paper is to illustrate that it may be vital to evaluate interventions in the broader policy context, rather than in isolation and independently from other public programs, keeping in mind the capacity constraints faced by existing personnel and infrastructure.

[4] Most — but not all — papers that evaluate the effects of monitoring find that it does improve the quality of public service delivery (Finan, Olken, and Pande 2015).

Our paper is also related to the literature on iron- and micronutrient-supplementation.[5] Although the MNM did not contain iron, it was intended to help children absorb iron and thereby lower iron-deficiency anemia, informed by literature documenting that multi-micronutrient supplementation is more effective at addressing anemia than iron and folic acid supplementation alone (Ramakrishnan et al. 2004, Best et al. 2011, Ahmed et al. 2010). Fawzi et al. (2007) and Mehta et al. (2011) also find that multi-micronutrient supplementation even without iron and folic acid can improve hemoglobin levels. In our case, we found no evidence of an effect of the micronutrient mix even with non-trivial take-up, possibly due to crowd-out of the IFA distribution, but also possibly due to the low dosage levels insisted upon by the National Institute of Nutrition.[6] While this suggests the study may be underpowered to detect an improvement in child health, it also points to limitations in the ability of large-scale distributions advocated by policy makers to provide sufficient nutrition, particularly for the sickest children. The dependence of the effect of school-based nutrition programs on consistent child attendance, as discussed above, exacerbates this implication. This conclusion relates closely to that of Banerjee, Barnhardt and Duflo (2018), who study the viability of double-fortified salt as a means to improve anemia levels. Even when provided for free, they find minimal effects on hemoglobin and attribute it to the low levels of iron that can be safely added to food intended for large-scale distribution.[7]

We proceed as follows: Section II provides context, first describing the school-based nutrition programs implemented by the government and then describing our interventions. Section III describes the experimental design, including the timeline, sample selection, and data collection. Section IV describes the results. Section V discusses external validity and policy implications. Section VI concludes.

---

[5] A major focus area of the nutrition literature is iron supplementation, with studies generally finding that supplementation decreases iron deficiency, across varying locations, populations, baseline rates of deficiency, and intervention methods (Hyder et al. 2007; Hirve et al. 2007; Ahmed et al. 2010; Tee et al. 1999; Gera et al. 2007; Abrams et al. 2008).

[6] We had initially intended to provide 100% of the recommended daily allowance (RDA) for this age group, but approval from the National Institute of Nutrition was conditional on reducing the RDA to approximately 50% and including calcium (which can inhibit the absorption of iron) in the mix.

[7] One noteworthy feature of the nutrition literature on iron supplementation is the amount of researcher control over consumption of supplements and adherence to the program. Banerjee, Duflo and Glennerster (2011) and Banerjee, Barnhardt and Duflo (2018) demonstrate take-up of iron-fortified food at the household level to be a challenge, limiting the potential effect of iron fortification in the field. While school-based distribution may have the potential to bypass household level take-up decisions, it introduces similar issues of inefficiency and corruption at the school level.

II. Background and description of interventions

A. India's school-based health interventions

The midday meal program

The predecessor to the current midday meal program was announced in 1995 when the Central Government of India mandated that primary school children in public and public-assisted schools receive lunch in school. Subsequent revisions to the scheme mandated the lunch be a cooked meal (instead of rations to take home), extended the program to upper primary students and further specified how it was to be implemented (e.g., number of calories).

Our study took place in the eastern state of Odisha, in the rural district of Kendujhar. During the period of our study, the midday meal program in Odisha was supervised by the state Department of School and Mass Education, in coordination with district-, block- (administrative unit smaller than a district) and cluster- (administrative unit smaller than block) level officials. The state government encouraged districts to delegate the preparation of meals to self-help groups (SHGs), groups of local women (and sometimes men) who organize as a financial intermediary for such purposes as running the school's midday meal or improving access to financial services for villagers. To this end, joint bank accounts were created between the president of the SHG and the school headmaster to facilitate the disbursement of funds. However, according to our survey data, SHGs were operating in only 43 percent of schools in our baseline sample. Finally, while the district was also supposed to train those responsible for providing the meals, in only 33 percent of our schools had anyone ever attended a training related to the midday meal program.

Rice was provided through the Food Corporation of India and passed through various administrative units before reaching schools. Almost all of the schools in our baseline sample had received rice (99.7 percent) but only 42 percent of schools received it on a regular schedule. The menu for the school meals was regulated by the state government, in order to ensure variety and prescribed levels of caloric and protein content (see Appendix Table 1). Other ingredients such as dal (pulses), eggs, vegetables, and cooking fuel were typically purchased at the school level, using funds disbursed from the district government. Since SHGs were not operating in most schools in our sample, either the headmaster or one or more of the teachers was responsible for purchasing food materials, obtaining cooking fuel and hiring and supervising cooks. More members of the teaching staff typically helped during lunch to organize the seating of students, distribution of meals, and washing of utensils before classes resume.

Iron and folic acid supplementation program

In 2012, India's Ministry of Health and Family Welfare introduced a national iron supplementation program to reduce the prevalence and severity of anemia among school children. Beginning in January 2013, according to the guidelines distributed by the central government, iron and folic acid supplements, as well as deworming medication, were to be distributed free of charge to all students attending public and public-assisted schools. During the year of our intervention, children aged 5-10 were to receive 45 mg of elemental iron and 400 µg of folic acid once a week, under supervision at school.[8] One tablet of deworming medication, Albendazole, was also to be administered to each child every six months. Headmasters received the medications and were expected to supervise the provision at school.

We surveyed schools and students about the implementation of the IFA program in the spring of 2014 (after the first year of the IFA program, but before our intervention year) and four times during the intervention in the 2014-2015 school year. In the first year, approximately 86% of schools received iron and folic acid tablets, ranging from 49% of schools in one block to 99% in another. In a companion paper (Berry et al. 2018) we use variation in implementation from the first year to estimate the impact of the IFA program in its first year after verifying that this variation appears to be quasi-random. By the start of the MNM intervention, virtually all schools in our sample had received the iron and folic acid tablets. Note that this also affects the interpretation of the MNM intervention: the comparison is between children whose schools received both iron supplementation and multi-micronutrient fortification and children whose schools only received iron supplementation.

[8] In the first year of the IFA program, the year *prior* to our intervention, the program was implemented differently. Children in grades 1-5 were to be given 30 mg of elemental iron and 250 µg of folic acid daily for a duration of 100 days. In both years, upper primary school children (those in grades 6 to 8) were to be given a higher dosage (100 mg of elemental iron and 500 µg of folic acid) each week.

B. Experimental interventions

The micronutrient mix program

The micronutrient mix (MNM) program was designed and implemented by the research team in consultation with the government of Odisha and the National Institute of Nutrition. We provided school headmasters and cooks with a multi-micronutrient mix, containing Vitamins A, C, D, B1, B2, B6, B12, Niacin, Zinc, Selenium and Calcium, to be added daily to the midday meal.

Note that the MNM we provided did not include iron or folic acid; we did not want to risk providing the children with too much iron, given the government IFA program implemented at the same time. Rather, the MNM was intended to help children absorb iron and complement the IFA distribution. This was motivated by the nutrition literature demonstrating that multi-micronutrient supplementation is more effective in combating anemia than iron and folic acid supplementation alone (Ramakrishnan et al. 2004, Best et al. 2011). It is also important to note that Fawzi et al. (2007) and Mehta et al. (2011) find that multi-micronutrient supplementation even without iron and folic acid can improve hemoglobin levels. As noted in the introduction, the dosage of these vitamins and minerals was restricted to approximately 50% of RDA by the National Institute of Nutrition to avoid over-supplementation.

In order to implement the intervention, we first trained headmasters, cooks, and other staff involved with meal preparation, such as self-help group members. During these trainings, we covered the health consequences of anemia and other forms of malnutrition, the health benefits of consuming the various vitamins and minerals in the MNM (also noting that they can aid the absorption of iron), and the directions for MNM use. We distributed the mix, plastic seal-able jars, and scoops that held 10 grams of the mix. The dosage approved by the National Institute of Nutrition meant that children were to receive 1.5 g of the MNM each day. Since the mix was to be added to the food before it was served to the children, it was necessary to estimate how many children would be eating the meal and multiply this by 1.5 to calculate the number of grams of the mix to add and then divide by 10 to calculate the number of scoops to add. During the training, we practiced calculating this number and found it necessary to involve the headmasters in the addition of the mix since cooks were usually not confident about performing this calculation. We also gave schools laminated fliers that clearly described the steps necessary to add the MNM to the food (see

Appendix Figure 1). Every month, we contacted schools to enquire whether they needed more of the MNM and, if so, delivered additional packets to the school.

High intensity monitoring

The second intervention involved earlier and more frequent monitoring of school meals during the study period. All schools in the study were visited during meal time on a random day once per month during the last three months of the study, but enumerators also visited the schools in the high intensity monitoring treatment group during the first two months of the intervention. As described in more detail below, monitoring involved observations of meal quality, child attendance, the distribution of food to the children, and how much of the food was consumed by the children. Enumerators also asked the headmasters and cooks about the preparation of the meal and storage of cooking equipment and ingredients, collected a sample of the meal for laboratory testing and measured the height of three randomly chosen students.

III. Experimental design

In this section, we describe the timeline, study sample, and experimental design for the randomized controlled trial evaluating the MNM provision and high intensity monitoring treatments.

A. Timeline

Figure 1 gives the chronology of key activities for the study. The original design was to fortify the school meals with iron. Three hundred seventy-five schools were selected for the study, and an initial baseline survey (Baseline 1) was conducted in these schools between September 2012 and January 2013. However, the plan was halted when the government's IFA program was announced, and the study was revised to evaluate the MNM program. Changing the intervention plan required securing approvals for the new design from a number of government agencies and took approximately 16 months, with final approvals received at the end of September 2014. While waiting for final approval, we conducted a survey measuring the intensity of IFA implementation, as well as a second baseline survey (Baseline 2) in a subset of sample schools during August and September of 2014, early in the 2014-2015 school year. Baseline 2 focused on the three administrative blocks (157 schools) with variation in IFA implementation in the first year in order to evaluate the impact of the government's IFA program on child health (see Berry et al. 2018).

The MNM and high intensity monitoring programs were launched at the end of November 2014 and continued through April 2015, in 150 schools across the 5 blocks represented in our original sample. During this period we also monitored school meals and conducted surveys to collect information on student attendance, MNM usage, and IFA tablet usage. Food samples were collected twice from each of the sample schools. The endline survey was conducted between April and July 2015.

B. School sample

The sample schools in Kendujhar district were selected for the study based on whether they satisfied the following conditions: (i) the school was located within 50 kms from the town of Kendujhar, the capital of the district and (ii) the school was located in one of five blocks: Banspal, Ghatagaon, Jhumpura, Sadar, or Patna. This minimized the fixed costs of dealing with government officials in charge of schools in each block. We initially had a larger sample of 377 primary schools that satisfied these conditions. One hundred and fifty schools were randomly selected out of the 377 in which to conduct the MNM and high intensity monitoring evaluation.[9] These schools are primarily rural with a high fraction of students from tribal/scheduled caste communities (approximately 95%). Households are relatively poor – 50% have electricity, 30% own a phone, and 50% of household heads are literate – and children are relatively unhealthy – 44% are underweight and 60% are anemic. In terms of child health, the sample is fairly representative of the state of Odisha, in which 41 percent of children under the age of five are underweight and 65 percent are anemic (International Institute for Population Sciences 2007).

C. Treatment assignment

Out of the 150 schools selected for the evaluation of the MNM and high intensity monitoring, 75 were randomly assigned to receive the MNM, stratified by block and school type (i.e., whether the school only had primary grades 1-5, or also had upper primary grades 6-8). Within each group of 75, half of the schools were randomly assigned to high intensity monitoring.

[9] The original sample of 377 schools was chosen to evaluate provision of fortification to schools as well as centralized school meal delivery operated by the NGO Naandi Foundation. Due to the various delays the project faced, the Naandi Foundation ultimately decided not to participate in the study or provide meals to the study schools.

Table 1 provides the number of schools and students in each group. While the original sample contained 150 schools, 2 schools refused to participate from the beginning of the study, before their treatment status was revealed. Thus, we were left with 75 MNM treatment schools and 73 comparison schools. Out of the 75 schools in the MNM treatment group, 37 were monitored intensely, while 38 were not, and out of the 73 schools that did not receive the MNM, 36 were monitored intensely while 37 were not.

Table 2 shows that the schools in each group are well balanced on a range of covariates measured during our first baseline school survey, right after randomization (top panel of Table 2), or measured during the first month of the intervention (bottom panel of Table 2). Each row shows the mean for that variable for the following groups: (i) schools that received neither the MNM treatment nor the high intensity monitoring, (ii) schools that only received the MNM treatment, (iii) schools that only received the high intensity monitoring, and (iv) schools that received both MNM as well as high intensity monitoring. The final column provides the p-value of the F-test of equality across all four groups. No individual covariate is significantly different across all groups.

D. Data

We collected data on a number of outcome variables at various points during the study. Specifically, we collected extensive data on (i) child-level outcomes (including hemoglobin levels; anthropometric measures such as height, weight, and mid-upper arm circumference; cognition; school attendance; and test scores), (ii) baseline household-level characteristics and school-level characteristics, (iii) the quality of midday meals and take-up of the micronutrient mix program (including the quantity of vitamin A and zinc in food samples), and (iv) the implementation of the IFA program. We briefly describe each survey and the variables of interest below. Figure 1 indicates the timing of these survey modules.

Child-level data

We randomly chose 15 students in each school for collection of health and education data. These students were chosen from the set of students enrolled in sample schools in grades 1 to 5 who lived with their parents. We excluded children who lived in school hostels due to the difficulty in locating parents to obtain consent. Students were randomly chosen, after stratifying by school and grade. The original baseline survey (Baseline 1) included 3 students per grade in grades 1 to 5 during the 2012-2013 school year. Because of the implementation delays described in Section III.A, children who were in grades 4 and 5 during Baseline 1 had finished primary school by the beginning of the intervention year and were ultimately excluded from the sample. During the 2014-2015 school year, we sampled an additional 3 students per grade in grades 1 and 2 so that the final sample at endline covered grades 1 to 5 during the intervention year. With attrition, there are on average 14 students per school surveyed at endline.

