Inputs, Monitoring, and Crowd-out in India's School-Based Health Interventions[1]

James Berry, Saurabh Mehta, Priya Mukherjee, Hannah Ruebeck, Gauri Kartini Shastry*

June 2017

PRELIMINARY: PLEASE DO NOT CITE

ABSTRACT

Nutritional deficiencies are widespread in developing countries, affecting child health and learning. In India, nutritional deficiencies have led to efforts by the government to try various strategies to address the issue. Many of these programs rely on school infrastructure for implementation, such as the government mandated Mid-day Meal program, the world's largest school meals program. We evaluate the impact of three such school-based interventions on program implementation and child health in a rural district in India: 1) the distribution of a micronutrient mix (MNM) to be added to the school mid-day meal, using a randomized controlled trial; 2) a new government-run program, the Iron and Folic Acid (IFA) Supplementation program, which provided students with iron tablets, using a difference-in-differences strategy; and 3) monitoring of school meals, using random variation in monitoring intensity. While we find significant and positive effects of distributing the MNM on micronutrients present in meals, we find no detectable effects on hemoglobin levels or on anthropometric measures of child health. We find suggestive evidence that the government's IFA supplementation improved hemoglobin, but our results indicate that the impact depends critically on how well the program is implemented. Increased monitoring of school meals, on the other hand, does improve hemoglobin levels. In addition, even though the additional monitoring only targeted meals, we find positive effects of the increased monitoring on how well the IFA program was implemented in schools. We find no effects of either intervention on learning, cognitive development or attendance. Finally, we find significant negative spillovers of the MNM intervention on how well the IFA program was implemented, suggesting that effort by school officials is crowded out by the introduction of the new MNM program.

[1] We thank The International Initiative for Impact Evaluation (3ie), the Douglas B. Marshall, Jr. Family Foundation and the International Growth Center at LSE 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: Cornell University; Saurabh Mehta: Cornell University; Priya Mukherjee: College of William and Mary, pmukherjee@wm.edu; Hannah Ruebeck: EdLabs, Harvard University; Gauri Kartini Shastry: Wellesley College, gshastry@wellesley.edu.

I. Introduction

Nutritional deficiencies are widespread in developing countries. In India, nutritional deficiencies are especially prevalent. According to the National Family Health Survey- 3 (NFHS-3) in 2005-2006, 43% of children under the age of five years in India are underweight. The same report shows that 70% of children under the age of five suffer from mild, moderate or severe anemia. In the state of Odisha, where we conducted our study, 41% children under the age of five are underweight and 65% are anemic. These deficiencies have substantial consequences for productivity, at the individual level (see, e.g., Thomas et al. 2006), and for economic growth, at the macro-level (see, e.g., Shastry and Weil 2003).

It is no surprise, then, that there is substantial policy interest in addressing these deficiencies through various programs, many of which use existing school infrastructure to distribute nutrients. In 1995, the Supreme Court of India mandated the provision of nutritious mid-day meals to every primary school student. Despite subsequent legislation, the availability and quality of these meals still varies greatly, potentially due to limited resources, inadequate infrastructure, and widespread corruption. In 2012, India's Ministry of Health and Family Welfare launched the Weekly Iron and Folic Acid program, which relies on school infrastructure to distribute iron and folic acid tablets to primary and upper primary school children.

Delivering nutrients through schools is promising because school-age children are an important demographic to target – recent work suggests that early childhood health interventions may have long-lasting effects (see, e.g., Hoddinott et al. 2008) – and because the school system is often the most comprehensive infrastructure available to reach children in remote areas. There are caveats to school-based distribution, however, since i) sicker and younger children may not attend school or pre-school, ii) breaks in the school year can erode health gains, iii) parents may substitute nutrients away from children who receive them in school (Afridi 2010; Jacoby 2002), and iv) new programs to be implemented in school may crowd out existing school activities (such as instruction time, Vermeersch and Kremer 2005).

In this, paper, we evaluate the efficacy of different school-based strategies to provide micronutrient fortification to children in India, using a randomized control evaluation in Keonjhar district in Odisha. In particular, we evaluate the fortification of school meals with a micronutrient mix (MNM) of vitamins and micronutrients, with and without intensive monitoring of the midday meals program in the study schools. 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. We also examine non-experimental variation in the government's implementation of the IFA program in the year before we implemented our intervention, to evaluate the effect of the iron fortification program itself. Specifically, we use a difference-in-differences strategy exploiting variation across schools in whether IFA tablets were received from government officials and time-series variation since we measured hemoglobin before the IFA program was implemented in our study schools.[2] We also exploit variation across children in initial hemoglobin status (those not anemic initially are less likely to respond to supplementation). During our study year, we also investigate interactions between the IFA program and each of the other two treatment arms.

The overall goal of these interventions was to combat micronutrient deficiencies, most notably anemia, among school-age children using the pre-existing school infrastructure. Improving child health through eliminating micronutrient deficiencies is expected to have the additional benefits of improved cognitive ability, school attendance and educational outcomes. The primary outcomes we consider are hemoglobin levels, anthropometric measures, cognitive ability, school attendance and performance on a test in reading and math. However, of first order importance are more intermediary outcomes that we measure through our surveys: take-up of the MNM, quality of the midday meal, and implementation of the IFA program.

Specifically, we begin by measuring whether inputs are consistently distributed: we organized and tracked the training of school staff and the school delivery of MNM mix and determined through surveys whether government officials distributed the IFA tablets as well as provided adequate training and instructions on how to distribute them. We also monitored meals to measure meal quality and consumption and conducted surveys asking students whether they received IFA tablets regularly. Another intermediary objective was for children to attend classes during the school day (not just the meal), which we measured by tracking attendance at a random time on a random day each month. In order to benefit from either intervention, children need to be attending (at least the meal) sufficiently often at baseline in order to increase nutrition; it is

[2] We discuss potentially confounding reasons why IFA tablets were not received at length in section III.B.

unlikely that either the IFA tablet or the MNM increase the incentive to attend school. We expect the interventions to improve nutritional status and health if children consume sufficient quantities of the MNM or IFA tablets. Previous research suggests that these health impacts can be identified within 5-6 months of increased consumption of micronutrients.

We find significant and positive effects of distributing the MNM on micronutrients present in meal samples that were tested in a laboratory. However, we find no detectable effects of the MNM treatment on hemoglobin levels or on anthropometric measures of child health. We find suggestive evidence that the government's IFA supplementation improved hemoglobin, but with the caveat that this impact depends critically on how well the program was implemented (e.g., whether the school ran out of tablets). Increased monitoring of school meals, on the other hand, has significant impacts on our primary outcome of interest: child hemoglobin levels. Consistent with this result, we find evidence that suggests that there was a positive effect of the increased monitoring on how well the IFA program was implemented, even though the increased monitoring specifically targeted the quality of school meals and not the IFA implementation.

Finally, we find significant negative spillovers of the MNM intervention on how well the IFA program was implemented. This suggests that effort exerted by school officials on distributing the iron tablets, was crowded out by the introduction of the new MNM program.

This paper relates to a number of literatures. First, our paper relates to a broad literature on the delivery of publicly provided goods and services. This is a large but growing literature using either field experiments, or administrative data, that evaluates multiple strategies to improve service delivery in the developing world (Olken, 2005; Bjorkman and Svensson, 2009; Duflo et. al., 2012; Miller et. al., 2012; Muralidharan et. al, 2014, Aker and Ksoll, 2015; Muralidharan et. al. 2016, among others). Such strategies include the use of technology in monitoring public officials, financial and non-financial incentives for service providers, and more traditional top down or bottom up monitoring. Perhaps the most closely related work is a recent study using administrative data (Debnath and Sekhri, 2016), which shows suggestive evidence that a mobile-based monitoring technology might reduce malfeasance and shirking in a school meal program. Our paper provides evidence that monitoring[3] leads to significant improvements in program

[3] The monitoring in this experiment most closely resembles top-down monitoring, even though it was not linked to higher officials in the government in any obvious way.

implementation in the world's largest school feeding program, using experimental variation, and direct measures of meal quality. Our study is also unique in that we evaluate the efficacy of two concurrently government-run programs (the mid-day meals, and the IFA), using quasi-experimental variation in implementation, while combining these with experimental variation in MNM fortification and monitoring.

A fairly unexplored area in the broad literature on service delivery is that of crowd-out and negative spillovers in the implementation of concurrently run programs. The literature has, up to now, largely focused on evaluating ways to increase effort by service providers or public officials. However, the potential negative effect on the quality of service delivery in contexts where service providers may be overburdened has remained largely understudied. Our study is, to our knowledge, one of the first to measure the cost (to already existing programs), of adding on additional programs within existing public infrastructure. In much of the developing world, schools remain the primary avenue for government programs to reach children of school-going age. However, nutritional and other programs are being added on through schools, without matching increases in infrastructural and personnel capacities at these schools. We believe our results provide a starting point to explore such issues in program implementation.

Finally, our paper is also related to the literature on efficacy trials of iron-, as well as micronutrient-supplementation, and fortification. Studies in this literature have consistently found a decrease in iron deficiency, despite varying locations, populations, baseline rates of iron deficiency, intervention methods, and length of intervention. Estimates of the effect on anemia prevalence are highly responsive to the percentage of the population with anemia at the outset of the trial. Studies with only iron deficient participants are likely to largely overstate the effect of iron supplementation for a population of both anemic and non-anemic children. Effects of supplementation on micronutrient deficiencies are largest for children who are deficient in that micronutrient and/or anemic at the outset of the study (Abrams et al. 2008; Tee et al. 1999). Across a variety of efficacy trials, the effect of iron supplementation on iron content in the blood ranged from 0.95-1.8 g/dL for anemic children (Hirve et al. 2007; Ahmed et al. 2010; Tee et al. 1999; Gera et al. 2007; Abrams et al. 2008) and from 0-0.5 g/dL for non-anemic children (Hyder et al. 2007; Gera et al. 2007). Studies also investigate the effects of providing multiple micronutrients rather than just iron or iron and folic acid, since many children are often deficient in more than one micronutrient. In general, multiple micronutrients are equally or more effective when compared with just iron supplementation or iron supplementation with folic acid (Miranda et al. 2014; Ahmed et al. 2010; Ramakrishnan et al. 2004).

In summary, most of the efficacy trials done to date show that increased iron intake, with or without other micronutrients, is most often beneficial or does no harm. However, most of these studies have small samples, are short in duration, and include a substantial amount of researcher control over consumption of supplements and adherence to the program. While they do offer causal interpretations due to their randomized design, they tell us little about the implementation by governments or non-governmental organizations on a larger scale, in less controlled settings.

We proceed as follows: Section II provides context for the three interventions. Section III describes the interventions and the experimental design, including the sample selection and timeline. Section IV describes the data collected and Section V describes the results. Section VI addresses various challenges to internal validity and external validity. Section VII concludes.

II. School-Based Nutrition Interventions in India

A. The Mid-day Meal Program

In this section, we briefly describe two large-scale school-based nutrition programs in India: the Mid-day Meal program and the Weekly Iron and Folic Acid (IFA) program. The predecessor to the current Mid-day 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 (number of calories, e.g.).[4]

In Odisha, during the period of the study, the mid-day meal program was supervised by the Department of School and Mass Education. The implementation of the program was stipulated as follows: the cost of the meal (Rs. 3.30 per meal for each primary child and Rs. 4.92 per meal for each upper primary child) was shared by the Government of India and the state of Odisha (Government of Odisha 2011). The District Education Officers (DEOs) in each district are

[4] http://mdm.nic.in/aboutus.html

responsible for drawing and disbursing funds to schools for purchasing ingredients for the mid-day meals. In addition, the district, block and cluster level administrative officers are responsible for supervising the mid-day meal disbursement. The state government encouraged districts to delegate the preparation of meals to self-help groups (SHG), groups of local women (and sometimes men) who organize as a financial intermediary for such purposes as improving access to financial services or facilitating participating in the community. 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, SHGs were operating in only 43% of schools in our baseline sample. While the district is also supposed to train those responsible for providing the meals, in only 33% of 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 the schools in our baseline sample had received rice (99.7%) but only 42% of schools receive 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. The appendix provides the menu that the government prescribes in the state. Other ingredients such as dal, eggs, vegetables and fuel are purchased at the school level, using funds disbursed from the district government. Thus, since SHGs were not operating in most schools in our sample, either the headmaster, or one or more of the teachers is responsible for purchasing food materials and obtaining fuel for the meals, and hiring and supervising cooks. All of the teaching staff typically helps during lunch to organize the seating of students, distribution of meals, and washing of utensils before classes resume.

B. IFA Supplementation Program

In 2012, India's Ministry of Health and Family Welfare introduced the national iron supplementation program through schools to reduce the prevalence and severity of anemia among school children. Beginning in January 2013, according to the guidelines distributed by central government, iron and folic acid supplements, as well as deworming medication, were to be distributed free of charge to all students attending government and government-aided schools. Children 6-10 years old were to receive 30 mg of elemental iron and 250 μg of folic acid daily for 100 days out of a year, under supervision at school. They were also supposed to receive tablets to take home with them over school vacations. One tablet of deworming medication was also to be administered to each child every six months. The guidelines suggest that teachers conduct monthly nutrition and health education sessions with their students. Headmasters in each received the medication, and were expected to supervise the provision at school. Every month, government officials were supposed to monitor school compliance of both the IFA programs and mid-day meal programs.

While the central government mandated program involved weekly distribution of iron tablets, the program in Odisha was implemented differently during the first year. Tablets containing iron and folic acid were distributed to school officials in our study district of Keonjhar in 2013-2014 with instructions to give one tablet to each primary school child each day. In the second year, schools were instructed to distribute these tablets weekly instead of daily. In both years, upper primary school children were to be given a higher dosage (Ferrous Sulfate equivalent to elemental Iron 100mg, Folic acid-0.5 mg) daily in the first year and weekly in the second year. We surveyed schools and students about the implementation of the IFA program in the spring of 2014 and four times during the randomized control evaluation in the 2014-2015 school year. As described in detail below, we use this variation to estimate the impact of the IFA program in its first year after verifying that the variation appears to be quasi-random.

III. Experimental Design and Empirical Strategy

In this section, we describe the experimental design for the randomized controlled trial conducted during the 2014-2015 schoolyear, including the various treatment arms, as well as the empirical strategy for evaluating the IFA program in its first schoolyear, 2013-2014. We also describe the timeline, the study sample and the sample covered by each of our surveys.

A. Randomized Controlled Trial: MNM and Additional Monitoring

In the randomized component of our study, we evaluate two treatments to improve the implementation of school-based nutrition programs: the provision of nutrition inputs, specifically a multiple micronutrient mix, and increased monitoring of school meals.

Micronutrient Mix (MNM) Provision: In this treatment, we provide school headmasters and cooks with a multi-micronutrient mix, containing Vitamins A, D, C, B1, B2, B6, B12, Niacin, Zinc, Selenium and Calcium. We conducted rigorous trainings for 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 MNM and directions for MNM use. We also gave schools pictorial charts that clearly described the steps necessary to add the MNM to the food. Every month, we contacted schools to enquire whether they needed more of the MNM and, if so, delivered additional packets to the school.

Motivating this treatment is previous nutrition literature demonstrating that multi-micronutrient supplementation is more effective in combating anemia than just iron and folic acid supplementation (Ramakrishnan et al. 2004, Best et al. 2011). Ahmed et al. (2010), for example, showed that the provision of multiple micronutrients to anemic adolescent girls in Bangladesh increased hemoglobin levels more than supplementation with just iron and folic acid. Other studies have focused on expecting and lactating mothers, highlighting the health benefits of multi-micronutrient supplementation, over and beyond iron and folic acid supplementation, for mothers and infants (Roberfroid et al. 2008; Roberfroid et al. 2012). Note that the MNM does not include iron and 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. It is 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.

High Intensity Monitoring: In the second treatment, school meals were monitored earlier during the intervention and more frequently. 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. During each visit, enumerators observed 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 and measured the height of three randomly chosen students.

The two treatments were cross-randomized, so approximately half the schools in the MNM treatment arm were monitored intensively and half the schools who did not receive the MNM treatment were monitored intensively. The table below gives the number of schools and students in each group.