As described in Section III.A, we conducted a second baseline (Baseline 2) at the beginning of the 2014-2015 school year, in about half of the schools in the original sample. This survey was conducted with 9 children per school. Appendix A describes the sampling procedure for students in Baseline 2.[10]

After obtaining parental consent, enumerators visited schools to measure the selected children's height, weight and hemoglobin levels during the Baseline 1, 2, and Endline surveys. During the Baseline 1 and Endline surveys, students in grades 1 to 4 were given reading and mathematics tests designed by Pratham, an India-wide NGO that works on child literacy and numeracy. Enumerators also conducted tests of cognitive development. We used two cognitive development tests: (i) a Digit Span Test (Pershad and Wig 1988) where children are asked to repeat sequences of numbers, ranging in length, both forwards and backwards and (ii) a Block Tapping Test (Kar et al., 2004, 2008) where children are asked to tap the top of four boxes in the same order in which a surveyor taps the boxes or in reverse order. The total number of points possible on these tests is 26 and 10, respectively. Scores on all four tests (Digit Span, Block Tap, Language and Mathematics) are normalized using the control group distribution by grade and survey round (Baseline 1 or Endline). School attendance data were also collected each month through random, unannounced visits. These checks were made at random times of the day in case children attend school just for the meal and leave immediately after.

We also collected baseline household-level characteristics, including demographic information, household assets, knowledge of anemia, and perceptions of the school's midday meal. At the same

[10] Recall that at Baseline 2, we only surveyed children in the 3 administrative blocks with variation in the IFA implementation. Data from Baseline 2 is primarily used in the companion paper evaluating the IFA program (Berry et al. 2018), but we control for these updated baseline hemoglobin measures in some specifications below.

time as Baseline 1 and again during the first two months of the intervention, we collected data on school characteristics and teacher demographic details and qualifications.

MNM take-up and midday meal monitoring

During the intervention year, we monitored take-up of the MNM program, observed meal quality, and tested for micronutrients in the meals. Take-up of fortification schemes is itself an important outcome. One measure of MNM program take-up is the amount of MNM each school added to midday meals during the school year. We calculate this as the amount of MNM received, relative to the amount we calculated they would need to serve their students, minus the amount of the MNM that remained at the end of the school year in April 2015.

In addition, trained enumerators made surprise visits to the study schools to observe the quantity and quality of school meals during the 2014-2015 school year. Schools in the low intensity monitoring treatment arm received these visits during the third, fourth and fifth months of the intervention. Schools in the high intensity monitoring arm received these visits every month (5 visits total). During the third and fifth months of the intervention, enumerators collected samples of the food being served and sent these samples to a laboratory for nutritional analysis. We have data on the amount of vitamin A and zinc in the food sample.[11]

IFA program implementation data

During March and April 2014 and the first, third, fourth, and fifth months of the intervention, enumerators visited each school to determine whether IFA tablets had been received from the government and how well the IFA program was being implemented. After speaking to the headmaster at each school, our enumerators randomly selected three children to answer additional questions about the IFA program. One student was randomly chosen from each of grades 2, 4, and 5. For each school, we calculate the fraction of those three children that reported receiving tablets regularly and the fraction that report receiving tablets recently.[12]

[11] We chose to test only these micronutrients for budgetary reasons and because pilot tests of samples of fortified food cooked by our research team and sent to this lab provided the most consistent results for vitamin A and for zinc.
[12] Students were also asked if they swallowed the tablets they received. Almost all students responded that they had.

E. Summary statistics

Table 3 checks balance on child health and demographics, across each of the treatment groups. Panels A and B focus on child health before the intervention; Panel A includes children who were in the sample at Baseline 1, while Panel B includes children surveyed at Baseline 2. For children added to the sample during the 2014-15 school year, we fill in demographic information collected at the endline survey if the variable is most likely time-invariant or unrelated to treatment (for example, we fill in the variable for mother's education but not whether the child takes any supplements). The groups are well balanced on most variables, including child health outcomes, household-level characteristics and demographic information on children, mothers, and heads of household, with a slight imbalance on a few of the 35 variables in the table. We cannot rule out that the significant differences in these cases exist merely by chance, but our preferred specifications include school or child fixed effects, effectively controlling for these possible differences.

Given that the sample changes over the two years between Baseline 1 and the intervention, as described above, Table 3 focuses on children who were in the sample at endline. Appendix B presents an analysis of attrition for our main outcome variables on child health. We find no significant differences in attrition between the schools that received MNM, schools that received high intensity monitoring, and the control group. However, adding an interaction term between the MNM and high-intensity treatments does reveal some significant differences. Appendix B presents several additional analyses that suggest that differential attrition does not substantially bias our results. We show that attriters have similar baseline characteristics across treatment groups. We also present Lee (2009) bounds of our results that account for potential differential attrition, and our results remain highly significant.

IV. Results

This section describes the impact of the MNM and high intensity monitoring interventions on various outcomes of interest, including MNM take-up, child health, and program implementation and spillovers.

A. MNM take-up

Our first outcome of interest is take-up of the MNM by schools in the MNM treatment group. Denoting a measure of take-up in school s in block b measured at time t as y_{sbt}, the basic specification in our analysis is as follows:

$$y_{sbt} = \beta_0 + \beta_1 MNM_{sb} + \beta_2 High_{sb} + \alpha_b + \lambda_t + \varepsilon_{sbt} \quad (1)$$

where MNM_{sb} is a dummy variable for schools that received the MNM fortification treatment, and $High_{sb}$ is a dummy variable for schools that received higher frequency monitoring visits. All our regressions contain fixed effects for administrative block, α_b. We include month fixed effects, λ_t, since some measures of take-up were collected multiple times during the intervention. Whenever we make use of multiple observations within a school, standard errors are clustered at the school level. In order to account for any differential impact that high intensity monitoring may have had on the MNM treatment, we include a specification that includes an interaction term

$$y_{sbt} = \beta_0 + \beta_1 MNM_{sb} + \beta_2 High_{sb} + \beta_3 (MNM_{sb} * High_{sb}) + \alpha_b + \lambda_t + \varepsilon_{sbt} \quad (2)$$

For some specifications, we also control for whether the school received IFA tablets during the previous school year to see if experience with nutrition supplements matters for implementation.

We consider three measures of take-up. First, we have data on the number of MNM deliveries made to the school, the amount of MNM delivered in kilograms, and the amount of MNM used in kilograms. Second, we have measures of take-up from the midday meal monitoring visits, such as whether the enumerator noticed a powdery addition to the meal. Finally, we have laboratory reports of the amount of vitamin A and zinc present in meal samples.

The estimates on take-up measured from our delivery records are reported in Table 4. We exclude schools not in the MNM treatment. In addition to block fixed effects, these regressions control for the number of children enrolled in the school at the start of the intervention. Schools assigned to the MNM treatment did take up the mix. The schools that were not monitored intensely received 2.8 deliveries of the mix during the study period (the dependent variable mean, presented at the bottom of Columns 1-2). On average, schools received approximately 0.6 kg of the mix per child enrolled and used almost all of it. This represents more than 58 percent of the amount we estimated they would use based on the number of students enrolled. Ninety percent of the schools used at least 25 percent of the amount we estimated they would need. The high intensity monitoring did not affect these measures of take-up.

Table 5 further reports take-up as inferred during the midday meal observation visits conducted by our enumerators. These measures allow us to compare take-up between the MNM treatment schools and the non-MNM treatment schools as well as across high and low intensity MNM treatment schools. We find that being in the MNM treatment group significantly increases (i) the likelihood of our enumerator being able to detect the MNM directly on inspection of the container in which the meal was cooked (Columns 1-3),[13] (ii) the likelihood that the cook reports that he/she added anything to the cooked food before serving it to the students (Columns 4-6), as well as (iii) the likelihood that our micronutrient mix was present in the room where food materials are stored (Columns 7-9). High intensity monitoring did not affect these measures of take-up.[14] The bottom panels of the table separately analyze meal observations from early in the intervention (when only highly monitored schools are observed, Panel B), the first months when all schools were observed (Panel C) and the last months of the intervention (Panel D). Comparing the magnitudes in Panels C and D suggests that take-up did not decline much over time.[15]

Finally, Table 6 reports take-up as measured by the amount of vitamin A and zinc present in food samples collected at each school and tested at a laboratory. Note that meals can contain vitamin A and zinc even if they do not contain the mix, so these measures could be considered a measure of meal quality, and not simply take-up of the intervention. However, dependent variable means from the control group indicate very little vitamin A and zinc present in the benchmark samples. We find large, significant increases in the amount of vitamin A and zinc for schools in the MNM treatment. The increase in zinc persists through April, the last month of the intervention, while the increase in vitamin A is still significant in April, although it is smaller than it was in February. We suspect this is due to higher stability of zinc than of vitamin A during storage (Kuong et al 2016).

---

[13] The treatment effect implied by how many enumerators detected a powdery addition in the meal is quite small relative to the other measures of take-up. Observing a powdery addition is not an ideal measure of treatment compliance since if the cook stirred the pot sufficiently, the MNM should not be easily detectable.

[14] Monitoring does affect whether cooks report using the mix ($p < .1$; Panel A, Columns 4 and 6), but this difference is economically quite small, does not persist throughout the school year (see Panel C) and is not robust to the other, more objective measures of take-up. The observed effect may therefore be due to reporting bias.

[15] Note that the coefficient on whether or not a school received IFA tablets in the previous year is often negative and sometimes marginally significant (only at 15 percent in Table 4 for the amount of MNM delivered to the school but at 10 percent in Table 5 for whether the MNM could be located in the storeroom). We return to possible spillovers between the two nutrition programs below.

As with the other measures of take-up, high intensity monitoring does not affect the amount of vitamin A or zinc found in the samples – the coefficients are small and of inconsistent sign. Note also that the low levels of vitamin A and zinc in the food samples suggest that spillovers between treatment arms were very unlikely – headmasters in the control schools did not obtain a similar mix to fortify their meals.

As noted above, the dosage agreed upon for our intervention would give children approximately 50% of RDA for the micronutrients listed, including vitamin A and zinc. We conduct back-of-the-envelope calculations based on these measures of take-up to get a better sense of how much children in MNM treatment schools received on average. Our midday meal observations document that children receive (and eat) approximately 130 mL of dal or vegetable curry each day, a little bit more than half a cup. We estimate that this weighs about 110 g. Using the range of estimates in Table 6, this means the MNM treatment increased vitamin A intake by 190-376 μg, roughly 30-60% of RDA for this age group, and zinc intake by 2 mg, roughly 20% of RDA for this age group.

B. Impact on child health

To estimate the health impacts of the interventions, we use a lagged dependent variable model:

$$y_{1isb} = \beta_0 + \beta_1 MNM_{sb} + \beta_2 High_{sb} + \beta_3 (MNM_{sb} * High_{sb}) + y_{0isb} + \alpha_b + age_{isb} + \varepsilon_{isb} \tag{3}$$

where y_{1isb} is the health outcome of child i in school s in block b at endline, and y_{0isb} is a baseline measure of the outcome variable. We include fixed effects for both block, α_b, and age, age_{isb}. As with the take-up analysis in the previous section, we estimate equation (3) with and without the interaction term $MNM_{sb} * High_{sb}$.

Table 7 presents these treatment effect estimates on measures of child health. Panel A focuses on a continuous measure of anemia status, hemoglobin levels (in g/dl), while Panel B focuses on a dummy variable for whether a child is anemic.[16] Panels C and D use, as the outcome variable, BMI-for-age (in z-scores) and height-for-age (in z-scores), respectively.[17] Column 1 includes no

[16] Hemoglobin level cutoffs used to classify children as anemic are those defined by the WHO at sea level by age group (WHO 2011). For the majority of the sample (ages 5-11), children with hemoglobin below 11.5 g/dL are anemic.
[17] In results available upon request, we replicate Table 7 for outcomes measuring school attendance, cognitive ability, and proficiency in reading and mathematics. Neither intervention has statistically significant effects on these

other controls while Column 2 includes the lagged dependent variable from Baseline 1. Columns 3-5 include the lagged dependent variables from both Baseline 1 and 2 and dummies for missing observations to allow for the inclusion of all children surveyed at endline. Recall that some children were included in the sample only at endline because they were too young to be enrolled in school during the Baseline 1 survey, two years prior to the intervention.

The results in Table 7 indicate that the MNM treatment had no effect on child health; in fact, the coefficients are often negative although always small and never significant. By contrast, high intensity monitoring increased hemoglobin levels by 0.17-0.24 g/dL and reduced the probability of being anemic by about 6-9 percentage points — a 10-15% decrease relative to the control mean. We look to other school-based iron supplementation programs to put the magnitude of this result in context. Krämer, Kumar and Vollmer (2018), using double-fortified salt in school midday meals in Bihar, find that hemoglobin increased by about 0.14 g/dL and the probability of being anemic fell by 9.3 percentage points. Luo et al. (2012) provided iron supplements in school to 4th graders in rural China and find that, on average, hemoglobin increased by 0.23 g/dL after one year. For an alternative comparison, a meta-analysis of 55 efficacy trials concludes that consistent iron supplementation increases children's hemoglobin levels by 0.74 g/dL — 1.1 g/dL for children with baseline hemoglobin levels below 11 g/dL and 0.49 g/dL for children with baseline hemoglobin above 11 g/dL (Gera et al. 2007). One might consider these highly-monitored efficacy trials an upper bound on the potential effect of school-based distribution.

Looking at the rest of Table 7, there are no significant impacts of the monitoring treatment on anthropometric outcomes in this time frame. The results also indicate no interactions between the MNM treatment and high-intensity monitoring (Column 4), with no significant impacts on either hemoglobin or anthropometric outcomes. These results are also robust to controlling for measures of IFA receipt during the previous year, specifically whether the school received IFA tablets (Column 5). In Appendix Tables 2 and 3, we verify that using additional anthropometric outcomes (BMI, weight, height, mid-upper-arm circumference and weight-for-age z-scores) or a differences-in-differences specification does not affect these conclusions.

---

outcomes. This is not particularly surprising given the short time horizon, the lack of an effect on child health for the MNM intervention and the fact that no other school characteristics likely changed (such as teacher motivation).

C. Mechanisms

In interpreting the results described above, it is important to consider how the three interventions – MNM provision, high intensity monitoring and the IFA program – may have interacted. Recall that the nutritional basis of the MNM intervention was that the vitamins and minerals in the mix would complement the iron from the IFA program (making the iron easier to absorb). It is also easy to imagine complementary effects if high intensity monitoring gave headmasters additional incentives to implement the MNM distribution or the IFA program more consistently. At the same time, the MNM treatment was a new program, implemented on top of the existing midday meals program as well as the IFA program. Since the addition of one more program at the school level increased the workload of the school staff, it is plausible that this might lead to negative effects on how well the other programs were implemented.

We first consider the positive effects of the high intensity monitoring in this light. There are a few explanations. First, it could be that these schools implemented the MNM fortification better. However, our results on take-up discussed above suggest this is not the case: high intensity schools were not more likely to take up the intervention or have more nutritious meals, and the effect of the monitoring on hemoglobin does not depend on whether the school had the MNM treatment.

A second explanation is that the high intensity monitoring may have led schools to implement the IFA program better. Table 8 displays treatment effects on measures of how well the IFA program was implemented. We focus on three measures of IFA implementation quality: (i) whether the headmaster is able to show the enumerator an IFA tablet (Columns 1-3), (ii) the number of tablets distributed per child in the past week (as reported by the headmaster, Columns 4-6), and (iii) the percent of students who say they get the tablets weekly or more frequently (out of three randomly chosen students spanning different grades, Columns 7-9). The results indicate that neither of our treatments affects whether the headmaster has a tablet available to show the enumerator or whether the headmaster reports distributing tablets the past week. However, students in the more intensely monitored schools are more likely to report getting IFA tablets regularly. Children were randomly chosen each month, making this outcome more difficult for the headmaster to manipulate. In addition, these results are driven by responses later in the school year. The difference is insignificant at the first IFA visit during the intervention (usually in December 2014), since many schools in the high intensity treatment arm had yet to receive a meal monitoring visit. By February

2015, however, the effects start to appear – most high intensity schools had received at least 2 and sometimes 3 midday meal visits while low intensity monitoring schools had received at most 1 visit. In fact, in the later part of the year, even headmasters are more likely to report distributing tablets in the highly monitored schools ($p < 0.1$). Note that in Panel B, during the first IFA implementation survey during the intervention year, students in schools that had received IFA tablets in the *prior* year are more likely to report receiving regular medications in school. This difference goes away over time, as all the schools receive IFA tablets during the intervention year. This finding provides a useful check on the reliability of student reports of tablet receipt.

At the same time, Table 8 reveals that schools receiving the MNM treatment do worse on IFA implementation. Students are less likely to report having received the IFA tablets regularly, again driven by the later visits. These results suggest that there is some crowding out of IFA implementation by the introduction of the MNM. As noted in footnote 15 above, there is suggestive evidence of crowd-out in the other direction in Tables 4 and 5: receiving IFA tablets during the previous year may reduce take-up of the MNM, although the coefficients tend to be significant only at 10 percent or 15 percent. Given these two results, one may wonder if some headmasters were intentionally choosing not to implement one of the nutrition interventions, perhaps out of concern that the children were getting too many supplements. The fact that the crowd-out result in Table 8 is driven by student reports of receiving tablets, and not headmaster reports of distributing them, does not support this explanation, but we are not able to rule it out entirely. A related possibility is that headmasters choose to sell the iron tablets since the students are now receiving other micronutrients. The fact that these tablets have very little market value in the region, and that we see no significant difference in the ability of headmasters to produce a tablet to show the enumerator, provides some evidence against this hypothesis.

Instead, we believe the crowd-out is driven by limited resources on the part of headmasters, either in terms of manpower or in terms of managerial capacity. Table 9 estimates regressions similar to those in Table 8 with student reports of tablet receipt as the dependent variable, but for subsets of schools to capture heterogeneous effects according to a school's resource constraints (measured at the beginning of the intervention). We create an index to proxy for a school's managerial capacity with regard to the implementation of the midday meals,[18] and find that both the monitoring impact and the crowd-out are driven by schools with below-median scores on the index. The crowd-out effect is statistically significantly different between above- and below-median managers. We also provide results for several components of the index, and find similar results if *low* managerial capacity is proxied by having fewer than 4 teachers[19] assigned to help run the midday meal, not having an external self-help group manage the meal, or by a school not having a treated water supply (treating a schools' water likely requires effort on the part of the headmaster and could proxy for managerial capacity). Note that most coefficients are not statistically different from each other across low and high managerial capacity schools, although the effects are driven by student reports later in the school year (Panel C), where the difference in the crowd-out estimates is starker.

D. Heterogeneous effects with respect to baseline hemoglobin status

In this section we explore heterogeneity in the impact of high intensity monitoring by student's baseline hemoglobin level. Figure 2 plots a local linear regression of the change in hemoglobin from Baseline 1 to Endline with respect to initial hemoglobin level separately for the control schools and the schools that received high intensity monitoring, along with 95% cluster-bootstrapped confidence intervals. We find that the positive effect of monitoring on hemoglobin levels is driven by children around the threshold of anemic (around 11.5-12 g/dL), rather than by the children with the lowest levels of hemoglobin.[20]

This is a surprising finding given the consistent result in the nutrition literature documenting larger effects of micronutrient supplementation for those who are initially more severely anemic.[21] Recall

[18] The components of the index are: having more than four teachers administer the midday meal (which is the median number of teachers), having an external self-help group administer the midday meal, having the school's water treated, having anyone from the school or self-help group attend a midday meal training, having a record of the most recent school management committee meeting, and reporting sufficient funds to administer the midday meal program. For each of these measures, a positive response indicates high managerial capacity and a negative response indicates low managerial capacity. Failing to answer any of these questions is recorded as low manager capacity. The index is calculated by standardizing each variable with respect to the control group distribution and summing the standardized variables. We obtain qualitatively similar results if we do not standardize the variables and simply sum the binary indicators, though the p-value on the difference between the two groups increases.

[19] Note that the specifications include school size (number of students) as a control variable.

[20] In Appendix Figure 2, we replicate Figure 2 for all four groups of schools. The confidence intervals overlap (and are not shown in the interest of readability), but the figure is not inconsistent with the conclusion that the MNM treatment crowded out some IFA implementation since the curves for the schools that received the MNM treatment lie between the control schools and the schools that only received the high intensity treatment.

[21] Similarly, it may be surprising that monitoring improved hemoglobin levels of children who were initially not anemic. Recall that baseline hemoglobin status was measured 2 years prior to the beginning of the intervention. We

that most of these studies invest substantial resources ensuring compliance; while we find consistent take-up of the program by schools, children who do not attend school regularly will miss fortified meals and are likely to miss receiving the iron tablet (as well as deworming medication, distributed twice a year). We do not have precise measures of daily attendance, but the coarse measures we have suggest that the more anemic children are more likely to miss school, possibly explaining this result. We use two measures of attendance to examine this. Our preferred measure is from Baseline 1, before the intervention. Since our enumerators measured hemoglobin levels at school, we can gauge consistent attendance by the number of visits our enumerators had to make before they found a particular child. Figure 3 shows that the number of visits required was monotonically falling in hemoglobin level. During the intervention, we also measured child attendance at two unannounced visits to school, where we recorded attendance for all the children on the roster, not just the children in the study sample. We can find approximately 66% of our endline sample in the roster attendance using a fuzzy matching algorithm to compare the roster and our sample by grade and name. Figure 4 shows that attendance is increasing in baseline hemoglobin level, even when baseline hemoglobin was measured two years prior to this attendance measure. While consistent with our hypothesis, it is important to note that the gradient is not particularly steep. Enumerators needed to visit a school less than 0.5 more times, on average, to find the most anemic children, relative to the least anemic children (Figure 3) and average attendance rates range from approximately 70% to 75% (Figure 4). Thus, while attendance may be part of the story, it is unlikely to entirely account for the fact that the increase in hemoglobin is driven by mildly anemic children.

V. Discussion and policy implications

The goal of this research was to study nutrient fortification and supplementation "in the field." While efficacy trials have convincingly demonstrated that fortification and supplementation can improve child health and school attendance, these studies are often highly monitored with compliance rates above 90 percent because researchers closely supervise the delivery and consumption of nutrients. This study, on the other hand, focused on programs that distributed

---

have too few observations overlapping between Baseline 2 and Endline (only 368 across all treatment arms) to generate a similar plot using an updated measure of initial hemoglobin, but we are able to investigate changes in hemoglobin from Baseline 1 to Baseline 2. We find that more than 40% of non-anemic children at Baseline 1 were anemic by Baseline 2. On average children who were not anemic at Baseline 1 experienced declines in hemoglobin large enough to make them marginally anemic by the time the intervention began (see Appendix Figure 3).

nutrients through existing infrastructure, specifically the Indian midday meal program, with an emphasis on program implementation. We start with some general considerations with respect to external validity and then discuss two policy-relevant puzzles raised by our results: (i) why the MNM did not affect hemoglobin, despite consistent take-up, and (ii) why the extra monitoring visits did improve hemoglobin, despite not affecting take-up of the MNM.

A number of elements of the study and the setting need to be considered when extrapolating from these results to other settings. First, the MNM and monitoring interventions were designed and implemented by the research team's field staff.[22] Thus, while our results regarding the responses by schools may be applicable to schools in other settings, it is important to note that the impacts of these interventions are conditional on consistent delivery of the MNM and, as the results suggest, actual visits by monitors. The analysis of IFA implementation in the year before our experiment provides some insight in this area: variation in the receipt of IFA tablets across blocks during this first year is likely due to incorrect estimates of the number of tablets each block needed. In the second year, this appears to have been resolved. Almost all schools received the IFA tablets within a few months of the start of the school year. This suggests that taking MNM provision to scale would be possible. At the same time, the impact of high intensity monitoring requires that enumerators actually visit the schools and that these visits are unannounced. Government audits are famously infrequent in India (Muralidharan et al. 2017). Taking intensity of meal monitoring to scale would require addressing the issues that currently limit effective monitoring.

Another element that might affect the generalizability of this study is the fact that school meals in Odisha are relatively consistent. Out of 732 unannounced visits to schools, only 12 times (1.6 percent) was a meal not served. This makes positive effects of fortification more likely, but may make it less easy to generalize the effects to a setting where school meals might be inconsistent.

A. Policy implications from the MNM distribution

The evaluation of the MNM distribution has several policy implications. First, note that while not perfect, take-up was relatively high: according to our records, only 3 schools out of 75 did not use

[22] MNM provision through midday meals is something that the government has been experimenting with in partnership with non-governmental organizations (NGOs) such as Naandi and Akshayapatra. These NGOs typically have a centralized kitchen where the MNM is mixed into the food, after which the food is distributed by the NGO staff to schools within a certain radius of the centralized kitchen.

any of the micronutrient mix. As described in Section IV.A, schools used about 58 percent of the MNM we estimated they would need. The range of take-up measures is similar across both the IFA program run by the government and the MNM program run by the researchers. For example, in 72 percent of midday meal visits, the cook reported adding a powder to the meal, while even in the first year of the IFA program's implementation when only 86 percent of schools received the tablets, 62 percent of children interviewed reported receiving the IFA tablet regularly. These take-up measures bode well for the potential of school-based health programs to improve child health.

That said, the MNM distribution did not actually improve measures of child health, despite previous literature that indicated multi-micronutrient supplementation is more effective than iron supplementation alone. There are many possible explanations. First, while we intended the MNM to (biologically) complement the IFA distribution, it actually crowded out implementation of the IFA program. We find evidence supporting this hypothesis, and while this may be somewhat specific to this context, it raises a policy-relevant concern about running multiple complementary programs through schools with limited resources (either in terms of labor input or managerial capacity). Anecdotal evidence suggests that the headmasters and teachers feel overburdened.[23]

Another possible explanation for the lack of an effect on child health is that the MNM treatment had less than perfect compliance, about 58% by some measures. One of the goals of this study was to look at how effective school-based distribution would be with minimal control by researchers. The lack of an effect therefore seems like the correct policy-relevant interpretation. With this in mind, it is useful to consider take-up of the government-run IFA intervention as a more generalizable measure of compliance in school-based nutrition programs. Even in the second year of the IFA program, when virtually all headmasters reported receiving the tablets from the government, approximately 80% of them reported distributing the tablets the previous week and only 62% of students reported receiving tablets.

Finally, the lack of an effect of MNM provision on child health outcomes could be due to the low dosage of micronutrients provided. As noted above, in order to obtain approval from the National Institute of Nutrition, we were required to halve the originally chosen dosage. The resulting dosage was about half of the recommended daily allowance (RDA) for children of this age, under the

[23] One of the most common concerns about the midday meal reported by the school officials during our field visits, was that it is takes up the headmasters' as well as teachers' time and mental energy.

assumption that these children would obtain additional micronutrients from other sources. This seems unlikely given the very low concentration of tested micronutrients in the meals provided in control schools (approximately 52-55 µg/100g vitamin A and 5-8 mg/kg zinc, about 10% of RDA each). Thus, the low quantities may not have been sufficient to impact iron absorption.

The fact that we had to halve the dosage indicates one disadvantage of general fortification or supplementation programs such as the IFA or the MNM distribution, because they require a one-size-fits-all-students approach. For safety reasons, micronutrient doses must be limited, but that also means the potential for impact is limited, especially for the sickest children (see Banerjee, Barnhardt & Duflo 2018 for a similar conclusion from a household-based program). A more customized program would allow for supplementation or fortification based on the micronutrient deficiencies a child exhibits, but may be prohibitively expensive to implement. Programs that target the general population are most likely to improve wellbeing for mildly malnourished children and perhaps reduce the probability that children develop mild forms of malnutrition, but may not be sufficient for children with more severe deficiencies. In sum, our results indicate that basic school-based programs may be effective at keeping relatively healthy children healthy, but substantially more resources would be required to treat less healthy children, such as initial testing, better tracking of children outside of school, and higher micronutrient doses.[24]

B. Policy Implications from Increased Monitoring

Finally, the robust positive impact of high intensity monitoring on child hemoglobin levels is particularly interesting and relevant for policy, and warrants some discussion. We find no evidence that the high intensity monitoring increased take-up of the MNM – high intensity schools did not request or use more of the MNM. Instead, we find evidence that high intensity monitoring improved implementation of the government's IFA program – students in high intensity schools were more likely to report receiving IFA tablets regularly. While our intent was to gather information on the quality of midday meals and take-up of the MNM, this intent was not conveyed to the schools. Since schools almost uniformly reported that they distributed the IFA tablets during

[24] Berry et al. (2018) explores another implementation-related concern regarding school-based nutrient distribution. We document variation in implementation in the IFA program's first year and use this variation to study the impact of the government's program itself. We find no evidence of an effect in the whole sample, which we attribute to a delay between when iron supplements were consistently distributed and when hemoglobin was measured. While fixable with consistent implementation, this suggests that scheduled breaks in the school year and irregular child attendance could still limit the effectiveness of school-based distribution.

meals, it is natural that they would have thought that one of the reasons for the unannounced visits at mealtime was to verify the distribution of IFA tablets.

We see two possible explanations for this response: First, the visits could have acted as reminders. However, we find that the effect of monitoring is driven by accumulated visits and not by proximity to the most recent visit (results available upon request), suggesting the visits did not simply act as reminders – but the experiment was not designed to have enough power to differentiate these effects. Second, visits may have acted as encouragement to implement the IFA program, or headmasters may have been concerned about how the information from the visits would be used. Even though enumerators did not provide direct encouragement and there were no explicit stakes associated with the data gathered, headmasters may have interpreted the visits as encouragement or assumed that the information would be shared with government officials.