Table: Treatment Arms

	Treatment	Number of Schools	Number of students enrolled in these schools	Number of students surveyed at endline
MNM	Meal provider education and micronutrient mix provision (MNM treatment)	75	6969	989
Control	Continuation of current meals	73	6723	997
Monitoring	High Intensity Monitoring (Cross Randomization across other groups)	73 (37 from MNM) (36 from Control)	6432	969

B. Empirical Strategy to Evaluate IFA Program

Variation in IFA implementation across schools in the year before our intervention allows for the analysis of the effect of the IFA program using a difference-in-differences strategy, comparing the changes in hemoglobin levels for students who experienced the program and those who did not. Since the government's IFA program was supposed to be implemented in all schools beginning in the 2013-2014 school year, experimental variation is unavailable. Using observational variation would be problematic if it was correlated with differential trends in child outcomes in the absence of the IFA program. The main concern is that implementation is correlated either directly with trends in anemia prevalence or with trends in some other predictor that also affects hemoglobin levels. In this subsection, we describe the variation in IFA implementation and provide a number of checks to support the assumption of parallel trends.

Survey Data on the IFA program's implementation

In the spring of 2014, we conducted an uptake survey to better understand how the program was implemented during its first year and gauge any sources of variation we could use to evaluate the program. This school-level survey asked detailed questions about IFA receipt from the government and distribution of the supplements to students attending school.[5]

In the first year of the IFA program (the year before our intervention), however, there was variation in how well the program was implemented that allows us to evaluate its impact. Approximately 86% of the schools in our 377 school sample received the IFA tablets that year, but this hides substantial variation across blocks (i.e., administrative units smaller than districts); two of the blocks in our sample have very little variation in IFA implementation – 95% and 99% of schools received the tablets – while the other three blocks have substantial variation in implementation: 49%, 62% and 83% of schools in these three blocks received the tablets. We focus on these three blocks in the first year in our analysis of the impact of the IFA program, after verifying that the variation appears to be quasi-random, potentially due to when the block officials ran out of tablets. Within the schools that received the tablets, there is additional variation in whether or not the school received deworming medication and the number of tablets per student a school reports having received.[6]

In addition, as part of the school uptake survey, three children per school were randomly selected to answer several questions about the IFA implementation in their school. One student was randomly chosen from each of grade 2, grade 4 and grade 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. In the first year of IFA implementation these questions focused

[5] We repeated this survey four times during the intervention during the 2014-2015 schoolyear; the data from which we use to evaluate spillovers between the different nutrition programs. We found that almost all schools received the IFA tablets during the intervention; thus, there is little variation with which to evaluate the IFA distribution in its second year. Note that this also affects the interpretation of the MNM intervention: the comparison is now between children who received both iron supplementation and multi-micronutrient fortification and children who only received iron supplementation.

[6] In schools that did not report the number of tablets received, the measure was replaced by the number of tablets schools reported distributing. Schools reported receiving between 0 and 150 tablets per student and distributing 0 to 100 tablets per student. The correlation coefficient between these two measures is 0.5403.

on daily receipt and the previous day's receipt. In the second year of IFA implementation these questions focused on weekly receipt and receipt in the previous 7 days, since the policy had changed about how often IFA tablets were to be given to each child.[7] In blocks with variation in IFA implementation, there were substantial differences between school-reported measures and child-reported measures. In schools that reported receiving IFA tablets in the first year, only 58 percent of schools had at least 2 out of 3 children reporting daily distribution. Only 24 percent of schools reporting IFA receipt had at least 2 out of 3 children reporting that they received tablets the day before the survey. Table 1 provides summary statistics on these measures of IFA implementation by block.

Quasi-Random Variation in the IFA program's implementation

It seems natural to expect that variation in implementation across schools is correlated with trends in household demographic characteristics and levels of corruption, both of which might have an independent effect on child health (and, thereby, bias our strategy). While we do not have data to study trends prior to the beginning of the study, we can look at differences at baseline. We first show that, indeed, the variation in IFA implementation across all five blocks does match patterns of household demographics and school resource allocation implied by high levels of inefficiency. However, we also show that the pattern of distribution within block in the three blocks with high variation in IFA implementation, the variation we use, is not correlated with potential confounders; it appears to be quasi-random, driven perhaps by when tablets were received at the block level and when the block officials ran out of tablets.

Table 2 shows that the two blocks with over 95 percent IFA implementation are different from the three blocks with more variation in IFA implementation on a range of measures. More than half of the observable characteristics (measured at the school-level) differ, statistically, between the two types of blocks.[8] High implementation blocks are more advantaged across a range of demographic variables, have parents that are more involved in implementation of the school lunch program, and are more likely to receive rice for that meal on a regular schedule from the government (although this last difference is not statistically significant). At the same time, these blocks also have slightly higher anemia rates among children.

[7] Students were also asked if the swallowed the tablets they received. Almost all students responded that they had.
[8] Table 2 tests 37 observable demographics; 21 are significantly different at the 5 percent level and 4 more at the 10 percent level.

However, the main concern for our strategy is whether or not schools within the high-variation blocks received IFA tablets systematically. Within these three blocks, there are two possible explanations for why some schools report receiving tablets and others do not that could be particularly worrisome. First, this variation could have been non-randomly influenced by the block officials, if any unobservable characteristics are correlated with whether the block official gave the school the right number of tablets. For example, block officials could choose to focus on certain types of schools. Second, this variation could have been non-randomly influenced by the schools, if unobservable characteristics are correlated with (a) how schools implement the IFA program or (b) how schools report receiving tablets in their response to the IFA survey. Our preferred measure of IFA implementation (whether or not the school received the tablets from the block official) helps minimize omitted variable bias from (a) since receipt of tablets does not rely on a school's ability to implement a program, but school characteristics may still be correlated with how schools respond to the question about IFA receipt on the IFA survey. However, further analysis supports none of these sources of bias (block- or school-induced). Finally, these systematic distribution patterns would only introduce bias if they were correlated with hemoglobin level trends in children. In addition to showing that there is little evidence to support any systematic distribution patterns, we also show that IFA receipt is not predicted by students' anemia status or hemoglobin levels and that observable characteristics of schools are not correlated with the percent of students that are anemic in each school.

To test these hypotheses for non-random tablet distribution we first check for differences between schools that received tablets and schools that did not, and then examine the ability of these measures to predict IFA receipt. To test the hypotheses for block-induced non-random distribution, we use the school's distance to the block headquarters and a range of demographic measures about each school.[9] Given data constraints it is impossible to fully untangle whether a block official targeted schools with a high ability to implement a government program or whether headmasters that are better implementers were better at reporting tablet receipt – we simply observe the ability of a school to implement a government program through their success at implementing the MDM. Similar to IFA distribution, supplies for the MDM are distributed by

[9] These measures include school-reported proxies for socioeconomic status, e.g. whether or not the school has a kitchen or sufficient water, and a range of household-reported proxies for socioeconomic status aggregated to the school level, e.g. the percent of families in agricultural work, the percent of families who own a phone, or the percent of families living in high- or low-quality housing.

block officials to each school. For each school, we observe the (parent-reported) mean number of lunches provided per week and the percent of parents who are satisfied with the implementation of the MDM. Further, we observe whether or not a school uses a self-help group to provide the MDM, whether or not anyone from the school attended government MDM training, and whether or not a school gets regular scheduled visits from the block officials to deliver the rice for the MDM. This final measure is the only one that contains information about block level decision-making; the rest simply measure the school's effectiveness at implementing the MDM.

As seen in Table 3, there is no significant difference in mean distance to the block headquarters among schools that got IFA tablets and those that did not within the three blocks with high variation in IFA implementation. Furthermore, over the range of observed demographics and measures of IFA implementation, there are few observable differences between schools that received IFA tablets and schools that did not. Out of all the observable demographic variables that the block officials likely knew, there are only three significant differences between these two types of schools: the percent of the population in a disadvantaged caste, the percent of villagers who report working in their own home not for pay, and the percent of villagers who report working in others' homes for pay. In addition, there are no significant differences between schools that received tablets and those that did not related to a school's ability to implement the MDM.

Furthermore, none of these variables overall are predictive of IFA receipt in a regression of IFA receipt on varying sets of demographic and school variables. Columns (1) and (2) of Table 4 regress IFA receipt on distance to the block headquarters and a host of observable demographic characteristics and MDM implementation variables (with and without block fixed effects). None of these observable characteristics significantly predict IFA receipt either across or within blocks in the three blocks with high variation in IFA implementation. However, this result may be caused by the smaller sample size due to missing data (n=124 schools). To account for this, the remaining columns consider subsets of the variables included in the first two columns. There are only two robustly significant predictors: the percent of villagers employed in housework outside the home and the percent of families in a non-disadvantaged caste. Given the number of variables tested, this is approximately the number we would expect to see significant by chance (at the 10% level). Overall, these regressions suggest that the block officials did not systematically target schools based on their observation of differences between schools or village populations.

If the tablets were disproportionately given to more disadvantaged schools within these three blocks (as mildly suggested in Tables 3 and 4), the estimated effects of the IFA could be biased in either direction, depending on how the trend in hemoglobin levels would have differed for advantaged and disadvantaged children in the absence of the IFA. However, in this sample, anemia and poverty are not strongly correlated, which supports the assumption of similar trends. Furthermore, controlling for whether a student is advantaged or disadvantaged only slightly changes the point estimate of the effect of the IFA (in fact, it increases slightly).

Finally, and perhaps most importantly, there is no evidence that schools receiving tablets had students that were disproportionately more or less anemic. Overall within the three blocks with high variation in implementation rates, as well as within each block separately, there is no statistical difference in the prevalence of anemia, mild anemia, or moderate anemia between schools that received the IFA tablets from the government and schools that did not (Table 3, Panel C). Additionally, there is no difference in the mean hemoglobin level or in standard nutritional markers like weight and height. Figure 1 plots the kernel densities of students' hemoglobin levels in schools that did and did not get IFA tablets and shows that the distribution of hemoglobin levels at baseline among study children is quite similar in both types of schools, for both anemic and non-anemic children.

More generally, anemia rates in each village are not correlated with any observable demographic characteristic (Figure 2): without testing hemoglobin levels, a block official could not target students who needed the iron supplements more even if he wanted to. This is consistent with the literature that shows that in settings with widespread anemia and poverty, it is difficult to identify those most in need of iron supplementation without actually measuring iron deficiency (WHO 2015). Further, in this sample, only 11% of parents know what the health condition called 'anemia' is (after implementation of the IFA). It is likely that even fewer adults are aware of the use of iron supplements to treat the micronutrient deficiency, and thus that there is no market for iron supplements, even if an official wanted to sell them. Overall, this suggests that the block official would distribute all of the provided tablets to schools and that he would do so in a way unrelated to underlying trends in children's hemoglobin levels.

In addition, school headmasters reported whether or not they had already run out of tablets (ran out, did not run out, or uncertain) at the time of the school uptake survey. In the sample of schools in blocks with high variation in IFA implementation that did receive tablets (111 schools), 39 schools report running out of tablets, 25 report still having tablets to distribute, and the remaining 47 are uncertain. Schools that report having already run out of tablets at the time of the school survey appear very similar to schools that do not run out on the same range of characteristics described above, suggesting that both the timing of tablet distribution and the number of tablets provided per student are also likely not systematically determined by the block officials and rather, were largely determined by chance.[10] These facts together with the descriptive analysis above, support the quasi-random distribution of tablets within each block.

C. Timeline

Figure 3 gives the chronology of key activities in the study. Figure 3 shows that the first baseline survey (Baseline 1) was completed by January 2013. The original design of this study was to provide iron-fortified meals in school, but the distribution of iron-fortified ingredients was halted when the Government's IFA program was announced during the 2012-2013 schoolyear. 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 the final approval, we conducted the Spring 2014 IFA implementation survey, described above as well as a second baseline survey (Baseline 2) in a subset of sample schools in order to update anthropometric and hemoglobin measures from the first survey. These two sources of data allow us to evaluate the impact of the government's IFA program on child health.

The MNM and high intensity monitoring evaluation was launched at the end of November 2014, and continued through April 2015. 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 launched in April 2015, and was completed in early July 2015. At the request of the

[10] Two differences are that schools that ran out of tablets for primary school children have much larger average secondary school enrollment (and perhaps they redistributed tablets designated for primary students to secondary students) and are also much less likely to receive their rice for the school lunch program on a regular schedule (indicating less frequent contact with the block officials). See Appendix Table A1 for the full range of statistics.

government agency in charge of school meals, focus group discussions were conducted with school children, their parents, and school officials from six sample schools in July 2015.

D. Sample

Schools

The sample schools in Keonjhar district were selected for the study based on whether they satisfied the following conditions: (i) the school is located within 50kms from the town, which was the capital of the district and (ii) the school is located in one of five blocks (i.e., administrative units smaller than districts): Banspal, Ghatagaon, Jhumpura, Keonjhar Sadar, 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 schools that satisfied these conditions, and the final sample for the study consisted of 150 randomly chosen schools from this list.[11] The sample of schools included in the evaluation of the government's IFA program, on the other hand, includes all 157 schools from the original list of 377 in the three more remote blocks, the blocks with variation in IFA implementation.

Out of the 150 study schools, 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. We stratified on block because block officials play a significant role in the mid-day meal program: funds and rice to schools for the mid-day meals are channeled through block offices, and schools are accountable to officials at the block level. We also stratified by school type because schools that have classes 6, 7 or 8 (in addition to classes 1 to 5) are accountable directly to officials at the district headquarters as well as block officials.

While the original list 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.

[11] The original study also intended to evaluate centralized school meal delivery operated by Naandi Foundation as well. Due to the various delays the project faced, Naandi ultimately decided not to participate in the study or provide meals to the study schools.

Children

All students in these schools received the treatments (conditional on attending), but conducting household and anthropometric surveys for all children would have been prohibitively expensive (specifically, getting parental permission). Thus, we randomly chose 15 students in each school to survey based on the power calculations described below. These students were chosen from the set of students enrolled in sample schools in classes 1 to 5 who live with their parents (and not in school hostels). We excluded children living in hostels due to the difficulty in locating their parents to obtain permissions. Students were randomly chosen, after stratifying by school and class. Our sample is most likely representative of children in Keonjhar schools that have some access to the town who live with their parents during the school year. Prevalence of malnutrition (underweight) and anemia among children under the age of five in Odisha is fairly representative of India, more generally: 41% of children in Odisha are underweight vs. 43% of children in India and 65% of children in Odisha are anemic (mild, moderate or severe) vs. 70% of children in India. While the NFHS-3 does not report hemoglobin levels for the school age population, data from our baseline survey indicates that approximately 60% of the children in our sample are anemic, a number very similar to the national average. Similarly, 44% of children in our sample are underweight.

Since we conducted multiple baseline surveys, and since we also conducted the surveys across school years, it is important to note here the samples surveyed during each survey round, as well as potential attrition across school years. First, the sample of children we surveyed during our second baseline survey (Baseline 2) was about 50% of our original sample, due to budgetary restrictions. Second, children who were in grade 5 during the first baseline survey (Baseline 1) had graduated from primary school before the 2013-2014 school-year, the first year of the IFA program, and the 2014-2015 school-year, the year of our intervention. Thus, at Baseline 2, we only surveyed children who had been enrolled in classes 1-4 during Baseline 1. We also sampled another 3 students who were enrolled in classes 1 and 2 during the 2014-2015 schoolyear. Thus, instead of 15 students per school there are on average 8 children per school for Baseline 2. Recall that at Baseline 2, we only surveyed children in the 3 administrative blocks with variation in the IFA implementation. At endline, we surveyed those enrolled in classes 1-3 during Baseline 1 and the students added to the sample during Baseline 2. We also sampled additional students enrolled in classes 1 and 2 during the 2014-2015 schoolyear to get a sample of 3 students per class. With some attrition, there are on average 14 students per school at endline.

Focus Groups

Finally, we also conducted focus group discussions in 6 randomly picked schools— two were MNM treatment schools with high MNM take-up, two were MNM treatment schools with low MNM take-up and two were comparison schools, in order to get focus group responses from a variety of perspectives and experiences.

IV. Data and Balance Checks

We collected data on a number of outcome variables at various points during the study. Specifically, we collected extensive data on i) school infrastructure in meal provision, ii) the implementation of the IFA program, iii) the quality of mid-day meals and the take-up of the micronutrient mix (including the quantity of vitamin A and zinc in food samples), iv) child-level outcomes (including hemoglobin levels, anthropometric measures, cognition, school attendance and test scores) and v) household-level demographic characteristics and information on assets and mid-day meal perceptions. We briefly describe each survey and the variables of interest below.

School survey data: Data on school characteristics and teacher demographic details and qualifications were collected during Baseline 1 and once again during the second month of the intervention.

Take-up rates for MNM: Take-up of fortification schemes is an important outcome in itself. One measure of take-up is the amount of MNM each school received, relative to the amount we calculated they would need to serve all their students. We subtract out the amount of the MNM that remained at the end of the school year in April 2015.