To understand the policy implications, it is worth thinking about the number of visits this study added to the school year (on top of the regular visits local officials might conduct). In every school, we conducted an initial training at the onset of the intervention, a school facilities and staffing survey at the beginning of the intervention, four visits to conduct IFA surveys in months 1, 3, 4 and 5, two visits to record attendance in months 3 and 5, and at least three visits to observe the midday meals in months 3, 4 and 5 (some of these visits overlapped to reduce transportation costs). The schools in the high intensity monitoring treatment received 2 additional visits to observe midday meals in months 1 and 2. It is noteworthy that the addition of 2 visits on top of a base of around 10 had such a substantial effect on the propensity of headmasters to distribute IFA tablets.

One significant difference between the midday meal visits and the other type of visits was timing – they were the only visits that occurred during the meal, which is when headmasters reported distributing IFA tablets. Another unusual element about the midday meal visits relative to most other visits was that enumerators spoke to randomly chosen students. While this also occurred during IFA survey visits, it is possible that the likelihood of discrepancies with student reports would have been more salient to the headmaster in the high intensity monitoring treatment schools (where we had spoken to students three times over the first two months of the intervention) than in other schools (where we had spoken to students only once). During the midday meal visits, enumerators also measured the students' heights.

VI. Conclusion

We study the impacts of two interventions aimed at improving implementation of India's school-based nutrition interventions. We find no evidence for effects of a micronutrient mix on its own. We believe the study may not have had sufficient power for the micronutrients to affect health given the short time period and the low dosage approved by the government, due to the safety restrictions necessary to provide the same intervention to all students. Nonetheless, the fact that we find strong take-up rates of the micronutrient mix, combined with efficacy trials demonstrating the effect of multi-micronutrient supplementation, suggests that school-based multi-micronutrient distribution remains a promising area.

However, we also find evidence supporting the effectiveness of frequent monitoring visits in improving implementation of public health programs. We contribute to the burgeoning literature on the effectiveness of monitoring visits, even with no stakes attached. Our results suggest that the exact timing of such visits and who the auditors speak to may have significant effects.

Finally, we demonstrate two important caveats to consider in the implementation of government programs through schools. First, because MNM schools had worse IFA implementation, the results suggest that implementing new programs may crowd-out other school activities. Second, our findings suggest that school-based programs are less likely to benefit the sickest children due to low micronutrient doses and because these programs rely on regular school attendance.

While schools are a natural setting for implementing social programs for children, it is unclear what the optimal number and types of programs should be, and how to hire and incentivize school officials to implement the programs effectively. Our study has highlighted several important mechanisms that can influence the effectiveness of such programs. As this is an area that is currently understudied in the literature, further research is needed to understand the functioning of similar programs in different contexts, which could lead to broader policy guidance on these issues.

REFERENCES

Abrams, S.A., M. Hettiarachchi, D.C. Hilmers, C. Liyanage, and R. Wickremasinghe. 2008. "The efficacy of micronutrient supplementation in reducing the prevalence of anaemia and deficiencies of zinc and iron among adolescents in Sri Lanka." *European Journal of Clinical Nutrition* 62 (7): 856

Ahmed, F., M.R. Khan, M. Akhtaruzzaman, R. Karim, G. Williams, H. Torlesse, I. D. Hill, N. Dalmiya, C.P. Banu, and B. Nahar. 2010. "Long-term intermittent multiple micronutrient supplementation enhances hemoglobin and micronutrient status more than iron and folic acid supplementation in Bangladeshi rural adolescent girls with nutritional anemia." *The Journal of Nutrition* 140 (10): 1879–86.

Banerjee, A., S. Barnhardt and E. Duflo. 2018. "Can iron-fortified salt control anemia? Evidence from two experiments in rural Bihar," *Journal of Development Economics*, 133: 127-146.

Banerjee, A., E. Duflo and R. Glennerster. 2011. **"**Is decentralized iron fortification a feasible option to fight anemia among the poorest?" In *Explorations in the Economics of Aging*, Ed. David Wise

Berry, J., S. Mehta, P. Mukherjee, H. Ruebeck, and G.K. Shastry. 2018. "Implementation and Effects of India's National School-Based Iron Supplementation Program." *Working paper, Wellesley College.*

Best, C., N. Neufingerl, J.M. Del Rosso, C. Transler, T. van den Briel, and S. Osendarp. 2011. "Can multi-micronutrient food fortification improve the micronutrient status, growth, health, and cognition of schoolchildren? A systematic review." *Nutrition Reviews* 69:186-204.

Chakraborty, T. and R. Jayaraman. 2016. "Impact of school feeding on cognitive skills: Evidence from India's midday meal scheme." *CESifo Working Paper 5994.*

Debnath, S. and S. Sekhri. 2016. "No free lunch: Using technology to improve the efficacy of school feeding programs," *Working paper, University of Virginia.*

Fawzi, W. W., G. I. Msamanga, R. Kupka, D. Spiegelman, E. Villamor, F. Mugusi, R. Wei, and D. Hunter. 2007. "Multivitamin supplementation improves hematologic status in HIV-infected women and their children in Tanzania." *The American Journal of Clinic Nutrition* 85: 1335-43.

Finan, F., B. Olken and R. Pande. 2015. "The personnel economics of the state." *NBER Working Paper 21825.*

Gera, T., H.P. Sachdev, P. Nestel, and S.S. Sachdev. 2007. "Effect of iron supplementation on haemoglobin response in children: Systematic review of randomised controlled trials." *Journal of Pediatric Gastroenterology and Nutrition* 44 (4): 468–86.

Hirve, S., S. Bhave, A. Bavdekar, S. Naik, A. Pandit, C. Schauer, A. Christofides, Z. Hyder, and S. Zlotkin. 2007. "Low dose sprinkles: An innovative approach to treat iron deficiency Anemia in infants and young children." *Indian Pediatrics* 44 (February).

Hyder, S., M. Ziauddin, F. Haseen, M. Khan, T. Schaetzel, C.S. Jalal, M. Rahman, B. Lönnerdal, V. Mannar, and H. Mehansho. 2007. "A multiple-micronutrient-fortified beverage affects hemoglobin, iron, and vitamin A status and growth in adolescent girls in rural Bangladesh." *The Journal of Nutrition* 137(9): 2147–53.

Hoddinott, J., J. Maluccio, J.R. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults." *The Lancet* 371(9610): 411-416.

Hoynes, H., D.W. Schanzenbach and D. Almond. 2016. "Long-run impacts of childhood access to the safetynNet." *American Economic Review* 106(4): 903-934.

International Institute for Population Sciences. 2007. National Family Health Survey (NFHS-3), 2005-06: India. Mumbai, India: International Institute for Population Sciences.

Kar B.R., Rao S.L., Chandramouli B.A., Thennarasu K. (2004). *NIMHANS Neuropsychological Battery for Children-Manual.* Bangalore: NIMHANS Publication Division

Kar B.R., Rao S.L., Chandramouli B.A. (2008). "Cognitive development in children with chronic protein energy malnutrition." *Behavioral and Brain Functions* 4:31

Krämer, M., S. Kumar and S. Vollmer. 2018. "Impact of delivering iron-fortified salt through a school feeding program on child health, education and cognition: Evidence from a randomized controlled trial in rural India." *Working paper, Sam Houston State University*.

Kuong, K., A. Laillou, C. Chea, C. Chamnan, J. Berger, and F.T. Wieringa. 2016. "Stability of vitamin A, iron and zinc in fortified rice during storage and its impact on future national standards and programs: Case study in Cambodia." *Nutrients*, 8(1): 51.

Lee, D.S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76(3): 1071-1102.

Luo, R., Y. Shi, L. Zhang, C. Lui, S. Rozelle, B. Sharbono, A. Yue, Q. Zhao, and R. Martorell. 2012. "Nutrition and educational performance in rural China's elementary schools: Results of a randomized control Ttial in Shaanxi province." *Economic Development and Cultural Change* 60 (4): 735–72.

Mehta, S., F.M. Mugusi, R.J. Bosch, S. Aboud, A. Chatterjee, J.L. Finkelstein, M. Fataki, R. Kisenge, and W.W. Fawzi. 2011. "A randomized trial of multivitamin supplementation in children with tuberculosis in Tanzania." *Nutrition Journal.* 10:120.

Miguel, E. and M. Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72(1): 159-217.

Muralidharan, K., J. Das, A. Holla, and A. Mohpal. 2017. "The fiscal cost of weak governance: Evidence from teacher absence in India," *Journal of Public Economics*, Elsevier, vol. 145(C), pages 116-135.

Pershad, D. and N. Wig. 1988. *Handbook for PGI memory scale clinical test.* Agra, Uttar Pradesh, India: National Psychological Corporation.

Ramakrishnan U., N. Aburto, G. McCabe, and R. Martorell. "Multi-micronutrient interventions but not vitamin A or iron interventions alone improve child growth: results of 3 meta-analyses." *Journal of Nutrition* 134:2592–2602.

Shastry, G.K. and D. Weil. 2003. "How much of income variation is explained by health?" *Journal of the European Economic Association* 1(2-3): 387-396.

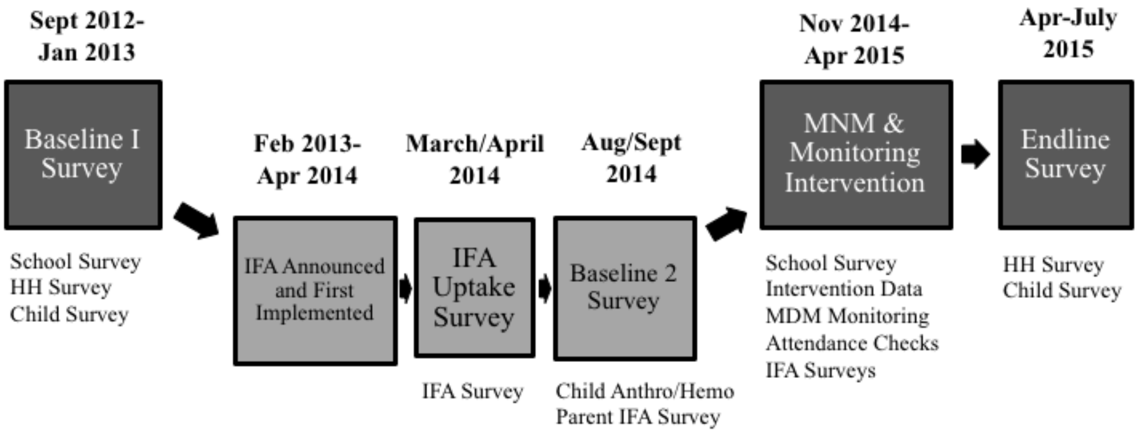
Tee, E.S., M. Kandiah, N. Awin, S. Chong, N. Satgunasingam, L. Kamarudin, S. Milani, A.E. Dugdale, and F.E. Viteri. 1999. "School-administered weekly iron-folate supplements improve hemoglobin and ferritin concentrations in Malaysian adolescent girls." *The American Journal of Clinical Nutrition* 69 (6):1249–56.

Thomas, D., E. Frankenberg, J. Friedman, J. Habicht, N. Ingwersen, C. McKelvey, M. Hakimi, Jaswadi, G. Pelto, B. Sikoki, T. Seeman, J.P. Smith, C. Sumantri, W. Suriastini, S. Wilopo 2006. "Causal effect of health on labor market outcomes: Experimental evidence," *Working paper, UCLA.*

Vermeersch, C. and M. Kremer. 2005 "School meals, educational achievement and school competition: Evidence from a randomized evaluation." *World Bank Policy Research Working Paper Series No. 2523.*

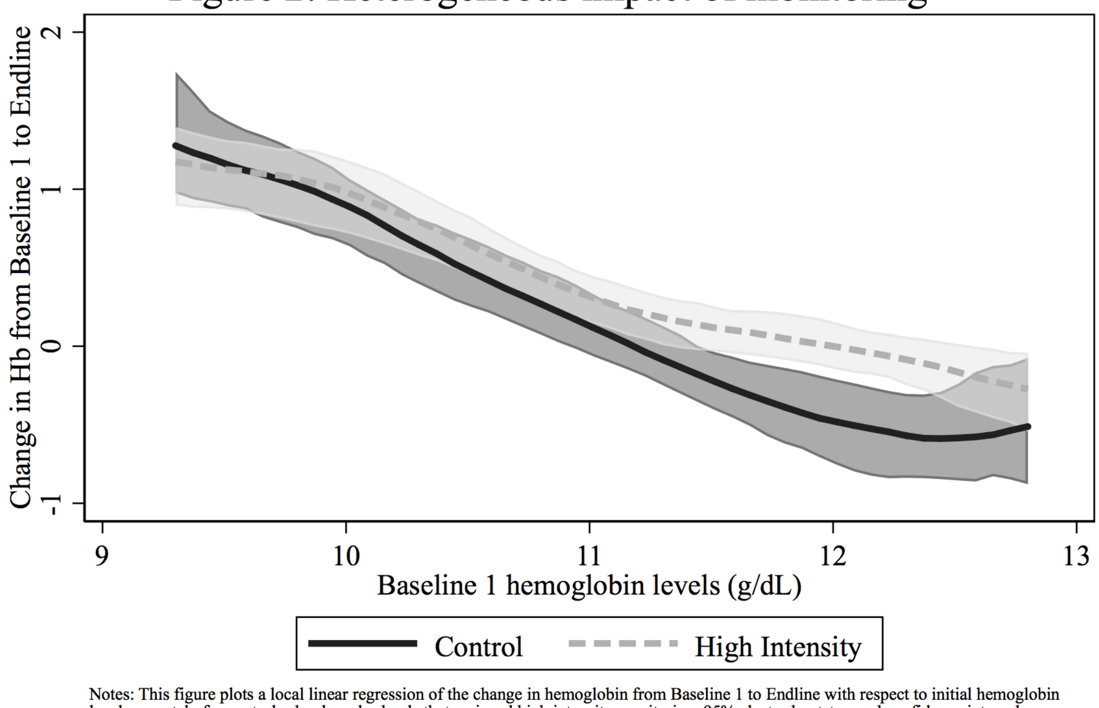
World Health Organization. 2011. "Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity." *Vitamin and Mineral Nutrition Information System.* Geneva, World Health Organization. http://www.who.int/vmnis/indicators/haemoglobin.pdf.

Figure 1: Timeline of key activities



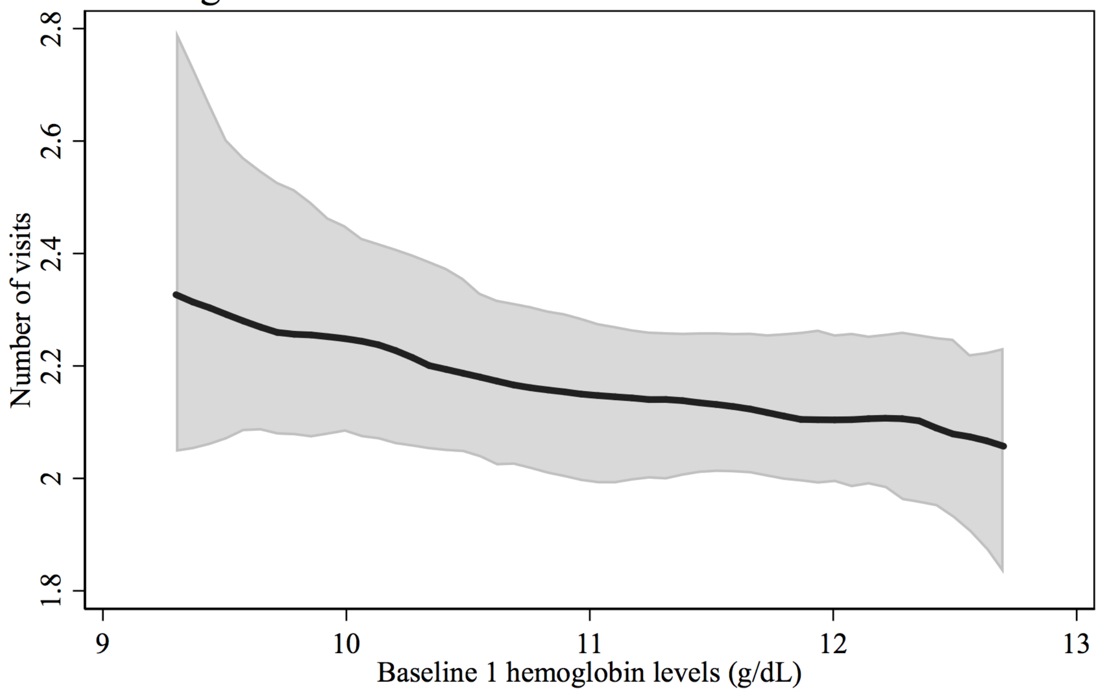
Notes: The household survey includes information about household demographics, assets, etc. The child survey includes health outcomes (height, weight, MUAC, hemoglobin level) and education and cognitive ability measures. The school survey includes information on teachers, school assets, infrastructure, and systems, and implementation of existing school programs.

Figure 2: Heterogeneous impact of monitoring



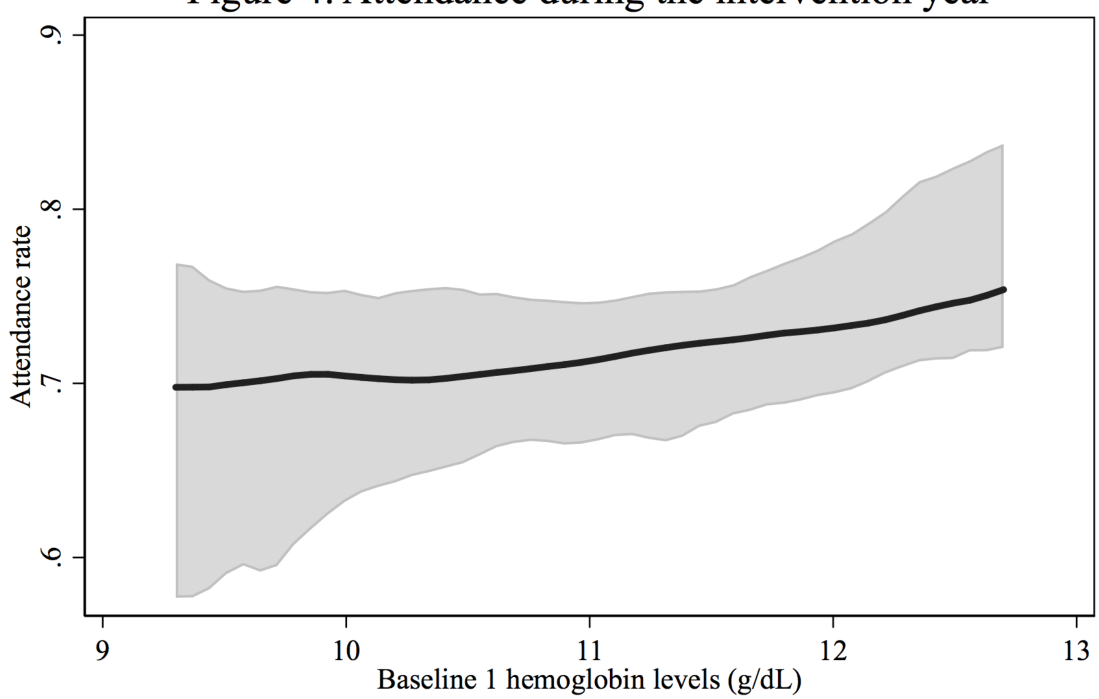
Notes: This figure plots a local linear regression of the change in hemoglobin from Baseline 1 to Endline with respect to initial hemoglobin level separately for control schools and schools that recieved high intensity monitoring. 95% cluster-bootstrapped confidence intervals are plotted along with each estimate.

Figure 3: Number of visits before Hb was measured



Notes: This figure plots a local linear regression of the number of visits made to school before a student's hemoglobin level was measured, with respect to initial hemoglobin level. 95% cluster-bootstrapped confidence intervals are plotted along with each estimate.

Figure 4: Attendance during the intervention year



Notes: This figure plots a local linear regression of student attendance during the intervention year with respect to initial hemoglobin level. 95% cluster-bootstrapped confidence intervals are plotted along with each estimate.

Table 1: Treatment arms

		Monitoring intensity	
		High	**Low**
MNM treatment	Schools:	37	38
	Students Targeted:	3358	3611
Meal provider education and micronutrient mix provision	Students Surveyed at Endline:	680	698
Status quo meals	Schools:	36	37
	Students Targeted:	3074	3649
	Students Surveyed at Endline:	672	670

Table 2: Balance across treatments at baseline: School characteristics

	Control	Only MNM	Only high intensity	Both	P-value of all 3 differences
Panel A: At Baseline 1					
Distance to the block headquarters (km)	22.973	22.789	24.861	24.889	0.815
Primary enrollment	85.351	84.868	74.028	74.378	0.686
Secondary enrollment	13.270	10.132	11.361	16.378	0.814
Number of teachers	2.514	2.421	2.472	2.486	0.994
Number of female teachers	2.757	2.868	2.528	2.676	0.641
Number of rooms	4.455	4.444	4.057	3.778	0.516
Percent of schools have a kitchen	0.784	0.833	0.800	0.676	0.462
Percent of schools have at least one latrine	0.838	0.789	0.889	0.865	0.693
Percent of schools have sufficient water	0.778	0.667	0.735	0.622	0.474
Percent of schools with treated water	0.324	0.263	0.286	0.243	0.887
Percent with parent group for MDM	0.394	0.444	0.471	0.343	0.713
Percent with MDM training	0.389	0.324	0.314	0.333	0.918
Percent receiving MDM rice on a regular schedule	0.472	0.486	0.400	0.278 *	0.225
Panel B: Before the intervention					
Received IFA during previous year	0.811	0.842	0.917	0.892	0.540
Panel C: First month of the intervention					
Primary enrollment	86.838	83.526	68.333	73.432	0.466
Secondary enrollment	10.595	10.395	11.278	15.622	0.823
Number of teachers	3.162	3.105	3.250	3.108	0.986
Number of female teachers	3.108	3.711	2.972	3.216	0.248
Number of rooms	4.611	4.514	4.556	4.622	0.995
Percent of schools have a kitchen	0.784	0.886	0.818	0.833	0.688
Percent of schools have at least one latrine	0.917	0.947	0.944	0.944	0.957
Percent of schools have sufficient water	0.706	0.556	0.667	0.706	0.531
Percent of schools with treated water	0.514	0.263 **	0.382	0.270 **	0.091
Percent with parent group for MDM	0.333	0.342	0.343	0.243	0.732
Percent with MDM training	0.559	0.343 *	0.441	0.545	0.232
Percent receiving MDM rice on a regular schedule	0.657	0.684	0.588	0.514	0.448
Received IFA tablets this year	1.000	1.000	1.000	0.973	0.319
Number of tablets distributed per child past week (school report)	0.838	1.105	0.889	0.889	0.621
Percent of students who say they get meds weekly or more frequently (out of 3)	0.417	0.356	0.455	0.480	0.713
Number of schools	37	38	38	37	

*Notes: This table presents balance checks on school characteristics at Baseline 1, across each of the treatment groups. Each row shows the mean for that variable for the following groups: (i) schools that received no treatment, (ii) schools that only received the MNM treatment, (iii) schools that only received the high intensity monitoring, and (iv) schools that received both MNM and high intensity monitoring treatments. Significance levels of the difference with the control group are indicated after each number, with standard errors robust to heteroskedasticity. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively. The final column provides the p-value for the F-test that the differences across all four groups are zero.*

Table 3: Balance across treatments at baseline: Child characteristics

	Control	Only MNM	Only high intensity	Both	P-value of all 3 differences
Panel A: Child health outcomes at Baseline 1					
Hemoglobin	11.097	11.063	11.170	11.027	0.698
z - weight	-1.839	-1.944	-1.811	-1.953	0.442
z - height	-1.351	-1.366	-1.511	-1.397	0.849
MUAC	15.066	15.174	15.180	15.106	0.807
Panel B: Child health outcomes at Baseline 2					
Hemoglobin	11.214	11.330	11.284	11.108	0.790
z - weight	-1.909	-1.778	-1.929	-2.072	0.532
z - height	-1.534	-1.495	-1.678	-1.891	0.291
MUAC	15.606	15.546	15.900	15.770	0.222
Panel C: Child demographics					
Age (Baseline 1)	6.749	6.720	6.995	6.614	0.753
Female dummy	0.475	0.483	0.480	0.499	0.921
Not child of head of household	0.135	0.123	0.136	0.124	0.902
Number of times child had MDM in past week	4.749	4.760	4.847	4.838	0.940
Takes any supplements	0.000	0.003	0.020 **	0.010 *	0.013
Has taken deworming pill in past year	0.128	0.117	0.101	0.122	0.803
Birth order	2.087	2.119	1.999	1.960	0.254
Panel D: Household demographics					
Non scheduled caste/tribe	0.050	0.030	0.072	0.060	0.157
Owns phone	0.422	0.418	0.413	0.415	0.997
Has electricity	0.531	0.505	0.616	0.504	0.115
House is *pucca*	0.117	0.118	0.106	0.112	0.972
Is satisfied with school meals	0.893	0.866	0.871	0.904	0.510
Has heard of anemia	0.094	0.076	0.088	0.067	0.676
Panel E: Mother demographics					
Age (Baseline 1)	31.276	31.206	30.955	30.805	0.858
Is literate	0.413	0.366	0.378	0.402	0.779
Completed primary school	0.027	0.026	0.023	0.019	0.851
Completed middle school	0.029	0.022	0.018	0.037	0.307
Completed high school	0.014	0.006	0.014	0.007	0.498
Not housewife	0.327	0.393	0.380	0.455 **	0.020
Has a job card	0.623	0.686	0.634	0.638	0.506
Panel F: Head of household demographics					
Age (Baseline 1)	38.990	37.646 *	38.990	37.794	0.144
Is literate	0.531	0.588	0.546	0.575	0.547
Completed primary school	0.028	0.038	0.049	0.060 **	0.077
Completed middle school	0.030	0.050	0.055 *	0.052 *	0.146
Completed high school	0.024	0.018	0.025	0.019	0.802
Occupation in sgriculture	0.495	0.479	0.460	0.450	0.730
Has a job card	0.720	0.783	0.738	0.698	0.119

*Notes: This table presents balance checks on demographic characteristics and child health at baseline, across each of the treatment groups for children who have endline data. Recall that not all children were surveyed at Baseline 1. Children that were added to the sample at Baseline 2 are not included in Panel A, and in Panels C-F, values for those children are filled in from the Endline survey if the variable is time-invariant or unrelated to treatment. Each row shows the mean for that variable for the following groups: (i) schools that received no treatment, (ii) schools that only received the MNM treatment, (iii) schools that only received the high intensity monitoring, and (iv) schools that received both MNM and high intensity monitoring treatments. Significance levels of the difference with the control group are indicated after each number, with standard errors clustered by school. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively. The final column provides the p-value for the F-test that the differences across all four groups are zero.*

Table 4: Take-up of MNM by schools

	Number of MNM deliveries		Amount of MNM delivered (kilos)		Amount of MNM used (kilos)	
	(1)	(2)	(3)	(4)	(5)	(6)
High intensity	0.063	0.062	-0.413	-0.392	-0.331	-0.311
	(0.122)	(0.122)	(3.798)	(3.748)	(4.649)	(4.627)
Number of children enrolled	-0.000	-0.000	0.646***	0.648***	0.637***	0.639***
	(0.001)	(0.001)	(0.050)	(0.049)	(0.056)	(0.055)
Received IFA during previous year		0.119		-9.984		-9.491
		(0.233)		(6.214)		(9.025)
N	73	73	72	72	72	72
R-squared	0.062	0.066	0.909	0.912	0.860	0.863
Dep. var mean, non-high intensity	2.757	2.757	64.324	64.324	58.635	58.635

*Notes: The dependent variables are: (i) the number of MNM deliveries made to the school, (ii) the amount of MNM delivered to the school in kilograms, and (iii) the amount of MNM used in kilograms. All columns include block fixed effects. Robust standard errors are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 5: Take-up of MNM, as seen in MDM observations