Meal quality: Trained enumerators made surprise visits to the study schools to observe the quantity and quality of school meals. 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.

Testing for micronutrients in meals: During the third and fifth months of the intervention, enumerators took 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 sample.

Implementation of the IFA Program: During April-May 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.

School attendance: Each month, an enumerator made random, unannounced visits to each school to take attendance. These checks were made at random times of the day in case children attend school just for the meal and leave immediately after.

Child health: Enumerators visited the schools and households to measure the selected children's height, weight and hemoglobin levels during the Baseline 1, 2 and Endline surveys.

Test scores and cognitive tests: During the Baseline 1 and Endline surveys, students in grades 1 to 4 were given mathematics and reading tests designed by Pratham, an India-wide NGO that works on child literacy. Enumerators also conducted tests of cognitive development. We used two cognitive development tests: 1) 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 2) a Block Tapping Test (Kar et al., 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 or endline).

Household data and health of family members*:* During Baseline 1, enumerators visited the households of the selected children. The survey conducted at this time included a module on demographic information, school attendance, assets, anemia and perceptions of the school's mid-day meal. The enumerators also conducted cognitive tests on household members, and were accompanied by specially trained enumerators, who measured the height, weight and hemoglobin level of younger siblings of the selected children (siblings in the 3-5 age range), all female siblings and their female guardian. A similar survey was administered at Endline, the main difference being that only the height, weight and hemoglobin levels of the selected schoolchild and his or her younger siblings were measured.

A. Balance in Randomized Controlled Trial

Table 5 checks balance on household characteristics and child health at baseline, across each of the treatment groups in the randomized controlled trial of the MNM and monitoring interventions. 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 all three differences. As shown in Table 5, the groups are well balanced on the child health outcomes of interest in Panel A, with a slight imbalance on a few of the 32 variables for which balance is checked. 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 across villages. We also present a similar balance table on school characteristics that were measured during our baseline school survey (Table 6). Our sample is well balanced on the variables measured at baseline.

V. Results

This section first describes the results on the impact of the MNM and high intensity interventions on various outcomes of interest, including take-up, meal quality and child health. Next, we discuss spillovers between the various interventions. Finally, we present our results from the evaluation of the government's IFA program.

A. MNM Intervention and High Intensity Monitoring

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* as y_s, the basic specification in our analysis is as follows:

$$y_s = \beta_0 + \beta_1 mnm_s + \beta_2 high_s + \delta controls_s + v_s \qquad (1)$$

In order to account for any differential impact that high intensity monitoring may have had on the MNM treatment, we also include a specification that includes an interaction term

$$y_s = \beta_0 + \beta_1 mnm_s + \beta_2 high_s + \beta_3 mnm_s * high_s + \delta controls_s + v_s \qquad (2)$$

where mnm_s is a dummy variable for the group that received the MNM fortification treatment, and $high_s$ is a dummy variable for the group that received higher frequency monitoring visits. All our regressions contain fixed effects for administrative block. Table-specific controls are indicated below. For outcomes that were measured at the child level, or whenever we make use of multiple observations within a school, standard errors are clustered at the school level. For some specifications, we also control for whether the school received IFA tablets during the previous school-year to see if familiarity with nutrition supplements matters for implementation.

We consider two 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 mid-day meal monitoring visits, such as whether the enumerator noticed a powdery addition to the meal.

The estimates on take-up measured from our delivery records are reported in Table 7. We exclude schools not in the MNM treatment since they did not receive any of the mix. In addition to block fixed effects, these regressions control for the number of children enrolled in the school as of the start of the intervention. Schools assigned to the MNM treatment did take up the mix. The schools that were not monitored intensely received 2.9 deliveries of the mix during the study period (the constant term in Columns 1-2), received approximately 0.6 kg of the mix per child enrolled in the school and used almost all of it. The high intensity monitoring did not affect these measures of take-up.

Table 8 further reports take-up as inferred during the MDM 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 it directly on inspection of the container in which the meal was cooked (Columns 1-3), (ii) the likelihood that the cook 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). High intensity monitoring did not affect take-up.

Note that the coefficient on whether or not a school received IFA tablets in the previous year is often negative and marginally significant (at 15% in Table 7 for the amount of MNM delivered to the school and at 10% in Table 8 for whether the MNM could be located in the storeroom). We come back to possible spillovers between the two nutrition programs in Section VI.E.

Effects on Micronutrients in School Meals

One of the main outcomes of interest was the nutritional quality of meals being served at schools. As described in previous sections, we measured micronutrients present in meals by collecting food samples at school during meal times, and tested these samples at a laboratory for zinc and vitamin A. These measures could also be considered indicative of take-up, with the caveat that meals can contain vitamin A and zinc even if they do not contain the mix.

Table 9 presents the results of the effect of the interventions on the micronutrients present in the meals. Both zinc and vitamin A levels increase significantly for schools in the MNM treatment, and we see no detectable differential effects for those schools that were also monitored at a higher frequency or those that received IFA tablets during the previous school year.

Effects on Child Health

The next set of results relates to the health outcomes of the children in our sample schools, our main outcome variable. Our measures of child health are (i) hemoglobin levels, as well as several anthropometric outcomes: (ii) weight, (iii) height, (iv) weight for age z-score, (v) height for age z-score, and (vi) mid-upper arm circumference.[12]

Tables 10 and 11 present results of the treatment effects on child health outcomes using simple difference, and difference in differences (DD) models, respectively. All of the specifications include block and age dummies, and Columns 2 and 3 include a lagged dependent variable from Baseline 1 and 2, respectively. Column 4 includes the lagged dependent variables from both Baseline 1and 2, and to allow for the inclusion of children included in the sample only at endline,

[12] In results available upon request, we replicate Tables 10 and 11 for outcomes measuring school attendance, cognitive ability and proficiency in reading and mathematics. Neither intervention has statistically significant effects on these outcomes. This is not particularly surprising given the lack of an effect on child health for the MNM intervention.

this also includes dummies for missing observations. Recall that some children were included in the sample only at endline because they had not been enrolled in school during the Baseline 1 survey, two years prior to the intervention.

The results in both Tables 10 and 11 indicate that the MNM treatment had no effect on hemoglobin, height or mid-upper arm circumference; in fact the coefficients are negative in the simple difference results for hemoglobin. While Table 10 does suggest that there is a decline in height and mid-upper-arm circumference due to MNM treatment, this finding is not robust to many specifications, particularly the difference in differences specification as seen in Table 11. All of the results are also robust to including controls for (i) whether the school received IFA tablets in the previous year, (ii) the fraction of children surveyed (out of three) during the first IFA survey that report receiving IFA tablets daily, and (iii) the fraction of children surveyed (out of three) during the first IFA survey that report having received IFA tablets on the previous day.

At the same time, both Tables 10 and 11 reveal robust, positive effects of high intensity monitoring on hemoglobin levels, although not on the other anthropometric outcomes. In the next section we discuss interactions between the three interventions that may help explain these findings.

B. Interactions between MNM, high intensity monitoring and IFA program

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 overlapped. It is easy to imagine complementary effects if high intensity monitoring gives 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 introduced in the treatment schools, on top of the existing mid-day meals program as well as the IFA program. Since the adding on of one additional program at the school level increases the workload of the school staff, it is plausible that this might lead to negative effects on how well other programs are implemented at school.

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 – 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 (Tables 8 and 9). The results in Tables 10 and 11 indicate no difference in the effect of high intensity monitoring in schools that also received the MNM mix and those that did not.

A second explanation is that the high intensity monitoring may have led schools to implement the IFA program better. Table 12 shows treatment effects on measures of how well the IFA program was implemented. We focus on four measures of IFA implementation quality: (i) whether the headmaster shows the enumerator an IFA tablet, (ii) the number of tablets distributed per child in the past week (as seen in the school report), (iii) the percent of students who say they get the tablets weekly or more frequently (out of three randomly chosen students spanning different grades), and (iv) whether at least 2 out of 3 students asked say they get the tablets at least weekly. High intensity monitoring has a positive effect on the implementation of the IFA program. Students in these schools are more likely to report getting the IFA tablets regularly. Interestingly these results are driven only by student-reported outcomes, which would be more difficult for the principal to manipulate (since the children were randomly chosen each month). In addition, these results are driven by responses later in the 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 show up – most high intensity schools had received at least 2 and sometimes 3 mid-day meal visits while low intensity monitoring schools had received at most 1 visit.

At the same time, schools receiving the MNM treatment seem to do worse on these IFA implementation outcomes. Students are less likely to report having received the IFA tablets regularly and the headmaster is less likely to be able to produce the IFA tablets to show the enumerator. Thus, these results suggest that there is some crowding out of IFA implementation by the introduction of the MNM. As discussed above, there is suggestive evidence of crowd-out in the other direction in Tables 7 and 8: receiving IFA tablets during the previous year may reduce take-up of the MNM, although the coefficients tend to be significant only at 10% or 15%.

Finally, we explore whether high intensity monitoring or the MNM introduction in schools affected the quality of the school meals (Table 13). We consider several measures of meal quality based on variables measured in our MDM monitoring surveys. These include: (i) whether a meal was served, (ii) whether vegetables had been added to the meal, (iii) whether any children received 2nd helpings of rice, or of the non-rice dish (egg curry, or dalma, or soybean curry), and (iii) the number of adults served. While there are significant impacts on whether a meal was served or vegetables were added, we refrain from concluding too much from these variables given the lack of variation – in 100% of visits to control group schools a meal was served and vegetables were added. While we see some effects on the number of adults who are served, we see no corresponding impact on the amount of food children receive.

C. Effects of the IFA Program

The results above related to the IFA program rely on variation in the MNM and monitoring treatments and do not enable us to evaluate the impact of the government's program itself. Recall that there was little variation in our measure of IFA implementation during the intervention. Thus, we use data from the first year's survey in April-May 2014 to evaluate the IFA program.

Empirical specification

As described above, the core of our strategy to identify the impact of the IFA program is a straightforward difference-in-differences (DD) model, comparing the change in hemoglobin levels for children who experienced the program relative to students who did not. This specification takes the form:

$$\text{Hb}_{ist} = \beta_0 + \beta_1 \text{IFA}_s + \beta_2 \text{post}_t + \beta_3 (\text{IFA}_s \text{ x post}_t) + \varepsilon_{ist} \quad (3)$$

where Hb_{ist} is the hemoglobin level of child i in school s at time t, IFA_s is a marker of IFA implementation, and post_t is an indicator for whether hemoglobin measurement was taken after IFA implementation. Additional control variables include the distance from a school to block headquarters, whether or not a school has a kitchen, the percent of parents satisfied with implementation of the school lunch program, the percent of families per school employed in housework outside the home, and the percent of families per school in a non-disadvantaged caste.[13] In the preferred specification with school fixed-effects, these control variables are interacted with the post indicator. Additional specifications include an indicator for whether or

[13] These control variables are intended to proxy for the different decision-making processes that could influence block officials, such as distance to the block headquarters, a school's implementing ability, village demographic indicators of socio-economic status, and a school demographic indicator of socio-economic status. The inclusion of these control variables does not substantially alter the coefficients, further supporting the conclusion that the manner in which the block official distributed tablets is not correlated with other differential trends in anemia.

not a school received deworming medication from the government, an interaction of that indicator with post, and an interaction to capture the joint effect of IFA receipt and deworming receipt. As discussed above, in order to infer that β_3 is the causal effect of the IFA, we assume that the health indicators of students in both IFA and non-IFA schools would have been on the same trend in the absence of the program.

Next, we estimate heterogeneous effects by comparing the difference in β_3 when Equation 3 is estimated separately for students at different points in the distribution of hemoglobin levels at baseline. The results of this estimation further support the validity of the identifying assumptions, since any differential trends across schools correlated with IFA implementation would also have to differ by baseline hemoglobin level.

Impact of IFA program

Table 14 presents results from specification (3), the DD analysis estimating the effectiveness of the IFA program in raising student hemoglobin levels. The dependent variable is a child's hemoglobin level and the key independent variable is an indicator for whether or not the school reported receiving IFA tablets from the government.[14] For this and all subsequent tables, the even-numbered columns include school fixed-effects and the additional control variables interacted with the "post" indicator. Columns 3-6 additionally control for the receipt of deworming medication from the government. The point estimates indicate that attending a school that reported receiving IFA tablets increases children's hemoglobin levels by 0.280-0.307 g/dL, once we control for receipt of deworming medication, fixed effects and the list of control variables described above. While not statistically significant at conventional levels, the p-values are suggestively close (0.118 and 0.130) for these two estimates. This effect is of the expected magnitude for combined anemic and non-anemic students in a real-world iron supplementation program.[15] Unexpectedly, there is a negative and significant effect of attending a school that

[14] Results are qualitatively similar when the independent variable is measure of the number of tablets received per student. This measure is noisy due to inconsistent reporting on the part of schools and therefore not our preferred measure.

[15] In the study most similar to this one in both supplementation program and empirical design, Luo et al. (2012) find the overall effect of school-based iron supplementation to be 0.23 g/dL for 4th graders in rural China.

received deworming medication, but no differential effect for schools that received both IFA and deworming medication (Columns 5 and 6).[16, 17]

Next, we estimate heterogeneous effects of the IFA program with respect to initial anemia status that are consistent with a causal interpretation of the DD specification. This supports the identification strategy because if the DD results were driven by differential trends instead of the IFA, we would have no reason to expect the results to be bigger for anemic students than non-anemic students. We divide the sample by baseline anemia status: non-anemic students (hemoglobin concentration over 12.5 g/dL at baseline), borderline anemic students (hemoglobin concentration between 11.5 and 12.5 g/dL at baseline), mildly anemic students (hemoglobin concentration between 11 and 11.5 g/dL at baseline), and moderately anemic students (hemoglobin concentration between 8 and 11 g/dL at baseline).[18] Note that "mild" anemia is a misnomer in that the negative effects of iron deficiency are already substantial by the time any level of anemia is diagnosed (WHO 2011). Similarly, borderline-anemic students are likely to be suffering from many of the negative effects of iron deficiency as well. The majority of the children in this sample are mildly or moderately anemic. Informed by previous highly-monitored trials of iron supplementation, we expect the effect of the IFA to be largest for more anemic students and smallest for the non-anemic students.

Table 15 presents the results from this heterogeneous effects model and illustrates that this expectation largely holds, providing additional support for the identification strategy. Focusing on the estimates that include school fixed effects, the IFA has an insignificant effect on the students with the highest baseline hemoglobin levels that fluctuates in sign between models. The effect is larger and positive across all specifications for non-anemic borderline-anemic students (0.09-0.37 g/dL with controls), but still insignificant. The largest and only statistically significant

[16] The negative effect of attending a school that received deworming medication could be a consequence of selection as well. As seen in Appendix Table A2, any differences in schools that got deworming medication indicate that those schools were more advantaged, even within these three blocks, and potentially on different trends.
[17] Appendix Table A3 presents results for this DD analysis with height and weight as the outcome variables. Existing literature shows no effect of iron supplementation on height and mixed, inconclusive effects on weight (Low et al. 2013; Vucic et al. 2013). Table A5 shows that the IFA had no effect on height and a small but significant effect on weight. In this context, iron supplementation could increase weight by reducing lethargy and increasing school attendance, thereby increasing weight if students receive more nutritional school lunches. Unfortunately, we do not have data on children's attendance in the first year.
[18] These hemoglobin cutoffs are as defined by WHO standards at sea level and apply to the majority of the sample (5-11 year olds). Students outside this age range are classified by alternate age-appropriate cutoffs.

effect of the IFA occurs for mildly anemic students: the IFA causes a significant increase in hemoglobin levels of 0.49-0.84 g/dL with the inclusion of school fixed effects and control variables including deworming receipt. This effect is about twice as large as the overall effect for all students reported in Table 14.

Finally, the effect of the IFA for moderately anemic students ranges from 0.218-0.248 g/dL with the inclusion of controls for deworming receipt and is insignificant. While the finding that the effect on moderately anemic students is smaller than the effect on mildly anemic students seems surprising, there are at least three possible related explanations. First, the most anemic students may not have received enough iron through the IFA to build up sufficient iron stores, for example, because of more infrequent school attendance due to the negative effects of anemia (such as like increased lethargy). Second, note that moderately anemic students are the only subgroup to have a positive (but insignificant) point estimate of the interaction effect of iron supplementation and deworming. This indicates that the most anemic students may have also been those with the highest worm loads. Both of these hypotheses imply that these children would therefore have experienced smaller immediate effects of iron supplementation as well as the most dramatic falls in hemoglobin levels when they ceased receiving iron supplements. On the other hand, students who were mildly anemic at baseline are likely less susceptible to timing discrepancies if they were more able to build up sufficient iron stores over the course of their supplementation. Since hemoglobin measurement was done over the summer vacation, it is possible the measureable effect for mildly anemic students persisted while the effect for moderately anemic students did not. A third explanation is that the most anemic students are likely to be more severely deficient in other vitamins and micronutrients, perhaps affecting their ability to absorb the iron in the supplements.

These results imply that school-based iron supplementation programs may not be sufficient to reduce the most severe cases of anemia (affecting one-third of children in this sample), but may be most effective in improving the hemoglobin levels of borderline or mildly anemic students and therefore preventing them from developing more severe levels of anemia. A main disadvantage of using the school system to distribute tablets is that the program only reaches kids who attend school frequently. These results are likely to generalize to other school-based nutrition programs, which would face many of the same constraints.

In our third set of results, we examine the heterogeneous effects for students whose schools report running out of tablets ahead of the school survey. Recall that schools reported whether they had run out of tablets, whether they still had tablets or whether they did not know if they had run out approximately two to four months before hemoglobin was measured. It is likely the impact of supplementation had fallen for children in schools that had run out of tablets. Table 16 presents the results, which support the hypothesis that students with more recent iron supplementation are driving the measurable effect of the IFA described above. In schools that still had tablets to distribute at the time of the school survey (the omitted category), the IFA increased children's hemoglobin levels by 0.414 g/dL (p-value 0.106). Students in schools that reported uncertainty regarding whether or not they had run out of tablets experienced an IFA effect of similar magnitude. However, the effect for students in schools that ran out of tablets at least two to four months before hemoglobin measurement was smaller by 0.311 g/dL (p-value 0.116), potentially because any effect they may have experienced eroded by the time of hemoglobin measurement. Note that the p-value from the F-test of all three IFA interactions presented at the bottom of the table rejects the null hypothesis that all three coefficients are zero at more conventional significance levels.

VI. Threats to Validity and Policy Implications

In this section, we discuss potential concerns with respect to internal and external validity, and present a discussion of implications of our results for policy.

A. Spillovers between Treatment Groups

First, we consider contamination of the MNM provision for schools that were not meant to receive the mix. While we did not deliver the mix to any schools in the comparison group, it is possible that they obtained a similar mix to add to their meal, perhaps inspired by the information on anemia provided to all schools at the onset of the intervention. We find very little evidence that this occurred, based on the food test results and the mid-day meal observations. Vitamin A content in the meals taken from schools with neither intervention were 52 µg/100g of food in February, and barely changed by April (55 µg/100g). Schools that received the MNM provided food with levels of vitamin A that were 7 times the food provided in schools that did not receive the MNM. Considering the observation of mid-day meal provision, the cook reported adding a powder to the meal only twice out of more than 400 visits to schools that did not receive the MNM. Similarly, the enumerator noted a powdery addition in the meal only twice out of more than 400 observations. None of the schools are repeated among these four observations suggesting that these are due to human or measurement error and not due to contamination. Note, in addition, that high intensity schools were not more likely to request and receive more of the mix (Table 7) and not more likely to serve food with higher vitamin A or zinc content (Table 9) – regardless of whether they received the mix or not – suggesting that contamination was unlikely.

B. Attrition

Another concern is overall as well as differential attrition (across treatment groups) of our sample. Table 17 provides an analysis of attrition for our main outcome variables on child health. This analysis is complicated by the fact that more than 2 years had passed between our baseline 1 survey and the endline survey and the fact that we only surveyed a subset of the respondents at the time of the baseline 2 survey. We focus on children who were surveyed at baseline 1 and still in primary school during the year of the intervention (those in classes 1-3 at baseline) in Columns 1-6 and additionally include children in the MNM intervention sample who were surveyed at baseline II in Columns 7-12. 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 minimal at only 9%. Column 1 in Table 17 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, class at baseline does not affect probability of attrition, and this does not differ by treatment group (Column 5-6, 11-12).

C. External Validity and Policy Implications

The goal of this research was to study iron supplementation or fortification ‘in the field.’ While efficacy trials have convincingly demonstrated that iron supplementation and fortification can improve child health and school attendance, these studies are often highly controlled with compliance rates above 95% because researchers closely monitor the delivery and consumption of iron supplements. This study, on the other hand, was able to evaluate the efficacy of a program that distributed iron through existing infrastructure, specifically the Indian mid-day meal program, in addition to studying a potential mechanism for the distribution of a micronutrient mix.

We were able to evaluate three policy-relevant interventions: 1) the government's as-is IFA program, 2) the provision of MNM to randomly chosen schools and 3) high intensity monitoring of mid-day meals, each with related but separate policy implications. We start with some general considerations with respect to external validity and policy implications and then discuss each intervention in separate subsections.

A number of elements about 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.[19] 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 provides some insight in this area: variation in the receipt of IFA tablets during the first year is likely due to incorrect estimates of the number of tablets each block needed because the blocks ran out of tablets. 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. Taking intensity of meal monitoring to scale would require addressing the issues that currently limit frequent monitoring.

[19] MNM provision through mid-day 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. Our intervention with the MNM mix on the other hand, was *decentralized* since the mix was distributed through the existing mid-day meal infrastructure in each school.

Another important 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%) was a meal not served. This bodes well for a possible effect of fortification, but may make it less easy to generalize the effects to a setting where school meals themselves might be inconsistent.

Policy implications from the IFA evaluation

The evaluation of the IFA program suggests that a school-based iron supplementation program has the potential to improve hemoglobin levels and reduce anemia prevalence for school-aged children in districts similar to Keonjhar District. A number of factors may have jointly contributed to this result. First, implementation of the program, while not perfect, did not appear to be plagued by systematic distribution by corrupt officials within each block. The main barrier to total coverage of the program seemed to be the misallocation of tablets relative to the number of actual intended recipients. In the second year of the program, 100% of the schools in our sample reported receiving tablets from the government. Of the schools that had received tablets in the first year, more than half of them received more tablets in the second year than they did in the year before. Overall, these data suggest that the administrative wrinkles were quickly and effectively ironed out of the IFA program, and that it therefore could, in principle, have a substantial impact on the prevalence of iron deficiency and anemia in its second (and subsequent) year(s).

However, there are still limitations facing programs like the IFA. The policy had larger measurable effects for students who received tablets closer to the time that their hemoglobin levels were measured. This is not unexpected, given the life cycle of a red blood cell and the low levels of iron naturally present in most Indian diets. This suggests that in the intervening time between rounds of supplementation in school, children's hemoglobin levels may fall. This could occur whenever schools run out of tablets or more systematically when students are out of school for long periods of time (e.g. the summer holiday). There are two obvious solutions, although they may be difficult to implement: first, ensuring that schools receive enough tablets, and second, providing students with tablets to take home over school vacations. While out-of-school tablet provision and student compliance may work differently than in-school provision and compliance, students would be less likely to experience falls in hemoglobin levels over the summer months.

Recall that the largest effects of the IFA program were concentrated among borderline anemic (i.e. iron deficient) children, suggesting that it could be particularly effective in reducing iron deficiency among children who are not yet presenting visible signs of moderate or severe anemia. At the same time, the IFA program was less effective in improving the hemoglobin levels of moderately anemic students. It is possible that the IFA program is not intensive enough to fully treat students who already present high degrees of iron deficiency, or that it does not reach those students as effectively because they are less likely to regularly attend school.

One persistent puzzle in the results is the potential negative effect of deworming on hemoglobin levels and the insignificant interaction between deworming and iron supplementation. While data constraints limit the further evaluation of these effects in this study, the phenomenon should be further studied in real-world programs that implement both biannual deworming regimens and weekly or daily iron supplementation in schools.

Policy implications from the MNM distribution

The evaluation of the MNM distribution has several policy implications. First, note that take-up was relatively high: according to our records, only 3 schools out of 75 did not use any of the micronutrient mix. On average, schools used more than 58% of the amount we estimated they would use based on the number of students enrolled; ninety percent of the schools used at least 40% of the amount we estimated they would need. One contribution of this study is the evaluation of nutrition programs run by different entities. The range of take-up measures is similar across both the nutrition program run by the government and the program run by researchers. For example, in 72% of mid-day meal visits, the cook reported adding a powder to the meal, while 62% of children interviewed reported receiving the IFA tablet regularly and 86% of schools received IFA tablets. In conjunction with the evaluation of the IFA program, 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 just iron supplementation. One likely reason for the lack of an effect of MNM provision on child health outcomes could be the low dosage of micronutrients provided. In order to obtain approval from the National Institute of Nutrition, we were required to halve the originally suggested dosage. The resulting dosage was well below the recommended daily allowance (RDA) for children of this age, under the 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). Thus, the low quantities may not have been sufficient to impact iron absorption. We hope future work evaluates an MNM supplementation with higher doses.

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 it requires a one-size-fits-all-students approach. 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. As indicated by the IFA results, 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.

Policy Implications from Increased Monitoring

Finally, the robust positive impact of high intensity monitoring on child hemoglobin levels (Tables 10 and 11) is particularly interesting and relevant for policy, especially in the absence of an interaction effect with the MNM treatment. 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 (Table 7) and did not have more mid-day meal visits where enumerators noted a powdery addition to the meal or cooks reporting the addition of the mix (Table 8). At the same time, the impact of high intensity monitoring on the quality of the meals provided was not particularly large and unlikely to affect hemoglobin levels (Table 13).

A likely explanation for the effect of the high intensity monitoring can be seen in Table 12: students in high intensity schools were more likely to report receiving IFA tablets regularly. While this could be seen as a spillover effect, it is also possible that it was a direct monitoring effect. Since schools almost uniformly reported that they distributed the IFA tablets during 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. Note that the stakes are extremely low – since tablets were to be distributed once a week, it would have been very easy for teachers to simply say they had distributed the tablets a previous day or were going to on a future day.

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 the block and other 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, 4 visits to conduct IFA surveys in months 1, 3, 4 and 5, and at least 3 visits to observe the midday meals in months 3, 4 and 5. The schools in the high intensity monitoring treatment received 2 additional visits to observe midday meals in months 1 and 2. It is puzzling that the addition of 2 visits on top of a base of 9 had such a substantial effect on the propensity of schools 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 to us that they distributed the IFA tablets. Another unusual element about the mid-day 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).

Spillovers

One last policy implication has to do with the negative spillovers of the MNM program on implementation of the IFA distribution. As seen in Table 12, students in MNM treatment schools are less likely to report receiving IFA tablets regularly than students in the comparison schools. Thus, the MNM treatment appears to crowd out the distribution of IFA tablets, perhaps because headmasters and teachers are overburdened, as they appear to be from anecdotal evidence.[20] Note that these negative spillovers are not as evident in the survey responses reported by the school officials, who could potentially get confused between the two programs. There is no other

[20] One of the most common concerns about the mid-day 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.

reason to think the students would report getting IFA tablets less frequently, besides crowding out of different school programs. An important policy implication is that giving additional responsibilities to school officials may lead to declines in the quality of teaching or implementation of other programs.

VII. Conclusion

To conclude, this study finds evidence supporting the effectiveness of a government program to distribute iron supplements in schools and the effectiveness of frequent monitoring visits at improving implementation of such public health programs. While we do not find evidence that the MNM distribution improved child health, this part of the intervention was not optimized to find an effect given the short time period within 1 school year and the restrictions necessary to provide the same intervention to all students. However, the fact that we do find strong take-up rates of the micronutrient mix, and significantly higher amounts of micronutrients in meals, combined with efficacy trials demonstrating the effect of multi-micronutrient supplementation, suggests that multi-micronutrient distribution is still a promising area and should be studied further.

This study contributes to the burgeoning literature on the effectiveness of monitoring visits, even with no stakes attached. Even more specifically, the results suggest that the exact timing of such visits and who the auditors speak to may have significant effects. Another important contribution of our study is the finding that adding on an increasing number of programs in schools may have negative spillovers on other programs if school officials already feel overburdened. While schools are a natural setting for implementing a number of 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. This is an area that is currently understudied in the literature, and warrants further research.

REFERENCES

Afridi, Farzana. 2010. "Child welfare programs and child nutrition: Evidence from a mandated school meal program in India." *Journal of Development Economics* 92(2), 152-165.Ahmed, Faruk, Moududur R. Khan, Mohammad Akhtaruzzaman, Rezaul Karim, Gail Williams, Harriet Torlesse, Ian Darnton-Hill, Nita Dalmiya, Cadi P. Banu, and Badrun Nahar. 2010. “Long-Term Intermittent Multiple Micronutrient Supplementation Enhances Hemoglobin and Micronutrient Status More Than Iron + Folic Acid Supplementation in Bangladeshi Rural Adolescent Girls with Nutritional Anemia.” *The Journal of Nutrition* 140 (10): 1879–86. doi:10.3945/jn.109.119123.

Aker, J. and C. Ksoll. 2015. “Call Me Educated: Evidence from a Mobile Monitoring Experiment in Niger”. Working Paper.

Berry, J., P. Mukherjee, and K. Shastry (2012). “Taken with a Grain of Salt? Micronutrient Fortification in South Asia”, *CESifo Economic Studies (2012) 58(2): 422-449*

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

Bjorkman, M. and Svensson, J. 2009. “Power to the people: Evidence from a random- ized field experiment on community-based monitoring in Uganda”. *The Quarterly Journal of Economics*, 124 (2), 735–769.

Debnath, S. and S. Sekhri (2016). ““No Free Lunch: Using Technology to Improve the Efficacy of School Feeding Programs”, Working Paper.

Gera, Tarun, Hps Sachdev, Penelope Nestel, and Sudeep Singh 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. doi:10.1097/01.mpg.0000243440.85452.38.

Goverment of Odisha, Director of Family Welfare. 2012. “Allocation of Required IFA & Albendazole Tablets for WIFS,” December 20.

Government of Odisha, Women and Child Development Department. 2011. “Guidelines for the implementation of Mid Day Meal programme in Primary and Upper Primary Schools of Odisha.”

Gupta, Anuradha, Suresh Mohammed, Anjana Saxena, Sheetal Rahi, and Anshu Mohan. 2012. “Operational Framework: Weekly Iron and Folic Acid Supplementation Programme for Adolescents.” Ministry of Health and Family Welfare. http://tripuranrhm.gov.in/Guidlines/WIFS.pdf.

Hirve, Siddhivinayak, Sheila Bhave, Ashish Bavdekar, Sadanand Naik, Anand Pandit, Claudia Schauer, Anna Christofides, Ziauddin Hyder, and Stanley 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, Farhana Haseen, Marufa Khan, Tom Schaetzel, Chowdhury S. B. Jalal, Mizanur Rahman, Bo Lönnerdal, Venkatesh Mannar, and Haile 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.

International Institute for Population Sciences (IIPS), 2010. “District Level Household and Facility Survey (DLHS-3), 2007-08: India.” Mumbai: IIPS.

Jacoby, Hanan. 2002. “Is there an intrahousehold flypaper effect? Evidence from a school feeding program.” *Economic Journal* 112 (476): 196-221.

Low, Michael, Ann Farrell, Beverley-Ann Biggs, and Sant-Rayn Pasricha. 2013. “Effects of Daily Iron Supplementation in Primary-School--Aged Children: Systematic Review and Meta-Analysis of Randomized Controlled Trials.” CMAJ: Canadian Medical Association Journal, November 19. General OneFile.

Luo, Renfu, Yaojiang Shi, Linxiu Zhang, Chengfang Lui, Scott Rozelle, Brian Sharbono, Ai Yue, Qiran Zhao, and Reynaldo Martorell. 2012. “Nutrition and Educational Performance in Rural China’s Elementary Schools: Results of a Randomized Control Trial in Shaanxi Province.” Economic Development and Cultural Change 60 (4): 735–72. doi:http://www.jstor.org/action/showPublication?journalCode=econdevecultchan.

Miller, G., R. Luo, L. Zhang, S. Sylvia, Y. Shi, P. Foo, Q. Zhao, R. Martorell, A. Medina, S. Rozelle Helen F Farnsworth. 2012. “Effectiveness of provider incentives for anaemia reduction in rural China: a cluster randomised trial”. BMJ 2012;345:e4809 doi: 10.1136/bmj.e4809

Muralidharan, K., P. Niehaus and S. Sukhtankar. 2014. “Building State Capacity: Evidence from Biometric Smartcards in India”. Working Paper 19999, National Bureau of Economic Research.