	Enumerator detected powdery addition in meal			Cook claims he/she added MNM mix			MNM present in store room		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All months									
MNM treatment	0.131***	0.132***	0.131***	0.715***	0.646***	0.715***	0.452***	0.411***	0.453***
	(0.023)	(0.037)	(0.023)	(0.035)	(0.064)	(0.035)	(0.033)	(0.054)	(0.032)
High intensity	-0.006	-0.006	-0.006	0.066*	0.012	0.067*	0.050	0.015	0.056
	(0.022)	(0.012)	(0.023)	(0.039)	(0.023)	(0.039)	(0.040)	(0.019)	(0.041)
MNM treatment * high intensity		-0.001			0.111			0.067	
		(0.047)			(0.075)			(0.068)	
Received IFA during previous year			-0.001			-0.027			-0.111*
			(0.032)			(0.081)			(0.061)
N	554	554	554	536	536	536	532	532	532
R-squared	0.087	0.087	0.087	0.573	0.576	0.573	0.297	0.299	0.303
p-value of F-test (high & interaction)		0.880	.	.	0.234	.	.	0.460	.
Panel B: December to January - High intensity schools (2 visits each)									
MNM treatment	0.152***	0.152***	0.153***	0.667***	0.667***	0.668***	0.450***	0.450***	0.449***
	(0.047)	(0.047)	(0.047)	(0.067)	(0.067)	(0.066)	(0.061)	(0.061)	(0.059)
N	139	139	139	133	133	133	131	131	131
R-squared	0.105	0.105	0.106	0.522	0.522	0.523	0.315	0.315	0.343
Panel C: February to March - All schools (2 visits each)									
MNM treatment	0.095***	0.123***	0.095***	0.751***	0.648***	0.752***	0.482***	0.436***	0.483***
	(0.028)	(0.044)	(0.028)	(0.045)	(0.068)	(0.045)	(0.043)	(0.059)	(0.043)
High intensity	-0.023	0.006	-0.022	0.074	-0.028	0.076*	0.056	0.007	0.059
	(0.026)	(0.008)	(0.027)	(0.045)	(0.025)	(0.044)	(0.045)	(0.010)	(0.045)
N	277	277	277	269	269	269	267	267	267
R-squared	0.075	0.080	0.076	0.623	0.634	0.624	0.319	0.322	0.320
Panel D: April to May - All schools (1 visit each)									
MNM treatment	0.179***	0.138**	0.177***	0.676***	0.629***	0.677***	0.401***	0.366***	0.402***
	(0.047)	(0.063)	(0.047)	(0.058)	(0.082)	(0.058)	(0.059)	(0.084)	(0.059)
High intensity	0.036	-0.003	0.031	0.049	0.001	0.052	0.030	-0.004	0.038
	(0.046)	(0.015)	(0.047)	(0.055)	(0.024)	(0.056)	(0.059)	(0.028)	(0.061)
N	138	138	138	134	134	134	134	134	134
R-squared	0.141	0.146	0.145	0.573	0.576	0.573	0.325	0.327	0.330

*Notes: This table reports take-up as inferred during the MDM observation visits conducted by our enumerators. The outcomes measured were (i) the likelihood of our enumerator being able to detect the MNM mix directly on inspection of the container in which the meal was cooked (Columns 1-3), (ii) the likelihood that the cook self reports that he/she added the mix (Columns 4-6), as well as (iii) the likelihood that the mix was present in the room where food materials are stored (Columns 7-9). All columns include block fixed effects, survey month fixed effects and a control for the school's total enrollment. While not always shown in the table, columns 2,5, and 8 always include the interaction term between the two treatments and columns 3, 6 and 9 always include a control for whether the school received the IFA tablets during the previous school year. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 6: Take-up of MNM, as seen in micronutrient levels from lab tests of food samples

	February						April					
	Vitamin A			Zinc			Vitamin A			Zinc		
MNM treatment	351.9***	347.4***	345.4***	16.6***	14.5***	14.6***	165.8***	181.9***	181.3***	15.6***	16.6***	16.5***
	(44.8)	(65.4)	(65.5)	(2.8)	(4.2)	(4.2)	(33.4)	(52.7)	(52.2)	(4.4)	(5.8)	(5.9)
High intensity	-5.3	-10.0	-16.2	1.3	-0.8	-0.6	-5.6	10.3	9.3	5.7	6.7	6.5
	(44.6)	(25.4)	(27.8)	(2.8)	(2.1)	(2.3)	(31.7)	(31.2)	(31.4)	(4.5)	(6.1)	(6.1)
MNM treatment * high intensity		9.3	12.7		4.1	4.0		-32.3	-31.4		-2.1	-1.9
		(90.8)	(91.7)		(5.6)	(5.6)		(67.2)	(66.5)		(8.9)	(8.9)
Received IFA during previous year			67.1			-2.5			11.0			1.7
			(77.6)			(5.8)			(57.4)			(5.8)
N	148	148	148	148	148	148	145	145	145	145	145	145
R-squared	0.307	0.307	0.311	0.214	0.217	0.219	0.154	0.156	0.156	0.101	0.101	0.101
Dep. var mean, control group	52.4	52.4	52.4	5.4	5.4	5.4	55.2	55.2	55.2	8.7	8.7	8.7

*Notes: This table presents the results of the effect of the MNM treatment on the micronutrients (namely, vitamin A and zinc) present in school meals, as measured in the laboratory using samples collected by enumerators during February and April of the treatment year. The recommended daily allowances (RDA) for this age group are 7-9 mg of zinc and 400-600 µg of vitamin A. A back-of-the-envelope calculation using estimates of how much food each child was given suggests that vitamin A intake increased by roughly 30-60% of RDA and zinc intake increased by roughly 20% of RDA. All columns include block fixed effects. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 7: Treatment effects on health outcomes - Lagged dependent variable (LDV) model

Lagged dependent variable from survey:	None	Just Baseline 1	Baseline 1 and Baseline 2 with dummies for missing		
	(1)	(2)	(3)	(4)	(5)
Panel A: Dep var: Hemogloblin (g/dL)					
MNM treatment	-0.044	-0.012	-0.018	0.032	-0.022
	(0.057)	(0.067)	(0.057)	(0.072)	(0.057)
High intensity	0.174***	0.244***	0.179***	0.229***	0.168***
	(0.058)	(0.067)	(0.058)	(0.079)	(0.059)
MNM treatment * high intensity				-0.101	
				(0.114)	
Received IFA during previous year					0.129
					(0.089)
N	1920	1118	1920	1920	1920
R-squared	0.024	0.173	0.129	0.130	0.131
Panel B: Dep var: Anemic					
MNM treatment	-0.000	-0.024	-0.009	-0.023	-0.009
	(0.026)	(0.031)	(0.026)	(0.035)	(0.026)
High intensity	-0.066**	-0.089***	-0.064**	-0.077**	-0.062**
	(0.027)	(0.030)	(0.027)	(0.035)	(0.027)
N	1920	1113	1920	1920	1920
R-squared	0.017	0.136	0.089	0.089	0.089
Panel C: Dep var: BMI-for-age (z-score)					
MNM treatment	-0.066	0.036	-0.033	0.017	-0.034
	(0.056)	(0.066)	(0.047)	(0.065)	(0.046)
High intensity	-0.037	-0.005	-0.059	-0.008	-0.063
	(0.055)	(0.066)	(0.045)	(0.067)	(0.045)
N	1743	964	1743	1743	1743
R-squared	0.009	0.250	0.204	0.205	0.205
Panel D: Dep var: Height-for-age (z-score)					
MNM treatment	-0.067	-0.049	-0.096	-0.109	-0.091
	(0.078)	(0.085)	(0.068)	(0.093)	(0.067)
High intensity	0.044	-0.059	0.062	0.049	0.080
	(0.078)	(0.083)	(0.068)	(0.102)	(0.070)
N	1869	1069	1869	1869	1869
R-squared	0.048	0.234	0.175	0.175	0.178

*Notes: The dependent variable in each specification is child's hemoglobin in g/dl (Panel A), an indicator for whether a child is anemic (Panel B), child's z-score for BMI for age (panel C), and child's z-score for height for age (Panel D). All columns include block and age fixed effects, in addition to the lagged dependent variable as described in the headers. While not always shown in the table, column 4 always includes the interaction term between the two treatments and column 5 always includes a control for whether the school received the IFA tablets during the previous school year. Standard errors, clustered by school, are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 8: Treatment effects on IFA program implementation

Dependent variable:	HM shows enumerator IFA tablet			Number of tablets distributed per child past week (school report)			Percent of students who say they get meds weekly or more frequently (out of 3)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All months (4 visits each)									
MNM treatment	-0.015	-0.039	-0.015	0.058	0.045	0.059	-0.062**	-0.073	-0.063**
	(0.022)	(0.028)	(0.022)	(0.052)	(0.072)	(0.052)	(0.031)	(0.045)	(0.031)
High intensity	-0.017	-0.042	-0.020	0.043	0.029	0.047	0.083**	0.072*	0.079**
	(0.022)	(0.030)	(0.022)	(0.053)	(0.073)	(0.053)	(0.032)	(0.039)	(0.031)
MNM treatment * high intensity		0.049			0.028			0.021	
		(0.044)			(0.100)			(0.064)	
Received IFA during previous year			0.042			-0.067			0.055
			(0.046)			(0.094)			(0.056)
N	557	557	557	555	555	555	538	538	538
R-squared	0.120	0.122	0.122	0.087	0.087	0.088	0.128	0.128	0.129
p-value of F-test (high & interaction)		0.363	.	.	0.702	.	.	0.035	.
Panel B: December-January (1 visit per school)									
MNM treatment	0.009	-0.007	0.009	0.121	0.253	0.125	-0.043	-0.056	-0.044
	(0.024)	(0.037)	(0.023)	(0.127)	(0.187)	(0.129)	(0.077)	(0.110)	(0.076)
High intensity	0.011	-0.005	0.011	-0.082	0.050	-0.056	0.066	0.053	0.041
	(0.024)	(0.036)	(0.023)	(0.138)	(0.177)	(0.131)	(0.078)	(0.111)	(0.077)
MNM treatment * high intensity		0.033			-0.273			0.028	
		(0.044)			(0.243)			(0.157)	
Received IFA during previous year			-0.001			-0.274			0.297***
			(0.043)			(0.253)			(0.082)
N	145	145	145	145	145	145	134	134	134
R-squared	0.041	0.044	0.041	0.100	0.106	0.109	0.139	0.139	0.174
Panel C: February - May (3 visits per school)									
MNM treatment	-0.024	-0.050	-0.024	0.032	-0.030	0.032	-0.065*	-0.069	-0.064*
	(0.030)	(0.037)	(0.029)	(0.048)	(0.064)	(0.048)	(0.035)	(0.054)	(0.035)
High intensity	-0.029	-0.055	-0.033	0.084*	0.021	0.083*	0.090**	0.086*	0.092**
	(0.029)	(0.041)	(0.029)	(0.049)	(0.069)	(0.050)	(0.036)	(0.045)	(0.036)
MNM treatment * high intensity		0.052			0.124			0.008	
		(0.058)			(0.095)			(0.071)	
Received IFA during previous year			0.056			0.010			-0.027
			(0.063)			(0.086)			(0.063)
N	412	412	412	410	410	410	404	404	404
R-squared	0.139	0.141	0.143	0.150	0.155	0.150	0.064	0.064	0.064

*Notes: This table shows treatment effects on measures of how well the government's IFA program was implemented. We use three measures of IFA implementation quality: (i) whether the headmaster shows enumerator an IFA tablet (Columns 1-3), (ii) the number of tablets distributed per child in the past week, as seen in the school report (Columns 4-6), and (iii) the percent of students who say they get the tablets weekly or more frequently, out of three randomly selected students that were asked the question (Columns 7-9). All columns include block fixed effects and survey month fixed effects. While not always shown in the table, columns 2, 5, and 8 always include the interaction term between the two treatments and columns 3, 6, and 9 always include a control for whether the school received the IFA tablets during the previous school year. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 9: Treatment effects on IFA program implementation by school characteristics

	Above-median 'school management quality' index			Four or more teachers help out with midday meal			Midday meal managed by a self-help group			School's drinking water is treated		
	No	Yes	P-value of diff	No	Yes	P-value of diff	No	Yes	P-value of diff	No	Yes	P-value of diff
Panel A: All months (4 visits each)												
MNM treatment	-0.114***	0.012	0.047	-0.128***	-0.047	0.190	-0.070*	-0.050	0.776	-0.091**	0.003	0.155
	(0.043)	(0.047)		(0.046)	(0.041)		(0.036)	(0.063)		(0.039)	(0.053)	
High intensity	0.126***	0.028	0.123	0.190***	0.013	0.003	0.095**	0.069	0.720	0.096**	0.060	0.616
	(0.042)	(0.047)		(0.043)	(0.040)		(0.038)	(0.061)		(0.043)	(0.058)	
N	304	234		209	329		364	167		348	182	
R-squared	0.153	0.131		0.144	0.202		0.158	0.126		0.176	0.092	
Panel B: December-January (1 visit per school)												
MNM treatment	-0.118	0.041	0.316	-0.219*	0.012	0.135	0.029	-0.168	0.283	-0.030	-0.006	0.883
	(0.103)	(0.120)		(0.128)	(0.086)		(0.091)	(0.163)		(0.096)	(0.134)	
High intensity	0.190*	-0.090	0.098	0.136	0.028	0.469	0.121	0.035	0.669	0.106	0.036	0.703
	(0.109)	(0.128)		(0.123)	(0.086)		(0.091)	(0.183)		(0.097)	(0.158)	
N	75	59		54	80		91	42		84	48	
R-squared	0.244	0.121		0.286	0.350		0.195	0.156		0.173	0.217	
Panel C: February - May (3 visits per school)												
MNM treatment	-0.110**	0.002	0.112	-0.098	-0.059	0.602	-0.101**	0.001	0.183	-0.114**	0.018	0.076
	(0.046)	(0.053)		(0.061)	(0.044)		(0.043)	(0.064)		(0.045)	(0.059)	
High intensity	0.116**	0.060	0.441	0.215***	0.008	0.005	0.087**	0.090	0.976	0.094*	0.066	0.714
	(0.049)	(0.054)		(0.057)	(0.045)		(0.044)	(0.061)		(0.049)	(0.060)	
N	229	175		155	249		273	125		264	134	
R-squared	0.099	0.055		0.149	0.066		0.066	0.129		0.086	0.055	

*Notes: This table shows treatment effects on measures of how well the IFA program was implemented in different groups of schools. The dependent variable in all regressions is the percent of students who say they get the tablets weekly or more frequently (out of three that were asked). All columns include block fixed effects and survey month fixed effects and a control for total enrollment at the school. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Appendix A. Sample selection for Baseline 2

As described in the text, Baseline 1 was conducted during the 2012-2013 school year and included 3 students per grade in grades 1 to 5 from all 300 study schools. Baseline 2 was conducted at the beginning of the 2014-2015 school year. This survey was designed primarily for the evaluation of the IFA program described in Berry et al. (2018) and covered the 157 sample schools from the three blocks that had high variation in IFA implementation. Due to budgetary restrictions, the sample of children we surveyed in this second survey was smaller than the sample of Baseline 1. We also took measures to account for children who would have graduated from primary school before the 2013-2014 school year (the first year of the IFA program). We randomly selected 6 students who were in grades 1 to 4 in the original sample in the 2012-2013 school year, and added 3 additional students who were in grades 1 and 2 in the 2014-2015 school year to yield a total of 9 targeted children per school.