Olken, B. 2007. “Monitoring Corruption: Evidence from a Field Experiment in Indonesia”. *Journal of Political Economy* 115 (2), pp. 200-249.

Ramakrishnan U, Aburto N, McCabe G, Martorell R. Multimicronutrient interventions but not vitamin A or iron interventions alone improve child growth: results of 3 meta-analyses. J Nutr. 2004;134:2592–2602.

Shastry, G.K and D. Weil (2003), Journal of the European Economic Association 1(2-3), pages 387-396.

Thomas, D. et al. (2006), “Causal effect of health on labor market outcomes: Experimental evidence,” mimeo, UCLA.

Vermeersch, Christel, and Michael Kremer. "School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation." World Bank Policy Research Working Paper Series No. 2523, 2005.

Vucic, Vesna, Cristiana Berti, Christiane Vollhardt, Katalin Fekete, Irene Cetin, Berthold Koletzko, Mirjana Gurinovic, and Pieter van’t Veer. 2013. “Effect of Iron Intervention on Growth during Gestation, Infancy, Childhood, and Adolescence: A Systematic Review with Meta-Analysis.” Nutrition Reviews 71 (6): 386–401. doi:10.1111/nure.12037.

WHO. 2011. “Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity.” WHO/NMH/NHD/MNM/11.1. Vitamin and Mineral Nutrition Information System. Geneva, World Health Organization. http://www.who.int/vmnis/indicators/haemoglobin.pdf.

"WHO | Micronutrient Deficiencies." 2015. WHO. Accessed October 29. http://www.who.int/nutrition/topics/ida/en/.

Figures

Figure 1: Distribution of child hemoglobin levels at baseline in IFA and non-IFA schools

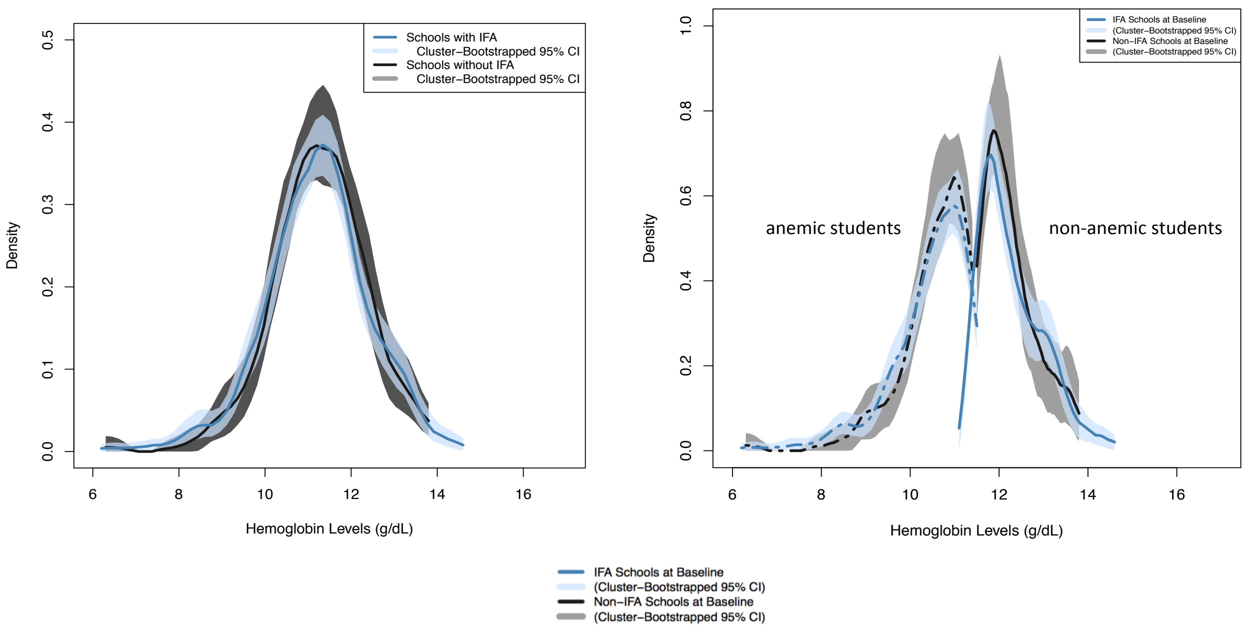
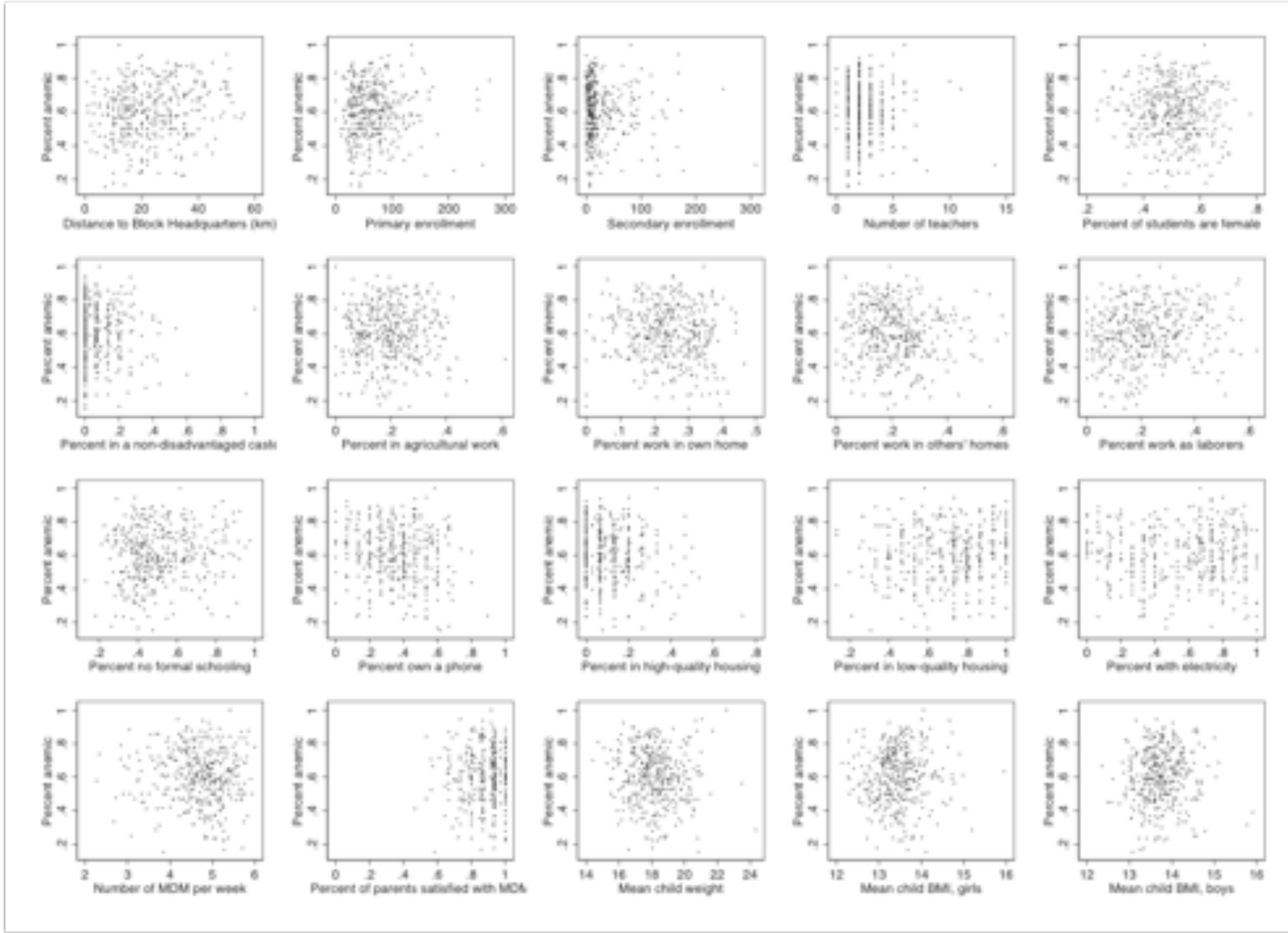
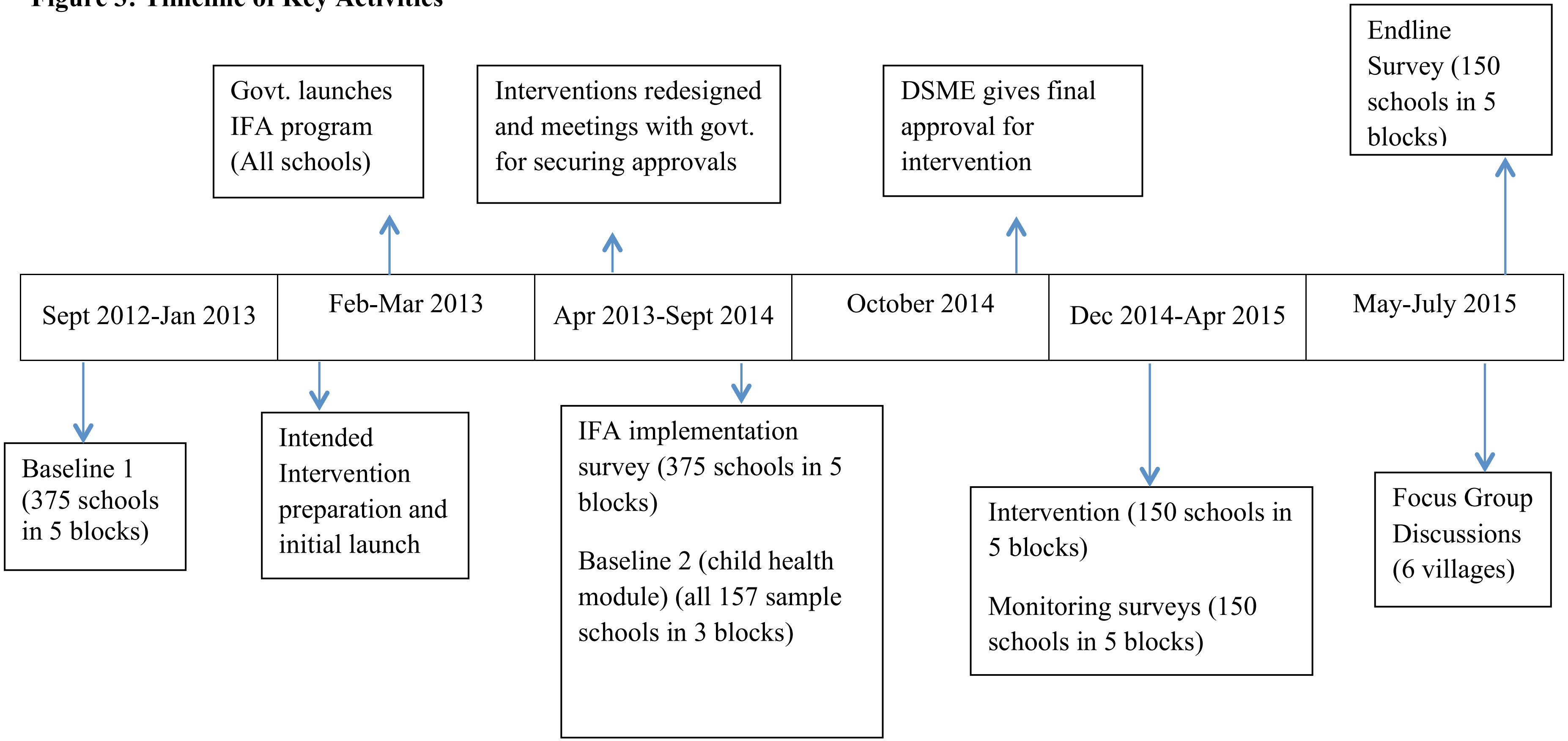


Figure 2: Observable demographic characteristics at school level are uncorrelated with anemia prevalence.



Note: Each graph plots the percent of students per school who are anemic versus some observable demographic characteristic. Anemia prevalence is not correlated with any observable school characteristics.

Figure 3: Timeline of Key Activities



Tables

Table 1: IFA implementation by block

		Blocks with high variation in IFA			Blocks with low variation in IFA	
		Block 1	**Block 2**	**Block 3**	**Block 4**	**Block 5**
SCH	Percent of schools received IFA	0.49	0.83	0.62	0.95	0.99
	Mean number of IFA tablets received per student (conditional)	15.12	57.47	38.45	65.74	101.27
	Mean percent of 3 kids saying they receive tablets daily	0.19	0.52	0.56	0.80	0.69
	Mean percent of 3 kids saying they received tablets the previous day	0.12	0.23	0.10	0.27	0.39
	Percent received deworming medication	0.42	0.77	0.67	0.93	0.81
	Mean number of deworming doses per student (conditional)	1.76	2.59	2.10	2.32	2.37
	Number of schools	43	93	21	117	103

Table 2: Comparison of high-variation and low-variation blocks

		High IFA variation	Low IFA variation	P-Value
	Panel A: Demographic Characteristics	**Blocks 1-3**	**Blocks 4-5**	
School-reported (SCH)	Distance to the block headquarters (km)	21.31	24.17	**0.0218**
	Primary enrollment	75.21	60.15	**0.0004**
	Secondary enrollment	30.09	24.99	0.2082
	Number of teachers	2.53	2.40	0.4332
	Percent of schools have a kitchen	0.73	0.81	**0.0984**
	Percent of schools have at least one latrine	0.85	0.86	0.8817
	Percent of schools have sufficient water	0.74	0.66	**0.0770**
Household-reported, aggregated to school level (HH)	Mean % of students are female	0.50	0.51	0.5813
	Mean % of families in a non-disadvantaged caste	0.04	0.10	**0.0000**
	Mean % of village adults in agricultural work	0.19	0.20	0.5387
	Mean % of village adults work in own home	0.25	0.23	**0.0325**
	Mean % of village adults work in others' homes	0.26	0.18	**0.0000**
	Mean % of village adults work as laborers	0.16	0.27	**0.0000**
	Mean % of village adults with no formal schooling	0.57	0.46	**0.0000**
	Mean % of village adults who own a phone	0.31	0.39	**0.0000**
	Mean % of families that live in high-quality housing	0.09	0.13	**0.0096**
	Mean % of families with electricity	0.52	0.55	0.2621
	Panel B: Implementer Variables			
SCH	Percent with parent group for MDM	0.12	0.63	**0.0000**
	Percent with MDM training	0.58	0.15	**0.0000**
	Percent receiving MDM rice on a regular schedule	0.38	0.46	0.1458
HH	Mean number of MDM per week	4.76	4.64	**0.0839**
	Mean % of parents satisfied with MDM	0.90	0.87	**0.0025**
	Panel C: Anthropometric Measures at Baseline			
Child-level measures aggregated to school level and averaged	Mean % of students with anemia	0.57	0.63	**0.0002**
	Mean % with mild anemia	0.23	0.24	0.4324
	Mean % with moderate anemia	0.33	0.38	**0.0007**
	Mean % with severe anemia	0.01	0.01	0.5127
	Mean child Hb level	11.20	11.04	**0.0003**
	Mean student BMI	13.68	13.55	**0.0909**
	Mean student weight	18.30	18.11	0.2002
	Mean BMI, girls	13.54	13.41	0.1900
	Mean BMI, boys	13.81	13.67	**0.0447**
	Panel D: IFA Implementation Variables			
SCH	Percent of schools received IFA	0.71	0.97	**0.0000**
	Mean number of IFA tablets received per student (conditional)	49.92	82.10	**0.0000**
	Mean percent of 3 kids saying they receive tablets daily	0.44	0.75	**0.0000**
	Mean percent of 3 kids saying they received tablets the previous day	0.18	0.33	**0.0007**
	Percent received deworming medication	0.66	0.88	**0.0000**
	Mean number of deworming doses per student (conditional)	2.42	2.34	0.7478
	Number of schools	157	220	

Note: P-value tests the difference in the two means, unconditional on block. Bolded p-values are significant at the 10% level.

Table 3: Comparison of IFA, non-IFA schools in high-variation blocks

		Blocks with high IFA variation		
	Panel A: Demographic Characteristics	Got IFA	No IFA	P-Value
School-reported (SCH)	Distance to the block headquarters (km)	20.59	23.02	0.1355
	Primary enrollment	75.02	75.67	0.9330
	Secondary enrollment	29.73	30.96	0.8691
	Number of teachers	2.61	2.33	0.3328
	Percent of schools have a kitchen	0.77	0.64	0.1085
	Percent of schools have at least one latrine	0.86	0.83	0.5644
	Percent of schools have sufficient water	0.76	0.69	0.3584
Household-reported, aggregated to school level (HH)	Mean % of students are female	0.50	0.50	0.6869
	Mean % of families in a non-disadvantaged caste	0.03	0.06	**0.0320**
	Mean % of village adults in agricultural work	0.20	0.18	0.2398
	Mean % of village adults work in own home	0.27	0.22	**0.0112**
	Mean % of village adults work in others' homes	0.25	0.30	**0.0057**
	Mean % of village adults work as laborers	0.15	0.18	0.1367
	Mean % of village adults with no formal schooling	0.56	0.60	0.3012
	Mean % of village adults who own a phone	0.30	0.33	0.4004
	Mean % of families that live in high-quality housing	0.09	0.11	0.2287
	Mean % of families with electricity	0.52	0.51	0.7547
	Panel B: Implementer Variables			
SCH	Percent with parent group for MDM	0.10	0.16	0.3282
	Percent with MDM training	0.61	0.51	0.2771
	Percent receiving MDM rice on a regular schedule	0.36	0.43	0.3551
HH	Mean number of MDM per week	4.81	4.63	0.1019
	Mean % of parents satisfied with MDM	0.90	0.90	0.7568
	Panel C: Anthropometric Measures at Baseline			
Child-level measures aggregated to school level and averaged	Mean % of students with anemia	0.56	0.58	0.5765
	Mean % with mild anemia	0.23	0.23	0.8608
	Mean % with moderate anemia	0.32	0.35	0.3013
	Mean % with severe anemia	0.01	0.00	**0.0210**
	Mean child Hb level	11.21	11.19	0.7580
	Mean student BMI	13.65	13.74	0.6224
	Mean student weight	18.33	18.21	0.6273
	Mean BMI, girls	13.51	13.60	0.6679
	Mean BMI, boys	13.77	13.90	0.3175
	Number of schools	111	46	

Note: P-value tests the difference in the two means, unconditional on block. Bolded p-values are significant at the 10% level.

Table 4: Predictors of IFA receipt in high-variation blocks

	Dependent Variable: Received IFA Indicator											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	All observed characteristics		**Demographic characteristics**		**Demographic characteristics**		**Sig. Predictors from Cols 1-6**		**Cols 7-8 + Implementer Variables**		**Final control variables**	
Distance to the block hq (km)	-0.004	-0.002	-0.003	-0.003	-0.003	0.006	--	--	--	--	-0.000	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.011)	--	--	--	--	(0.004)	(0.004)
Mean % of pop in non-disadvanta	-0.909*	-0.745	-1.074**	-0.907*	-1.352**	-1.086	-1.061**	-0.853*	-0.929*	-0.710	-0.887**	-0.778*
	(0.532)	(0.524)	(0.517)	(0.504)	(0.659)	(1.062)	(0.475)	(0.458)	(0.492)	(0.473)	(0.440)	(0.420)
Percent of students are female	-0.066	-0.149	-0.012	-0.094	-0.320	1.926	--	--	--	--	--	--
	(0.475)	(0.464)	(0.470)	(0.456)	(0.430)	(1.162)	--	--	--	--	--	--
Primary enrollment	-0.001	-0.001	-0.001	-0.001	-0.002	0.004	--	--	--	--	--	--
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	--	--	--	--	--	--
Secondary enrollment	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	--	--	--	--	--	--
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	--	--	--	--	--	--
Number of teachers	0.062	0.049	0.055	0.046	0.061	-0.027	--	--	--	--	--	--
	(0.048)	(0.048)	(0.047)	(0.047)	(0.060)	(0.085)	--	--	--	--	--	--
Percent has a kitchen	0.034	-0.033	0.052	-0.022	0.254*	-0.149	--	--	--	--	0.086	-0.074
	(0.109)	(0.109)	(0.107)	(0.106)	(0.137)	(0.177)	--	--	--	--	(0.093)	(0.096)
Percent has at least one latrine	0.163	0.083	0.167	0.096	0.345**	-0.105	0.138	0.058	0.142	0.049	--	--
	(0.127)	(0.127)	(0.125)	(0.124)	(0.154)	(0.218)	(0.115)	(0.113)	(0.116)	(0.114)	--	--
Percent has sufficient water	0.043	0.066	0.037	0.047	-0.060	0.164	--	--	--	--	--	--
	(0.103)	(0.100)	(0.100)	(0.097)	(0.099)	(0.188)	--	--	--	--	--	--
Mean % of village adults in agric	-1.415	-0.949	-1.532*	-1.198	-1.312*	-0.852	-0.919	-1.064	-0.820	-0.899	--	--
	(0.888)	(0.877)	(0.863)	(0.840)	(0.719)	(3.020)	(0.703)	(0.678)	(0.720)	(0.689)	--	--
Mean % of village adults work in	-0.318	-0.070	-0.435	-0.032	0.542	1.470	--	--	--	--	--	--
	(0.727)	(0.726)	(0.710)	(0.711)	(0.625)		--	--	--	--	--	--
Mean % of village adults work in	-0.982	-0.689	-1.133*	-0.715	-0.418	0.008	-0.822**	-0.735**	-0.840**	-0.731**	-0.843**	-0.551*
	(0.646)	(0.638)	(0.627)	(0.624)	(0.552)	(2.272)	(0.361)	(0.347)	(0.364)	(0.350)	(0.337)	(0.329)
Mean % of village adults work as	-1.385*	-0.738	-1.570**	-0.910	-1.540**	0.530	-1.050*	-0.755	-0.972	-0.572	--	--
	(0.785)	(0.795)	(0.760)	(0.765)	(0.679)	(2.568)	(0.582)	(0.569)	(0.606)	(0.593)	--	--
Percent of villagers with no forma	-0.385	0.026	-0.381	0.037	-0.034	-1.286	-0.561	-0.096	-0.529	-0.145	--	--
	(0.468)	(0.498)	(0.450)	(0.477)	(0.421)	(1.003)	(0.412)	(0.437)	(0.433)	(0.445)	--	--
Percent of villagers who own a pl	-0.357	-0.430	-0.475	-0.504	0.222	-1.910**	-0.448	-0.463	-0.413	-0.420	--	--
	(0.366)	(0.359)	(0.357)	(0.347)	(0.319)	(0.812)	(0.340)	(0.326)	(0.342)	(0.328)	--	--
Percent of villagers who live in h	-0.903	-0.534	-0.896	-0.532	-0.715	-1.695	--	--	--	--	--	--
	(0.623)	(0.620)	(0.614)	(0.606)	(0.612)	(1.315)	--	--	--	--	--	--
Percent of villagers who live in lo	-0.196	-0.169	-0.253	-0.173	0.503	-2.168**	0.196	0.181	0.204	0.202	--	--
	(0.430)	(0.424)	(0.425)	(0.415)	(0.389)	(1.055)	(0.303)	(0.298)	(0.304)	(0.298)	--	--
Percent of villagers with electrici	0.086	0.099	0.050	0.081	0.100	-0.160	--	--	--	--	--	--
	(0.159)	(0.155)	(0.156)	(0.151)	(0.168)	(0.279)	--	--	--	--	--	--
Mean number of MDM per week	0.040	0.030	--	--	--	--	--	--	0.052	0.032	--	--
	(0.067)	(0.066)	--	--	--	--	--	--	(0.063)	(0.061)	--	--
Mean % of parents satisfied with	0.290	0.598	--	--	--	--	--	--	0.298	0.602	0.444	0.856*
	(0.508)	(0.507)	--	--	--	--	--	--	(0.476)	(0.461)	(0.456)	(0.441)
Percent with parent group for M	-0.170	-0.061	--	--	--	--	--	--	--	--	--	--
	(0.130)	(0.134)	--	--	--	--	--	--	--	--	--	--
Percent with MDM training	0.095	0.047	--	--	--	--	--	--	--	--	--	--
	(0.092)	(0.090)	--	--	--	--	--	--	--	--	--	--
Percent receiving MDM rice on a	-0.056	0.037	--	--	--	--	--	--	--	--	--	--
	(0.094)	(0.099)										
Constant	1.518	0.467	2.173***	1.294	0.660	2.334	1.531***	1.025**	0.943	0.263	0.531	-0.047
	(1.002)	(1.042)	(0.817)	(0.839)	(0.740)	(2.567)	(0.469)	(0.485)	(0.722)	(0.713)	(0.460)	(0.455)
Blocks	No Block F.E.	Block F.E.	No Block F.E.	Block F.E.	Block 2	Blocks 1 and 3	No Block F.E.	Block F.E.	No Block F.E.	Block F.E.	No Block F.E.	Block F.E.
N	124	124	124	124	76	48	124	124	124	124	124	124
R-squared	0.212	0.273	0.183	0.252	0.392	0.397	0.145	0.231	0.155	0.246	0.096	0.213

*Note: Standard errors in parentheses. The dependent variable is the indicator for receiving IFA tablets in every column. Columns 1 and 2 include every observable characteristic of each school; the remaining columns restrict to particular subsets of observable characteristics. Columns 3 and 4 only include demographic characteristics (at both the school and household level, see Table 1,3, or 4 for a description of which are reported by the household and aggregated to the school level and which are reported by the school). Columns 5 and 6 repeat the analysis for demographic characteristics but separate schools by block (blocks 1 and 3 are combined due to the small number of schools in each block). Columns 7 and 8 restrict the independent variables to those that are significant in any of the previous columns. Columns 9 and 10 repeat columns 7 and 8 but add indicators of implementing ability of each school. Finally, Columns 11 and 12 include the final control variables used in the main analysis of the paper. These include a proxy for each of the three possible ways in which BEOs could have distributed tablets systematically (for full discussion of these decision-making processes see Section 4B) as well as the two robustly significant variables in the previous columns. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 5: Balance across treatments at baseline: Household Characteristics

	Control	Only MNM	Only High	Both	p-value of all 3 differences
Child health outcomes					
Hemoglobin	11.107	11.081	11.170	11.000	0.54
z - weight	-1.851	-1.930	-1.810	-1.957	0.49
z - height	-1.367	-1.355	-1.491	-1.420	0.86
MUAC	15.052	15.201	15.167	15.101	0.75
Household-level data					
Non scheduled caste/tribe	0.057	0.045	0.094	0.087	0.23
Owns phone	0.328	0.375	0.346	0.323	0.60
Has electricity	0.517	0.497	0.592	0.480	0.35
House is *pucca*	0.109	0.088	0.097	0.100	0.86
Is satisfied with school meals	0.895	0.868	0.868	0.901	0.53
Has heard of anemia	0.086	0.084	0.087	0.070	0.88
Child demographics					
Age	6.756	6.859	6.984	6.642	0.78
Female Dummy	0.486	0.458	0.489	0.510	0.68
Not Child of Head of Household	0.128	0.110	0.134	0.121	0.78
Number of times child had MDM in past week	4.760	4.773	4.820	4.811	0.99
Takes any supplements	0.003	0.003	0.020 **	0.007	0.21
Has taken deworming pill in past year	0.126	0.112	0.103	0.119	0.85
Mother demographics					
Age	31.209	31.173	30.867	30.795	0.87
Is literate	0.387	0.357	0.378	0.411	0.86
Completed Primary School	0.015	0.025	0.026	0.027	0.71
Completed Middle School	0.019	0.014	0.019	0.027	0.81
Completed High School	0.004	0.007	0.008	0.008	0.90
More than High School	0.004	0.022 *	0.008	0.015	0.24
Not housewife	0.305	0.355	0.352	0.462 **	0.09
Has a Job Card	0.637	0.734 *	0.683	0.663	0.25
Head of household demographics					
Age	38.990	37.845	38.921	37.950	0.34
Is literate	0.505	0.594 *	0.520	0.542	0.26
Completed Primary School	0.020	0.044 *	0.040	0.047 *	0.16
Completed Middle School	0.031	0.047	0.050	0.050	0.52
Completed High School	0.014	0.016	0.020	0.020	0.88
More than High School	0.020	0.047 *	0.030	0.040	0.33
Occupation in Agriculture	0.470	0.478	0.431	0.367 *	0.19
Has a Job Card	0.723	0.783	0.776	0.686	0.07

Notes: This table presents balance checks on household characteristics and child health at baseline, across each of the treatment groups for those who have endline data as well. Each row shows the mean for that variable for the following groups: (i) control group, (ii) schools that only received the MNM treatment group, (iii) schools that only received the high intensity monitoring group, and (iv) the group that received both MNM as well as high intensity monitoring. Significance levels of the difference with the control group are indicated after each number. The final column provides the p-value of the F-test of all three differences.

*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: Balance across treatments at baseline: School characteristics

	Control	Only MNM	Only High	Both	p-value of all 3 differences
School-level variables					
Distance to the block headquarters (km)	22.155	22.789	23.383	24.889	0.59
Primary enrollment	64.720	70.763	67.980	63.324	0.80
Secondary enrollment	24.760	24.263	30.313	27.432	0.64
Number of teachers	2.307	2.421	2.600	2.486	0.45
Number of female teachers	2.700	2.868	2.747	2.676	0.90
Number of rooms	3.986	4.444	3.854	3.778	0.47
Percent of schools have a kitchen	0.738	0.833	0.823 *	0.676	0.13
Percent of schools have at least one latrine	0.865	0.789	0.857	0.865	0.77
Percent of schools have sufficient water	0.671	0.667	0.739	0.622	0.43
Percent with parent group for MDM	0.401	0.444	0.466	0.343	0.51
Percent with MDM training	0.368	0.324	0.299	0.333	0.67
Percent receiving MDM rice on a regular schedule					

Notes: This table presents balance checks on school characteristics at baseline, across each of the treatment groups for those who have endline data. Each row shows the mean for that variable for the following groups: (i) control group, (ii) schools that only received the MNM treatment group, (iii) schools that only received the high intensity monitoring group, and (iv) the group that received both MNM as well as high intensity monitoring. The final column provides the p-value of the F-test of all three differences.

*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: 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. The even columns also include a dummy for whether the school received IFA 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.*

Table 8: 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***	0.131***	0.131***	0.715***	0.646***	0.715***	0.452***	0.412***	0.453***
	(0.023)	(0.037)	(0.023)	(0.035)	(0.064)	(0.035)	(0.033)	(0.054)	(0.032)
High Intensity	-0.007	-0.007	-0.007	0.066*	0.012	0.067*	0.051	0.017	0.058
	(0.022)	(0.012)	(0.023)	(0.040)	(0.023)	(0.040)	(0.041)	(0.020)	(0.041)
MNM treatment * High Intensity		-0.000			0.111			0.066	
		(0.046)			(0.075)			(0.068)	
Received IFA during previous year			0.000			-0.027			-0.114*
			(0.032)			(0.082)			(0.062)
N	554	554	554	536	536	536	532	532	532
R-squared	0.088	0.088	0.088	0.573	0.576	0.573	0.298	0.299	0.303
p-value of F-test (high & interaction)		0.847	.	.	0.238	.	.	0.446	.
Panel B: December to January - High Intensity Schools (2 visits each)									
MNM treatment	0.152***	0.152***	0.152***	0.667***	0.667***	0.668***	0.450***	0.450***	0.449***
	(0.047)	(0.047)	(0.047)	(0.067)	(0.067)	(0.067)	(0.061)	(0.061)	(0.060)
N	139	139	139	133	133	133	131	131	131
R-squared	0.106	0.106	0.107	0.522	0.522	0.523	0.319	0.319	0.348
Panel C: February to March - All Schools (2 visits each)									
MNM treatment	0.095***	0.123***	0.095***	0.751***	0.649***	0.752***	0.482***	0.437***	0.482***
	(0.028)	(0.045)	(0.028)	(0.045)	(0.068)	(0.045)	(0.043)	(0.059)	(0.043)
High Intensity	-0.023	0.006	-0.022	0.075*	-0.026	0.078*	0.058	0.010	0.061
	(0.026)	(0.008)	(0.027)	(0.045)	(0.025)	(0.044)	(0.046)	(0.011)	(0.046)
N	277	277	277	269	269	269	267	267	267
R-squared	0.075	0.080	0.076	0.624	0.635	0.625	0.319	0.322	0.321
Panel D: April to May - All Schools (1 visit each)									
MNM treatment	0.181***	0.138**	0.179***	0.679***	0.629***	0.680***	0.401***	0.366***	0.403***
	(0.047)	(0.062)	(0.047)	(0.057)	(0.082)	(0.058)	(0.059)	(0.084)	(0.059)
High Intensity	0.032	-0.009	0.027	0.047	-0.004	0.049	0.028	-0.007	0.037
	(0.046)	(0.018)	(0.047)	(0.056)	(0.026)	(0.057)	(0.060)	(0.030)	(0.062)
N	138	138	138	134	134	134	134	134	134
R-squared	0.156	0.162	0.161	0.577	0.580	0.577	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 it 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 the (iii) likelihood that the mix was present in the room where food materials are stored (Columns 7-9). All columns include block 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 schoolyear.