Appendix B. Attrition

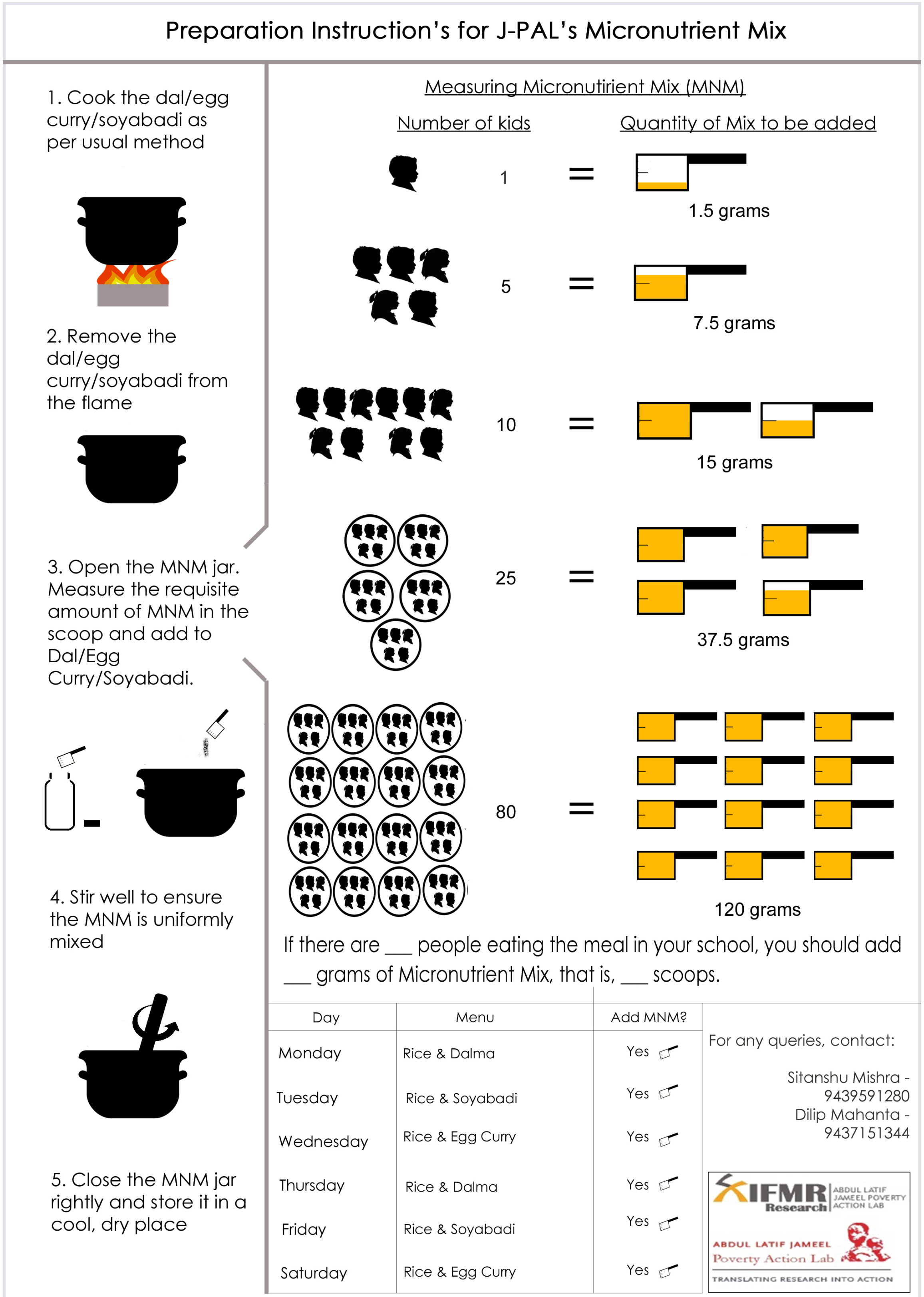
Appendix Table 4 presents results of regressions of attrition at Endline on treatment status. In Columns 1-6, we focus on children who were surveyed at Baseline 1 and are still in primary school during the year of the intervention (those initially enrolled in grades 1-3) and, in Columns 7-12, we include children who were newly sampled at Baseline 2. In our implementation of the survey, we went to great lengths to visit children at both home and at school (after getting the necessary permissions from the parents) and conduct multiple visits if we were initially unable to find the child. Attrition is therefore quite low at only 9 percent, on average. Column 1 in Table 4 indicates that whether the child attended a school that received the MNM or a school that was monitored intensely does not affect the probability of attrition. In Column 2, we add an interaction term between the two treatments. While the coefficients become significant, Columns 3-6 show that the composition of those who attrited is not significantly different across groups. Baseline hemoglobin levels do not affect probability of attrition, and this does not differ by treatment group (Columns 3-4, 9-10). Similarly, grade at baseline does not affect probability of attrition, and this does not differ by treatment group (Columns 5-6, 11-12).

We also compute bounds on our main treatment effect estimates, as developed by Lee (2009). These bounds provide 'worst case' estimates – the sample is trimmed to achieve equal attrition between the treatment and control groups. If there is more attrition in the control group than the treatment group, we calculate lower (upper) bounds by dropping the children in the treatment group with the highest (lowest) hemoglobin levels. If there is more attrition in the treatment group than the control group, we calculate lower (upper) bounds by dropping the children in the control group with the lowest (highest) hemoglobin levels.

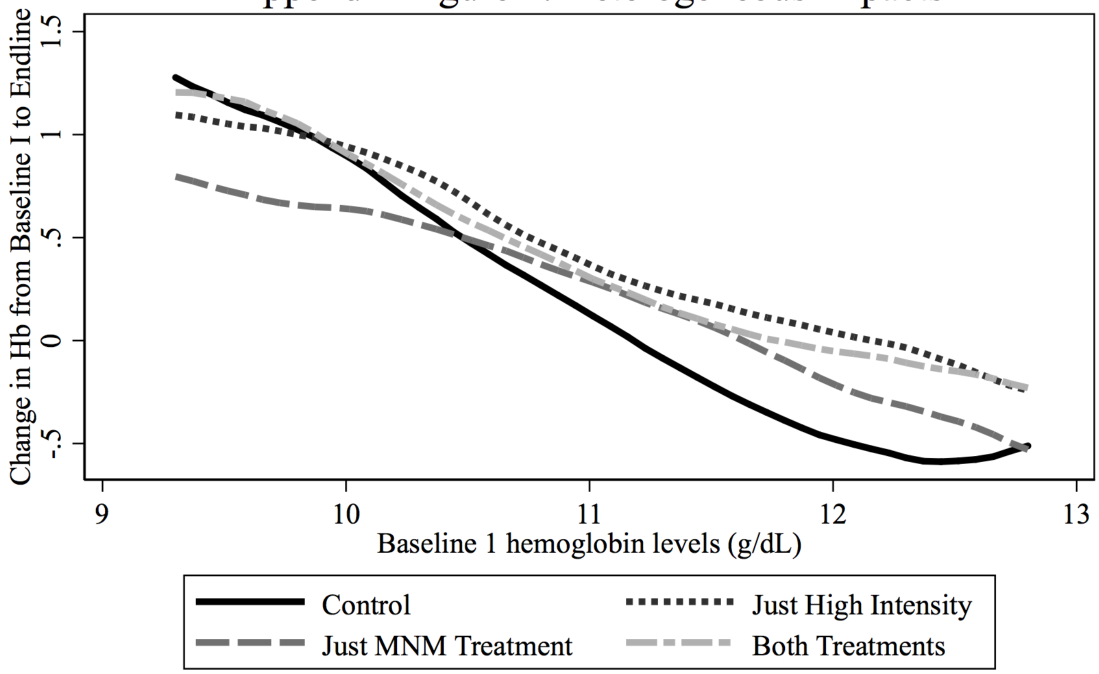
The methods in Lee (2009) are described for a single treatment group and for specifications including no control variables. Therefore, the bounds provided in Appendix Table 5 are estimated separately for each of the three treatment arms compared to the control group (unlike in the main results in the paper where we interact the two treatment variables). They are estimated first with no control variables, using the exact procedure described in Lee (2009), and again using a method in the spirit of Lee (2009) that allows for the inclusion of controls. In this specification, the outcome measure is regressed on all control variables and the trimming described above is done with the residuals rather than with the outcome variable itself. Finally, we estimate the results on three different samples: all students surveyed at Endline, where “attrition” means failure to collect a hemoglobin sample; and all students surveyed at Baseline 1, or all students surveyed in Baseline 1 or 2, where “attrition” means either failure to locate the child at Endline or failure to collect a hemoglobin sample at Endline.

The positive effect of just high intensity monitoring on children’s hemoglobin levels remains highly significant for all samples and specifications, even after accounting for potential differential attrition. The positive effect of high intensity monitoring with the MNM treatment maintains significance when controls are included, but the lower bound is insignificant without the inclusion of controls.

Appendix Figure 1: MNM flier

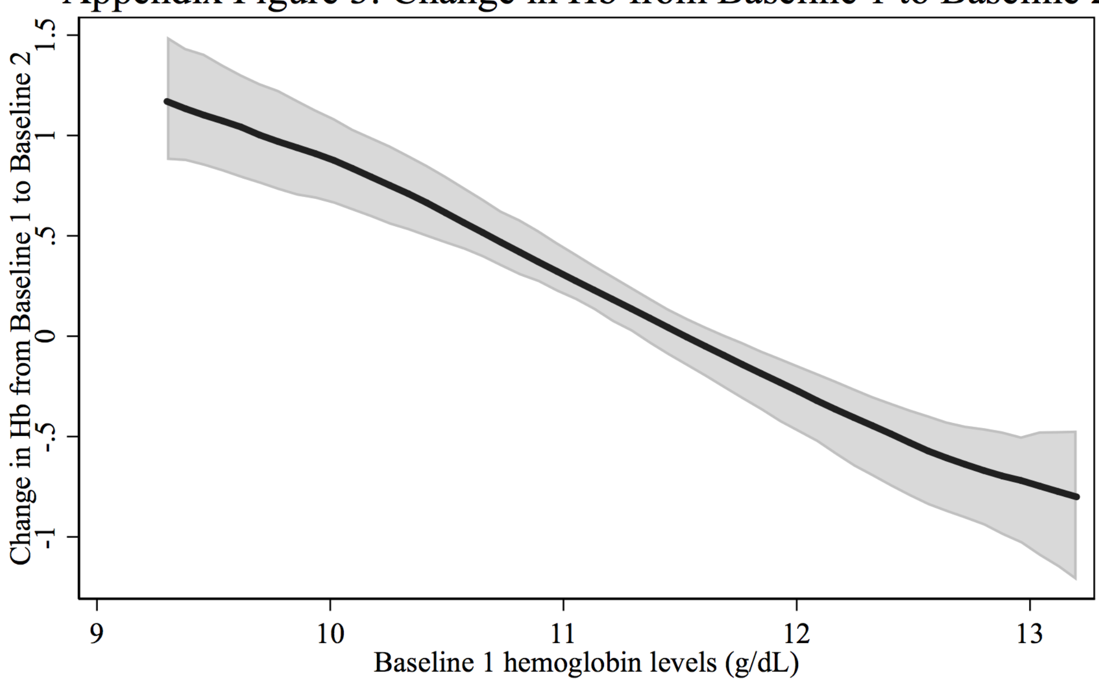


Appendix Figure 2: Heterogeneous impacts



Notes: This figure plots a local linear regression of the change in hemoglobin from Baseline 1 to Endline with respect to initial hemoglobin level separately for each of the four treatment groups. 95% cluster-bootstrapped confidence intervals not shown for readability, but all four confidence intervals overlap.

Appendix Figure 3: Change in Hb from Baseline 1 to Baseline 2



Notes: This figure plots a local linear regression of the change in hemoglobin from Baseline 1 to Baseline 2 (two measurement periods both occuring before the intervention) with respect to initial hemoglobin level. 95% cluster-bootstrapped confidence intervals are plotted along with each estimate.

Appendix Table 1: Midday meal menu for the state of Odisha

DAY	MENU
MONDAY	Rice and Dalma
TUESDAY	Rice and Soyabadi
WEDNESDAY	Rice and Egg Curry
THURSDAY	Rice and Dalma
FRIDAY	Rice and Soyabadi
SATURDAY	Rice and Egg Curry

Appendix Table 2: Treatment effects on health outcomes - LDV Model - Additional outcomes

	None	Just Baseline 1	Baseline 1 and Baseline 2 with dummies for missing		
	(1)	(2)	(3)	(4)	(5)
Panel A: Dep var: BMI (kg/m^2)					
MNM treatment	-0.078	-0.001	-0.053	0.006	-0.055
	(0.076)	(0.101)	(0.064)	(0.079)	(0.064)
High intensity	-0.075	-0.029	-0.100	-0.040	-0.104*
	(0.077)	(0.100)	(0.063)	(0.096)	(0.062)
MNM treatment * high intensity				-0.119	
				(0.129)	
Received IFA during previous yea					0.061
					(0.092)
N	1942	1118	1942	1942	1942
R-squared	0.145	0.309	0.336	0.336	0.336
Panel B: Dep var: Weight (kilos)					
MNM treatment	-0.187	0.068	-0.140	-0.083	-0.131
	(0.220)	(0.137)	(0.125)	(0.179)	(0.123)
High intensity	-0.118	-0.110	-0.063	-0.006	-0.037
	(0.223)	(0.134)	(0.127)	(0.190)	(0.131)
MNM treatment * high intensity				-0.114	
				(0.252)	
Received IFA during previous yea					-0.375
					(0.239)
N	1946	1123	1946	1946	1946
R-squared	0.480	0.757	0.737	0.737	0.737
Panel C: Dep var: Height (cm)					
MNM treatment	-0.352	-0.504	-0.618*	-0.713	-0.579
	(0.461)	(0.461)	(0.373)	(0.497)	(0.367)
High intensity	0.203	-0.082	0.312	0.216	0.428
	(0.465)	(0.450)	(0.382)	(0.563)	(0.388)
MNM treatment * high intensity				0.191	
				(0.750)	
Received IFA during previous yea					-1.662**
					(0.640)
N	1942	1128	1942	1942	1942
R-squared	0.473	0.476	0.584	0.584	0.586

Appendix Table 2: Treatment effects on health outcomes - LDV Model - Additional outcomes

	None	Just Baseline 1	Baseline 1 and Baseline 2 with dummies for missing		
	(1)	(2)	(3)	(4)	(5)
Panel D: Dep var: Mid-upper-arm circumference (cm)					
MNM treatment	-0.123	-0.137	-0.156**	-0.141	-0.154**
	(0.090)	(0.093)	(0.071)	(0.102)	(0.071)
High intensity	-0.018	-0.078	-0.069	-0.055	-0.063
	(0.091)	(0.103)	(0.075)	(0.109)	(0.076)
MNM treatment * high intensity				-0.029	
				(0.140)	
Received IFA during previous yea					-0.093
					(0.125)
N	1946	1127	1946	1946	1946
R-squared	0.315	0.407	0.497	0.497	0.497
Panel E: Dep var: Weight-for-age (z score)					
MNM treatment	-0.120*	0.012	-0.065	0.045	-0.063
	(0.066)	(0.062)	(0.053)	(0.076)	(0.053)
High intensity	-0.038	-0.104	-0.050	0.061	-0.046
	(0.068)	(0.069)	(0.053)	(0.075)	(0.054)
MNM treatment * high intensity				-0.217**	
				(0.104)	
Received IFA during previous yea					-0.054
					(0.092)
N	1158	476	1158	1158	1158
R-squared	0.033	0.496	0.308	0.311	0.308

*Notes: The dependent variable in each specification is child's BMI in kg/m^2 (Panel A), child's weight measured in kg (Panel B), child's height in cm (Panel C), mid-upper arm circumference in cm (Panel D), and child's z-score for weight for age (Panel E). All columns include block and age fixed effects, in addition to the lagged dependent variable as described in the headers. Column 5 includes a dummy variable for receiving the IFA tablets during the previous year. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Appendix Table 3: Treatment effects on health outcomes - difference in difference estimates

	Age, block F.E.	School F.E.			Child F.E.
	(1)	(2)	(3)	(4)	(5)
Panel A: Dep var: Hemogloblin (g/dL)					
Endline * MNM treatment	0.049	0.025	0.016	0.020	0.035
	(0.081)	(0.084)	(0.104)	(0.083)	(0.123)
Endline * High intensity	0.189**	0.209**	0.200*	0.188**	0.243*
	(0.082)	(0.085)	(0.115)	(0.084)	(0.124)
Endline * High intensity * MNM treatment			0.018		
			(0.168)		
Baseline 2 * Received IFA during previous year				0.030	
				(0.243)	
Endline * Received IFA during previous year				0.225	
				(0.159)	
N	3452	3452	3452	3452	3452
R-squared	0.041	0.109	0.109	0.110	0.737
p-value of F-test (high & interaction)	.	.	0.052	.	.
Panel B: Dep var: Anemic					
Endline * MNM treatment	-0.027	-0.016	0.002	-0.016	-0.026
	(0.036)	(0.038)	(0.047)	(0.038)	(0.059)
Endline * High intensity	-0.056	-0.065*	-0.046	-0.061	-0.062
	(0.036)	(0.038)	(0.050)	(0.038)	(0.059)
Endline * High intensity * MNM treatment			-0.036		
			(0.074)		
Baseline 2 * Received IFA during previous year				-0.014	
				(0.118)	
Endline * Received IFA during previous year				-0.046	
				(0.083)	
N	3452	3452	3452	3452	3452
R-squared	0.028	0.092	0.093	0.093	0.707
p-value of F-test (high & interaction)	.	.	0.224	.	.
Panel C: Dep var: BMI-for-age (z-score)					
Endline * MNM treatment	0.047	0.045	0.097	0.045	0.052
	(0.054)	(0.057)	(0.080)	(0.057)	(0.088)
Endline * High intensity	-0.061	-0.051	0.001	-0.050	-0.015
	(0.051)	(0.055)	(0.078)	(0.056)	(0.087)
Endline * High intensity * MNM treatment			-0.104		
			(0.115)		
Baseline 2 * Received IFA during previous year				0.041	
				(0.106)	
Endline * Received IFA during previous year				0.009	
				(0.105)	
N	3176	3176	3176	3176	3176
R-squared	0.014	0.103	0.103	0.103	0.809

Appendix Table 3: Treatment effects on health outcomes - difference in difference estimates

	Age, block F.E.	School F.E.			Child F.E.
	(1)	(2)	(3)	(4)	(5)
Panel D: Dep var: Height-for-age (z-score)					
Endline * MNM treatment	-0.072	-0.067	-0.031	-0.064	-0.145
	(0.091)	(0.095)	(0.112)	(0.095)	(0.131)
Endline * High intensity	0.206**	0.158	0.194	0.174*	0.055
	(0.094)	(0.099)	(0.150)	(0.100)	(0.120)
Endline * High intensity * MNM treatment			-0.071		
			(0.191)		
Baseline 2 * Received IFA during previous year				0.013	
				(0.173)	
Endline * Received IFA during previous year				-0.172	
				(0.161)	
N	3391	3391	3391	3391	3391
R-squared	0.044	0.131	0.131	0.131	0.807
Panel E: Dep var: BMI (kg/m^2)					
Endline * MNM treatment	0.005	-0.002	0.052	-0.001	-0.010
	(0.067)	(0.072)	(0.086)	(0.072)	(0.132)
Endline * High intensity	-0.110*	-0.102	-0.048	-0.099	-0.039
	(0.065)	(0.070)	(0.102)	(0.071)	(0.130)
Endline * High intensity * MNM treatment			-0.108		
			(0.145)		
Baseline 2 * Received IFA during previous year				0.065	
				(0.132)	
Endline * Received IFA during previous year				-0.003	
				(0.139)	
N	3463	3463	3463	3463	3463
R-squared	0.134	0.209	0.209	0.209	0.809
p-value of F-test (high & interaction)	.	.	0.259	.	.
Panel F: Dep var: Weight (kilos)					
Endline * MNM treatment	0.055	0.054	-0.032	0.061	-0.093
	(0.168)	(0.176)	(0.242)	(0.175)	(0.203)
Endline * High intensity	-0.078	-0.142	-0.229	-0.108	-0.172
	(0.168)	(0.179)	(0.274)	(0.187)	(0.201)
Endline * High intensity * MNM treatment			0.174		
			(0.355)		
Baseline 2 * Received IFA during previous year				0.103	
				(0.414)	
Endline * Received IFA during previous year				-0.339	
				(0.363)	
N	3474	3474	3474	3474	3474
R-squared	0.508	0.567	0.567	0.567	0.947

Appendix Table 3: Treatment effects on health outcomes - difference in difference estimates

	Age, block F.E.	School F.E.			Child F.E.
	(1)	(2)	(3)	(4)	(5)
Panel G: Dep var: Height (cm)					
Endline * MNM treatment	-0.501	-0.440	-0.466	-0.419	-0.856
	(0.508)	(0.531)	(0.624)	(0.525)	(0.742)
Endline * High intensity	0.995*	0.756	0.730	0.867	0.247
	(0.522)	(0.553)	(0.847)	(0.559)	(0.687)
Endline * High intensity * MNM treatment			0.052		
			(1.064)		
Baseline 2 * Received IFA during previous year				-0.153	
				(0.943)	
Endline * Received IFA during previous year				-1.342	
				(0.926)	
N	3473	3473	3473	3473	3473
R-squared	0.466	0.514	0.514	0.514	0.895
Panel H: Dep var: Mid-upper-arm circumference (cm)					
Endline * MNM treatment	-0.126	-0.127	-0.187	-0.125	-0.194
	(0.085)	(0.089)	(0.147)	(0.089)	(0.131)
Endline * High intensity	-0.090	-0.086	-0.146	-0.076	-0.094
	(0.091)	(0.097)	(0.122)	(0.100)	(0.143)
Endline * High intensity * MNM treatment			0.119		
			(0.185)		
Baseline 2 * Received IFA during previous year				-0.027	
				(0.164)	
Endline * Received IFA during previous year				-0.124	
				(0.150)	
N	3477	3477	3477	3477	3477
R-squared	0.322	0.392	0.392	0.392	0.868
Panel I: Dep var: Weight-for-age (z score)					
Endline * MNM treatment	-0.005	0.003	0.104	0.004	-0.005
	(0.076)	(0.079)	(0.115)	(0.079)	(0.068)
Endline * High intensity	-0.020	-0.021	0.083	-0.014	-0.073
	(0.077)	(0.081)	(0.113)	(0.082)	(0.067)
Endline * High intensity * MNM treatment			-0.203		
			(0.160)		
Baseline 2 * Received IFA during previous year				0.038	
				(0.168)	
Endline * Received IFA during previous year				-0.063	
				(0.149)	
N	2437	2437	2437	2437	2437
R-squared	0.022	0.145	0.146	0.146	0.961

*Notes: The dependent variable in each specification is child's hemoglobin in g/dl (Panel A), an indicator for whether a child is anemic (Panel B), child's z-score for BMI for age (Panel C), child's z-score for height for age (Panel D), child's BMI in kg/m2 (Panel E), child's weight measured in kg (Panel F), child's height in cm (Panel G), mid-upper arm circumference in cm (Panel H), and child's z-score for weight for age (Panel I). All columns include fixed effects for age and block interacted with Baseline 2 or Endline in addition to the controls and fixed effects indicated in the headers. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Appendix Table 4: Analysis of differential attrition

Dependent variable:	Attrited From Baseline 1						Attrited From Baseline 1 or Baseline 2					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MNM treatment	-0.014	-0.052**	0.270	-0.005	-0.028	-0.025	-0.007	-0.023	0.219	0.312	-0.023	-0.013
	(0.016)	(0.022)	(0.194)	(0.313)	(0.044)	(0.069)	(0.019)	(0.027)	(0.209)	(0.345)	(0.041)	(0.064)
High Intensity	-0.007	-0.043*	0.128	-0.135	-0.053	-0.049	-0.021	-0.037	-0.098	-0.002	-0.044	-0.032
	(0.016)	(0.024)	(0.193)	(0.323)	(0.044)	(0.068)	(0.019)	(0.026)	(0.209)	(0.354)	(0.042)	(0.061)
MNM treatment * High Intensity		0.073**		0.501		-0.008		0.032		-0.188		-0.023
		(0.032)		(0.388)		(0.089)		(0.038)		(0.419)		(0.083)
Baseline hemoglobin level			0.020	0.008					0.022	0.028		
			(0.020)	(0.027)					(0.022)	(0.028)		
MNM treatment * baseline hemoglobin level			-0.012	0.008					0.007	-0.003		
			(0.017)	(0.029)					(0.019)	(0.032)		
High Intensity * baseline hemoglobin level			-0.026	-0.004					-0.020	-0.030		
			(0.017)	(0.028)					(0.019)	(0.031)		
MNM treatment * High Intensity * baseline hemoglobin level				-0.039						0.020		
				(0.035)						(0.038)		
Class in school as of baseline					-0.005	0.005					0.001	0.013
					(0.020)	(0.025)					(0.019)	(0.023)
MNM treatment * class at baseline					0.023	0.003					0.020	-0.004
					(0.021)	(0.030)					(0.020)	(0.029)
High Intensity * class at baseline					0.006	-0.013					0.003	-0.020
					(0.021)	(0.032)					(0.020)	(0.031)
MNM treatment * High Intensity * class at baseline						0.041						0.048
						(0.042)						(0.040)
Constant	0.104***	0.123***	-0.122	0.030	0.114**	0.112**	0.158***	0.166***	-0.089	-0.143	0.100**	0.094*
	(0.015)	(0.018)	(0.225)	(0.296)	(0.044)	(0.055)	(0.016)	(0.018)	(0.240)	(0.314)	(0.040)	(0.049)
N	1246	1246	1237	1237	1245	1245	1484	1484	1474	1474	1387	1387
R-squared	0.001	0.005	0.004	0.008	0.003	0.007	0.001	0.002	0.005	0.006	0.003	0.008
p-value from F-test of all regressors	0.650	0.105	0.523	0.156	0.554	0.226	0.494	0.525	0.258	0.350	0.419	0.125

*Notes: This table presents an analysis of differential attrition in the child healh surveys. In Columns (1)-(6) the dependent variable is an indicator for attriting from Baseline 1 while in Columns (7)-(12) the dependent variable is an indicator for attriting from either Baseline 1 or Baseline 2. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Appendix Table 5: Lee Bounds on Impact of Treatments on Hemoglobin

	No Controls			Block and Age F.E. + Lagged Hb		
	OLS	Lee (2009) Bounds		OLS	Manual Trimming Bounds	
		Lower	Upper		Lower	Upper
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sample should have Endline						
MNM Treatment Only	0.020	-0.060	0.110	0.040	0.013	0.061
	(0.068)	(0.072)	(0.073)	(0.070)	(0.064)	(0.064)
N	973	--	--	973	969	969
High Intensity Treatment Only	0.255***	0.148**	0.359***	0.240***	0.207***	0.285***
	(0.069)	(0.073)	(0.073)	(0.080)	(0.065)	(0.064)
N	957	--	--	957	952	952
MNM and High Intensity Treatment	0.120*	0.077	0.173**	0.149*	0.128*	0.163**
	(0.068)	(0.076)	(0.076)	(0.083)	(0.066)	(0.066)
N	940	--	--	940	937	937
Panel B: Sample has Baseline 1, should have Endline						
MNM Treatment Only	0.042	-0.102	0.183*	0.082	0.036	0.114
	(0.091)	(0.103)	(0.104)	(0.089)	(0.083)	(0.084)
N	561	--	--	561	557	557
High Intensity Treatment Only	0.344***	0.199*	0.520***	0.343***	0.324***	0.367***
	(0.092)	(0.105)	(0.100)	(0.091)	(0.086)	(0.085)
N	546	--	--	546	545	545
MNM and High Intensity Treatment	0.153*	0.136	0.174	0.232**	0.174**	0.291***
	(0.091)	(0.123)	(0.133)	(0.089)	(0.084)	(0.085)
N	528	--	--	528	521	521
Panel C: Sample has Baseline 1 or Baseline 2, should have Endline						
MNM Treatment Only	0.070	-0.073	0.211**	0.101	0.044	0.144*
	(0.085)	(0.102)	(0.102)	(0.083)	(0.076)	(0.078)
N	633	--	--	633	627	627
High Intensity Treatment Only	0.346***	0.244**	0.455***	0.331***	0.285***	0.384***
	(0.086)	(0.106)	(0.106)	(0.089)	(0.078)	(0.077)
N	611	--	--	611	607	607
MNM and High Intensity Treatment	0.164*	0.073	0.231**	0.224**	0.191**	0.255***
	(0.085)	(0.109)	(0.117)	(0.089)	(0.080)	(0.080)
N	600	--	--	600	596	596

*Notes: The dependent variable in each specification is child's hemoglobin level in g/dl. Each specification is the lagged dependent variable specification, estimated separately for each treatment group, relative to the pure control group. In columns (1)-(3), there are no control variables and the bounds are classic Lee (2009) bounds, estimated using the Stata command described in Tauchmann (2014). In columns (4)-(6), we use a manual trimming method in the spirit of Lee bounds that allows for the inclusion of continuous control variables (hemoglobin levels taken at Baseline 1 or 2), plus indicators for missing lagged hemoglobin as well as age and block fixed effects. See Appendix B for details. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Designed by soapbox.co.uk