*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: Micronutrient Levels in MDM from Lab Tests of Food Samples

	February						April					
Panel B	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, zinc and vitamin A) present in school meals, as measured in the laboratory using samples collected by enumerators during February and April. 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 10: Treatment Effects on Health Outcomes - Lagged Dependent Variable Model

Lagged dep var from survey	None	Baseline I	Baseline II	Both with dummies for missing				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Dep var: hemogloblin (g/dl)				
MNM treatment	-0.042	-0.009	-0.081	-0.017	0.029	0.023	-0.024	0.026
	(0.057)	(0.067)	(0.100)	(0.057)	(0.072)	(0.073)	(0.078)	(0.072)
High Intensity	0.170***	0.224***	0.156	0.177***	0.223***	0.211***	0.198**	0.225***
	(0.058)	(0.067)	(0.105)	(0.058)	(0.079)	(0.080)	(0.080)	(0.079)
MNM treatment * High Intensity					-0.092	-0.087	-0.037	-0.091
					(0.114)	(0.114)	(0.119)	(0.113)
N	1921	1108	349	1921	1921	1921	1769	1921
R-squared	0.024	0.169	0.173	0.127	0.128	0.129	0.128	0.128
				Panel B: Dep var: weight (kilos)				
MNM treatment	-0.210	0.100	-0.024	-0.162	-0.108	-0.091	-0.151	-0.125
	(0.222)	(0.144)	(0.167)	(0.129)	(0.184)	(0.182)	(0.194)	(0.186)
High Intensity	-0.119	-0.101	0.105	-0.017	0.037	0.070	0.028	0.048
	(0.224)	(0.139)	(0.169)	(0.129)	(0.195)	(0.193)	(0.200)	(0.194)
MNM treatment * High Intensity					-0.110	-0.126	-0.115	-0.104
					(0.261)	(0.258)	(0.272)	(0.261)
N	1947	1114	355	1947	1947	1947	1795	1947
R-squared	0.480	0.743	0.891	0.719	0.719	0.720	0.718	0.719
				Panel C: Dep var: height (cm)				
MNM treatment	-0.396	-0.480	-0.589	-0.635*	-0.763	-0.685	-0.849	-0.843*
	(0.462)	(0.464)	(0.745)	(0.369)	(0.500)	(0.479)	(0.520)	(0.493)
High Intensity	0.211	-0.097	0.049	0.361	0.231	0.382	0.208	0.278
	(0.466)	(0.452)	(0.738)	(0.377)	(0.551)	(0.534)	(0.561)	(0.543)
MNM treatment * High Intensity					0.258	0.179	0.217	0.282
					(0.746)	(0.729)	(0.768)	(0.745)
N	1943	1119	354	1943	1943	1943	1791	1943
R-squared	0.474	0.474	0.692	0.581	0.581	0.583	0.573	0.582
				Panel D: mid-upper-arm circumference (cm)				
MNM treatment	-0.128	-0.120	-0.154*	-0.160**	-0.143	-0.139	-0.204*	-0.155
	(0.090)	(0.094)	(0.083)	(0.072)	(0.102)	(0.101)	(0.109)	(0.102)
High Intensity	-0.021	-0.071	0.025	-0.061	-0.043	-0.035	-0.078	-0.036
	(0.091)	(0.103)	(0.085)	(0.074)	(0.110)	(0.111)	(0.115)	(0.109)
MNM treatment * High Intensity					-0.034	-0.038	0.042	-0.030
					(0.143)	(0.142)	(0.152)	(0.142)
N	1947	1118	355	1947	1947	1947	1795	1947
R-squared	0.316	0.394	0.782	0.483	0.484	0.484	0.482	0.484

	Panel E: Weight-for-age (z score)							
MNM treatment	-0.117*	0.028	-0.084	-0.068	0.035	0.038	0.042	0.045
	(0.067)	(0.063)	(0.111)	(0.053)	(0.074)	(0.073)	(0.077)	(0.076)
High Intensity	-0.043	-0.113	-0.028	-0.049	0.055	0.061	0.032	0.046
	(0.068)	(0.070)	(0.111)	(0.053)	(0.072)	(0.071)	(0.073)	(0.072)
MNM treatment * High Intensity					-0.204**	-0.206**	-0.196*	-0.205**
					(0.103)	(0.102)	(0.106)	(0.103)
N	1161	475	188	1161	1161	1161	1066	1161
R-squared	0.033	0.493	0.481	0.302	0.305	0.305	0.314	0.305
	Panel F: Height-for-age (z score)							
MNM treatment	-0.069	-0.045	-0.107	-0.098	-0.115	-0.103	-0.141	-0.124
	(0.078)	(0.086)	(0.149)	(0.067)	(0.092)	(0.089)	(0.096)	(0.092)
High Intensity	0.047	-0.053	0.059	0.075	0.058	0.082	0.040	0.064
	(0.078)	(0.084)	(0.151)	(0.067)	(0.100)	(0.099)	(0.102)	(0.100)
MNM treatment * High Intensity					0.034	0.021	0.042	0.037
					(0.134)	(0.131)	(0.139)	(0.134)
N	1873	1064	342	1873	1873	1873	1723	1873
R-squared	0.047	0.231	0.198	0.172	0.172	0.175	0.172	0.172

Notes: The dependent variable in each specification is child's hemoglobin in g/dl (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), child's z-score for weight for age (Panel E), and child's z-score for height for age (Panel F). All columns include block and age fixed effects, in addition to the lagged dependent variable as described in the headers. . Columns 6-9 include measures of IFA receipt during the previous year (a dummy for receiving the IFA tablets in Column 6, the percent of students who say they get IFA tablets regularly in Column 7 and the percent of students who say they got IFA tablets yesterday in Column 8).

*Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 11: Treatment Effects on Health Outcomes - Difference in Difference Estimates

				Block fixed effects interacted with midline & endline						
Additional controls								Measures of IFA receipt during previous year		
Fixed effects		School	Child		School	Child	School	School		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Dep var: hemogloblin (g/dl)										
Endline * MNM treatment	0.041	0.020	0.010	0.061	0.040	0.033	0.014	0.036	0.031	0.027
	(0.089)	(0.090)	(0.138)	(0.081)	(0.083)	(0.125)	(0.104)	(0.082)	(0.087)	(0.085)
Endline * High Intensity	0.198**	0.208**	0.223	0.204**	0.215**	0.228*	0.188*	0.196**	0.243***	0.225***
	(0.089)	(0.090)	(0.137)	(0.081)	(0.084)	(0.125)	(0.113)	(0.083)	(0.087)	(0.085)
Endline * High Intensity * MNM treatment							0.052			
							(0.166)			
N	3489	3489	3489	3489	3489	3489	3489	3489	3192	3489
R-squared	0.024	0.099	0.732	0.040	0.109	0.740	0.109	0.110	0.108	0.109
p-value of F-test (high & interaction)		.	.	.	.	.	0.040	.	.	.
Panel B: Dep var: weight (kilos)										
Endline * MNM treatment	0.055	0.043	-0.034	0.042	0.031	-0.040	-0.110	0.037	-0.020	0.020
	(0.181)	(0.190)	(0.214)	(0.175)	(0.180)	(0.211)	(0.244)	(0.180)	(0.187)	(0.188)
Endline * High Intensity	0.010	-0.047	-0.109	-0.048	-0.115	-0.153	-0.258	-0.082	-0.198	-0.100
	(0.179)	(0.188)	(0.215)	(0.173)	(0.181)	(0.208)	(0.274)	(0.189)	(0.192)	(0.183)
Endline * High Intensity * MNM treatment							0.283			
							(0.357)			
N	3511	3511	3511	3511	3511	3511	3511	3511	3215	3511
R-squared	0.504	0.565	0.945	0.508	0.567	0.945	0.568	0.568	0.572	0.568
Panel C: height (cm)										
Endline * MNM treatment	-0.468	-0.409	-0.795	-0.469	-0.404	-0.798	-0.619	-0.386	-0.462	-0.497
	(0.522)	(0.548)	(0.733)	(0.514)	(0.537)	(0.750)	(0.625)	(0.532)	(0.555)	(0.526)
Endline * High Intensity	1.049**	0.808	0.198	1.044**	0.782	0.225	0.564	0.891	0.557	0.865
	(0.529)	(0.557)	(0.717)	(0.526)	(0.556)	(0.691)	(0.847)	(0.560)	(0.581)	(0.554)
Endline * High Intensity * MNM treatment							0.432			
							(1.069)			
N	3511	3511	3511	3511	3511	3511	3511	3511	3215	3511
R-squared	0.460	0.512	0.894	0.466	0.514	0.895	0.514	0.514	0.511	0.514
Panel D: mid-upper-arm circumference (cm)										
Endline * MNM treatment	-0.131	-0.131	-0.171	-0.140	-0.141	-0.175	-0.216	-0.140	-0.183*	-0.154*
	(0.090)	(0.094)	(0.135)	(0.089)	(0.091)	(0.134)	(0.148)	(0.091)	(0.098)	(0.090)
Endline * High Intensity	-0.065	-0.048	-0.082	-0.079	-0.069	-0.086	-0.144	-0.060	-0.053	-0.057
	(0.092)	(0.095)	(0.137)	(0.094)	(0.098)	(0.145)	(0.124)	(0.100)	(0.108)	(0.096)
Endline * High Intensity * MNM treatment							0.149			
							(0.186)			
N	3515	3515	3515	3515	3515	3515	3515	3515	3219	3515
R-squared	0.315	0.388	0.865	0.323	0.392	0.866	0.392	0.392	0.388	0.392
Panel E: Weight-for-age (z score)										
Endline * MNM treatment	-0.009	-0.009	0.008	-0.007	-0.005	0.009	0.083	-0.004	0.007	0.006
	(0.075)	(0.078)	(0.078)	(0.075)	(0.077)	(0.069)	(0.111)	(0.077)	(0.081)	(0.080)
Endline * High Intensity	-0.008	-0.007	-0.061	-0.023	-0.021	-0.077	0.069	-0.015	-0.043	-0.030
	(0.075)	(0.079)	(0.079)	(0.075)	(0.079)	(0.068)	(0.108)	(0.080)	(0.085)	(0.078)
Endline * High Intensity * MNM treatment							-0.177			
							(0.155)			
N	2471	2471	2471	2471	2471	2471	2471	2471	2255	2471
R-squared	0.015	0.141	0.959	0.021	0.147	0.960	0.148	0.147	0.156	0.148

					Panel F: Height-for-age (z score)					
Endline * MNM treatment	-0.060	-0.054	-0.142	-0.061	-0.053	-0.141	-0.050	-0.050	-0.057	-0.069
	(0.092)	(0.097)	(0.128)	(0.091)	(0.095)	(0.132)	(0.110)	(0.095)	(0.099)	(0.093)
Endline * High Intensity	0.213**	0.168*	0.048	0.213**	0.165*	0.058	0.168	0.181*	0.130	0.181*
	(0.093)	(0.098)	(0.123)	(0.093)	(0.098)	(0.119)	(0.149)	(0.099)	(0.104)	(0.098)
Endline * High Intensity * MNM treatment							-0.006			
							(0.190)			
N	3432	3432	3432	3432	3432	3432	3432	3432	3138	3432
R-squared	0.032	0.126	0.807	0.044	0.130	0.807	0.130	0.130	0.133	0.130

Notes: The dependent variable in each specification is child's hemoglobin in g/dl (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), child's z-score for weight for age (Panel E), and child's z-score for height for age (Panel F). All columns include age fixed effects, in addition to the controls and fixed effects indicated in the headers. Columns 9-11 include measures of IFA receipt during the previous year (a dummy for receiving the IFA tablets in Column 9, the percent of students who say they get IFA tablets regularly in Column 10 and the percent of students who say they got IFA tablets yesterday in Column 11).

*Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 12: Spillover Effects on IFA Program

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	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)			At least 2 out of 3 students asked say they get meds weekly or more frequently		
Panel A: All Months (4 visits each)												
MNM treatment	-0.023	-0.047*	-0.024	0.052	0.039	0.053	-0.059*	-0.070	-0.060*	-0.067*	-0.088*	-0.067*
	(0.022)	(0.028)	(0.022)	(0.051)	(0.071)	(0.051)	(0.032)	(0.046)	(0.032)	(0.036)	(0.053)	(0.036)
High Intensity	-0.012	-0.036	-0.016	0.045	0.031	0.048	0.085***	0.074*	0.081**	0.085**	0.063	0.080**
	(0.022)	(0.029)	(0.022)	(0.052)	(0.071)	(0.052)	(0.032)	(0.040)	(0.032)	(0.037)	(0.045)	(0.036)
MNM treatment * High Intensity		0.049			0.027			0.022			0.044	
		(0.043)			(0.098)			(0.065)			(0.073)	
Received IFA during previous year			0.062			-0.053			0.052			0.073
			(0.047)			(0.091)			(0.060)			(0.076)
N	557	557	557	555	555	555	538	538	538	538	538	538
R-squared	0.113	0.115	0.118	0.088	0.089	0.089	0.112	0.112	0.113	0.094	0.095	0.096
p-value of F-test (high & interaction)		0.428	.	.	0.678	.	.	0.034	.	.	0.071	.
Panel B: December-January (1 visit per school)												
MNM treatment	-0.017	-0.032	-0.018	0.103	0.232	0.106	-0.024	-0.037	-0.024	-0.016	-0.028	-0.015
	(0.024)	(0.044)	(0.024)	(0.127)	(0.184)	(0.128)	(0.077)	(0.108)	(0.076)	(0.082)	(0.113)	(0.081)
High Intensity	0.039*	0.024	0.033*	-0.069	0.063	-0.048	0.038	0.024	0.023	0.041	0.028	0.023
	(0.022)	(0.024)	(0.020)	(0.140)	(0.180)	(0.134)	(0.079)	(0.111)	(0.078)	(0.084)	(0.119)	(0.083)
MNM treatment * High Intensity		0.032			-0.268			0.027			0.026	
		(0.042)			(0.242)			(0.156)			(0.165)	
Received IFA during previous year			0.080			-0.255			0.208**			0.242**
			(0.073)			(0.260)			(0.101)			(0.117)
N	145	145	145	145	145	145	134	134	134	134	134	134
R-squared	0.062	0.065	0.089	0.098	0.104	0.106	0.146	0.146	0.163	0.135	0.135	0.155
Panel C: February - May (3 visits per school)												
MNM treatment	-0.026	-0.052	-0.026	0.036	-0.025	0.036	-0.072**	-0.076	-0.071**	-0.085**	-0.105*	-0.085**
	(0.030)	(0.038)	(0.029)	(0.049)	(0.064)	(0.049)	(0.036)	(0.054)	(0.036)	(0.041)	(0.062)	(0.041)
High Intensity	-0.032	-0.058	-0.035	0.083*	0.020	0.083*	0.090**	0.085*	0.092**	0.089**	0.069	0.090**
	(0.029)	(0.040)	(0.029)	(0.048)	(0.067)	(0.049)	(0.036)	(0.045)	(0.036)	(0.041)	(0.049)	(0.040)
MNM treatment * High Intensity		0.052			0.124			0.010			0.041	
		(0.058)			(0.094)			(0.071)			(0.082)	
Received IFA during previous year			0.057			0.009			-0.024			-0.006
			(0.063)			(0.086)			(0.064)			(0.080)
N	412	412	412	410	410	410	404	404	404	404	404	404
R-squared	0.132	0.134	0.135	0.159	0.163	0.159	0.064	0.064	0.064	0.059	0.059	0.059

Notes: This table shows treatment effects on measures of how well the IFA program was implemented. We focus on four measures of IFA implementation quality: (i) whether HM shows enumerator IFA tablet, (ii) the number of tablets distributed per child in the past week (as seen in the school report), (iii) the percent of students who say they get the tablets weekly or more frequently (out of three that were asked), and (iv) whether at least 2 out of 3 students asked say they get the tablets at least weekly. All columns include block fixed effects and survey round fixed effects. While not always shown in the table, columns 2, 5, 8, 11, and 14always include the interaction term between the two treatments and columns 3, 6, 9, and 12 always include a control for whether the school received the IFA tablets during the previous schoolyear.

*Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 13: Treatment Effects on the Quality of Midday Meals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Was a meal served?			Were vegetables added to the meal?			Any children received 2nd helpings of ...						Number of adults served		
							Rice			Curry					
Panel A: All Months (5 visits each for High Intensity Schools, 3 visits each for Low Intensity Schools)															
MNM treatment	-0.021*	-0.036**	-0.021*	0.011	-0.009	0.011	0.016	0.021	0.016	0.008	0.006	0.008	0.141*	0.077	0.140*
	(0.012)	(0.017)	(0.011)	(0.008)	(0.009)	(0.008)	(0.045)	(0.065)	(0.045)	(0.035)	(0.054)	(0.035)	(0.083)	(0.163)	(0.082)
High Intensity	-0.012	-0.024**	-0.013	-0.016	-0.033**	-0.016	0.003	0.007	0.005	0.021	0.019	0.020	-0.215**	-0.267*	-0.223**
	(0.014)	(0.012)	(0.014)	(0.011)	(0.015)	(0.011)	(0.047)	(0.063)	(0.049)	(0.037)	(0.049)	(0.037)	(0.100)	(0.148)	(0.100)
MNM treatment * High Intensity		0.024			0.032**			-0.008			0.004			0.103	
		(0.023)			(0.014)			(0.090)			(0.070)			(0.186)	
Received IFA during previous year			0.026			0.000			-0.023			0.021			0.127
			(0.030)			(0.017)			(0.090)			(0.055)			(0.169)
N	581	581	581	568	568	568	565	565	565	561	561	561	567	567	567
R-squared	0.030	0.031	0.032	0.027	0.034	0.027	0.075	0.075	0.076	0.073	0.073	0.073	0.051	0.051	0.052
p-value of F-test (high & interaction)		0.123	.	.	0.065	.	.	0.993	.	.	0.849	.	.	0.100	.
Dep. var mean, control group	1.000	1.000	1.000	1.000	1.000	1.000	0.219	0.219	0.219	0.155	0.155	0.155	0.472	0.472	0.472
Panel B: Dec-Jan (only High Intensity Schools, 2 visits each)															
MNM treatment	0.025	0.025	0.025	.	.	.	-0.002	-0.002	-0.004	0.025	0.025	0.028	0.357*	0.357*	0.360*
	(0.017)	(0.017)	(0.017)	.	.	.	(0.084)	(0.084)	(0.083)	(0.072)	(0.072)	(0.070)	(0.185)	(0.185)	(0.186)
N	145	145	145	.	.	.	141	141	141	139	139	139	143	143	143
R-squared	0.072	0.072	0.073	.	.	.	0.121	0.121	0.124	0.103	0.103	0.115	0.099	0.099	0.102
Panel C: Feb-May (All Schools, 3 visits each)															
MNM treatment	-0.036***	-0.035**	-0.036***	0.014	-0.010	0.014	0.022	0.024	0.022	0.001	0.006	0.002	0.064	0.068	0.064
	(0.013)	(0.017)	(0.013)	(0.010)	(0.009)	(0.010)	(0.045)	(0.064)	(0.045)	(0.035)	(0.053)	(0.035)	(0.096)	(0.160)	(0.096)
High Intensity	-0.013	-0.012	-0.016	-0.017	-0.041**	-0.017	0.003	0.005	0.003	0.023	0.028	0.024	-0.199**	-0.196	-0.203**
	(0.014)	(0.010)	(0.014)	(0.011)	(0.019)	(0.012)	(0.047)	(0.064)	(0.048)	(0.036)	(0.050)	(0.036)	(0.099)	(0.150)	(0.099)
N	436	436	436	425	425	425	424	424	424	422	422	422	424	424	424
R-squared	0.039	0.039	0.046	0.030	0.042	0.030	0.049	0.049	0.049	0.047	0.047	0.047	0.067	0.067	0.067

Notes: This table presents treatment effects of the MNM introduction in schools and increased monitoring affected multiple measures of quality of the school meals. We consider several measures of meal quality based on variables measured in our MDM monitoring surveys. These include: (i) whether a a meal served, (ii) whether vegetables had been added to the meal, (iii) whether any children received 2nd helpings of rice, and of the non-rice dish (egg curry, or dalma, or soybean curry), (iii) the number of children served, and (iv) the number of adults served. 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, 8, 11, 14, and 17 always include the interaction term between the two treatments and columns 3, 6, 9, 12, 15, and 18 always include a control for whether the school received the IFA tablets during the previous schoolyear.

*Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 14: Overall effect of the IFA

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Children's hemoglobin levels					
IFA*Post	-0.047	0.053	0.199	0.280	0.283	0.307
	(0.141)	(0.165)	(0.150)	(0.178)	(0.202)	(0.202)
Deworming*Post	--	--	-0.345**	-0.329**	-0.210	-0.286
	--	--	(0.140)	(0.152)	(0.254)	(0.315)
IFA*Deworming*Post	--	--	--	--	-0.205	-0.066
	--	--	--	--	(0.304)	(0.354)
N	1459	1413	1459	1413	1459	1413
School fixed effects?	No	Yes	No	Yes	No	Yes
Added controls?	No	Yes	No	Yes	No	Yes

Note: The dependent variable is child's hemoglobin level measured in g/dL. IFA is a dummy variable that is one if a school reported recieiving IFA tablets and zero otherwise. All regressions include an indicator for whether hemoglobin measurement was taken after IFA implementation and the other relevant main effects of each interaction term. "Added controls" include the following variables interacted with "post": distance to block headquarters, whether or not a school has a kitchen, the percent of parents satisfied with MDM implementation, the percent of families employed in housework outside the home, and the percent of families in a non-disadvantaged caste. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.

Table 15: Heterogeneous effect of the IFA by anemia level at baseline

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Children's hemoglobin levels					
Panel A: Non-Anemic and Non-Borderline Anemic Students (Hb>=12.5 g/dL at baseline)						
IFA*Post	-0.387	-0.101	-0.118	0.141	0.214	0.192
	(0.257)	(0.284)	(0.284)	(0.368)	(0.362)	(0.379)
Deworming*Post	--	--	-0.356	-0.318	0.106	-0.221
	--	--	(0.284)	(0.347)	(0.435)	(0.788)
IFA*Deworming*Post	--	--	--	--	-0.743	-0.147
	--	--	--	--	(0.558)	(0.934)
N	196	186	196	186	196	186
Panel B: Non-Anemic Borderline-Anemic Students (11.5 <= Hb < 12.5 g/dL at baseline)						
IFA*Post	0.034	0.114	0.090	0.187	0.353	0.375
	(0.176)	(0.182)	(0.188)	(0.210)	(0.305)	(0.303)
Deworming*Post	--	--	-0.087	-0.115	0.214	0.117
	--	--	(0.190)	(0.193)	(0.250)	(0.280)
IFA*Deworming*Post	--	--	--	--	-0.518	-0.389
	--	--	--	--	(0.358)	(0.379)
N	420	410	420	410	420	410
Panel C: Mildly Anemic Students (11 <= Hb < 11.5 g/dL at baseline)						
IFA*Post	0.028	0.307*	0.243	0.499**	0.496*	0.842***
	(0.173)	(0.177)	(0.225)	(0.231)	(0.299)	(0.305)
Deworming*Post	--	--	-0.015	-0.432***	-0.011	0.225
	--	--	(0.032)	(0.142)	(0.064)	(0.300)
IFA*Deworming*Post	--	--	--	--	-0.532	-0.753*
	--	--	--	--	(0.440)	
N	280	272	280	272	280	272
Panel D: Moderately Anemic Students (8 <= Hb < 11 g/dL at baseline)						
IFA*Post	-0.003	0.066	0.218	0.277	0.180	0.248
	(0.169)	(0.185)	(0.242)	(0.243)	(0.326)	(0.317)
Deworming*Post	--	--	-0.288	-0.290	-0.371	-0.357
	--	--	(0.238)	(0.234)	(0.315)	(0.337)
IFA*Deworming*Post	--	--	--	--	0.114	0.093
	--	--	--	--	(0.436)	(0.452)
N	539	521	539	521	539	521
School fixed effects?	No	Yes	No	Yes	No	Yes
Added controls?	No	Yes	No	Yes	No	Yes

*Note: Receiving IFA is a dummy variable that is one if a school reported recieiving IFA tablets and zero otherwise. The dependent variable is a child's hemoglobin levels, measured in g/dL. All regressions include an indicator for whether hemoglobin measurement was taken after IFA implementation and the other relevant main effects of each interaction term. Anemia levels are defined by the WHO standards at sea level. "Added controls" include the following school-level variables interacted with 'post': distance to block headquarters, the percent of parents satisfied with MDM, whether a school has a kitchen, the percent of families employed in housework outside the home, and the percent of students in a non-disadvantaged caste. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 16: Heterogeneous effect of the IFA students receiving tablets more recently

	(1)	(2)	(3)	(4)
	Dependent variable: Children's hemoglobin levels			
IFA*Post	0.006	0.128	0.303	0.414
	(0.206)	(0.226)	(0.234)	(0.254)
IFA*Post*UncertainTabletStatus	0.130	0.045	0.091	0.008
	(0.196)	(0.201)	(0.198)	(0.205)
IFA*Post*RanOutOfTablets	-0.288	-0.259	-0.334*	-0.311
	(0.201)	(0.196)	(0.198)	(0.197)
Deworming*Post	--	--	-0.370**	-0.363**
	--	--	(0.154)	(0.159)
N	1459	1413	1459	1413
P-value (F-test of 3 IFA*Post coefficients):	0.059	0.210	0.020	0.058
P-value (IFA*Post + IFA*Post*RanOut =0)	0.091	0.470	0.861	0.595
School fixed effects?	No	Yes	No	Yes
Added controls?	No	Yes	No	Yes

*Note: The dependent variable is child's hemoglobin level measured in g/dL. IFA is a dummy variable that is one if a school reported recieiving IFA tablets and zero otherwise. Uncertain Tablet Status is a dummy variable that is a one if the school reported not knowing if they had run out of tablets, and Ran Out of Tablets is a dummy variable that is a one if the school reported running out of tablets. All regressions include an indicator for whether hemoglobin measurement was taken after IFA implementation and the other relevant main effects of each interaction term. "Added controls" include the following variables interacted with "post": distance to block headquarters, whether or not a school has a kitchen, the percent of parents satisfied with MDM implementation, the percent of families employed in housework outside the home, and the percent of families in a non-disadvantaged caste. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 17: Analysis of differential attrition

Dep variable:	Attrited From Baseline I						Attrited From Baseline I or Baseline II					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MNM treatment	-0.016	-0.055**	0.267	0.040	0.005	-0.006	-0.004	-0.028	0.198	0.350	0.004	-0.002
	(0.016)	(0.022)	(0.193)	(0.317)	(0.044)	(0.068)	(0.019)	(0.027)	(0.201)	(0.324)	(0.040)	(0.060)
High Intensity	-0.011	-0.049**	0.122	-0.093	-0.040	-0.050	-0.021	-0.044*	-0.161	-0.006	-0.037	-0.042
	(0.016)	(0.023)	(0.192)	(0.328)	(0.044)	(0.067)	(0.019)	(0.026)	(0.201)	(0.336)	(0.040)	(0.058)
MNM treatment * High Intensity		0.076**		0.409		0.019		0.047		-0.301		0.009
		(0.031)		(0.388)		(0.089)		(0.038)		(0.399)		(0.081)
Baseline hemoglobin level			0.022	0.012					0.017	0.026		
			(0.020)	(0.027)					(0.021)	(0.027)		
MNM treatment * baseline hemoglobin level			-0.012	0.004					0.013	-0.003		
			(0.017)	(0.030)					(0.018)	(0.031)		
High Intensity * baseline hemoglobin level			-0.026	-0.008					-0.018	-0.034		
			(0.017)	(0.029)					(0.018)	(0.029)		
MNM treatment * High Intensity				-0.031						0.031		
* baseline hemoglobin level				(0.035)						(0.036)		
Class in school as of baseline					0.004	0.010					0.004	0.012
					(0.020)	(0.024)					(0.018)	(0.021)
MNM treatment * class at baseline					0.016	0.003					0.016	0.001
					(0.021)	(0.030)					(0.019)	(0.027)
High Intensity * class at baseline					-0.010	-0.023					-0.009	-0.023
					(0.021)	(0.031)					(0.019)	(0.029)
MNM treatment * High Intensity						0.027						0.031
* class at baseline						(0.041)						(0.039)
Constant	0.103***	0.123***	-0.135	-0.011	0.093**	0.100*	0.154***	0.166***	-0.034	-0.123	0.086**	0.089**
	(0.015)	(0.018)	(0.227)	(0.302)	(0.044)	(0.054)	(0.016)	(0.018)	(0.228)	(0.296)	(0.037)	(0.045)
N	1241	1241	1232	1232	1239	1239	1481	1481	1471	1471	1380	1380
R-squared	0.001	0.006	0.004	0.009	0.002	0.007	0.001	0.002	0.005	0.006	0.002	0.006
p-value from F-test of all regressors	0.544	0.074	0.461	0.144	0.757	0.340	0.545	0.401	0.241	0.233	0.722	0.305

Notes: This table presents an analysis of attrition for our main outcome variables on child health. Columns 1-6 study attrition from Baseline I while Columns 7-12 study attrition from either Baseline I or Baseline II. While column 2 suggests differential attrition by treatment arm, 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, class at baseline does not affect probability of attrition, and this does not differ by treatment group (Column 5-6, 11-12).

*Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table A1: Comparison of schools that ran out of tablets to those that didn't

		Blocks with high IFA variation			
	Panel A: Demographic Characteristics	Didn't Run Out	Ran Out	Uncertain	P-Value
School-reported (SCH)	Distance to the block headquarters (km)	18.21	18.56	23.49	**0.0132**
	Primary enrollment	77.48	88.92	62.17	**0.0115**
	Secondary enrollment	23.48	50.97	15.43	**0.0005**
	Number of teachers	2.56	3.15	2.19	**0.0270**
	Percent of schools have a kitchen	0.76	0.82	0.74	0.7063
	Percent of schools have at least one latrine	0.92	0.84	0.85	0.6325
	Percent of schools have sufficient water	0.83	0.79	0.70	0.4480
Household-reported, aggregated to school level (HH)	Mean % of students are female	0.51	0.52	0.49	0.2814
	Mean % of families in a non-disadvantaged caste	0.03	0.05	0.02	0.2927
	Mean % of village adults in agricultural work	0.15	0.19	0.23	**0.0020**
	Mean % of village adults work in own home	0.26	0.27	0.26	0.8171
	Mean % of village adults work in others' homes	0.25	0.24	0.25	0.9121
	Mean % of village adults work as laborers	0.19	0.16	0.13	0.1014
	Mean % of village adults with no formal schooling	0.64	0.55	0.54	**0.0442**
	Mean % of village adults who own a phone	0.28	0.30	0.32	0.6020
	Mean % of families that live in high-quality housing	0.09	0.08	0.09	0.9090
	Mean % of families that live in low-quality housing	0.80	0.79	0.73	0.1441
	Mean % of families with electricity	0.50	0.54	0.52	0.8735
	Panel B: Implementer Variables				
SCH	Percent with parent group for MDM	0.00	0.17	0.10	0.1324
	Percent with MDM training	0.50	0.63	0.65	0.4478
	Percent receiving MDM rice on a regular schedule	0.58	0.22	0.35	**0.0129**
HH	Mean number of MDM per week	4.72	4.87	4.81	0.5724
	Mean % of parents satisfied with MDM	0.92	0.90	0.89	0.3251
	Panel C: Anthropometric Measures at Baseline				
Child-level measures aggregated to school level and averaged	Mean % of students with anemia	0.59	0.54	0.57	0.3850
	Mean % with mild anemia	0.24	0.23	0.23	0.8318
	Mean % with moderate anemia	0.34	0.29	0.33	0.3065
	Mean % with severe anemia	0.01	0.01	0.01	0.7241
	Mean child Hb level	11.17	11.31	11.15	0.1882
	Mean student BMI	13.60	13.54	13.78	0.5894
	Mean BMI, anemic students	13.52	13.59	13.68	0.7782
	Mean BMI, nonanemic students	13.62	13.55	13.95	0.4484
	Mean student weight	18.26	18.32	18.38	0.9448
	Mean weight, anemic students	17.36	18.10	17.71	0.3497
	Mean weight, nonanemic students	19.35	19.16	19.35	0.9162
	Mean BMI, girls	13.43	13.47	13.59	0.8722
	Mean BMI, boys	13.74	13.63	13.89	0.4017

Note: P-value corresponds to F-test of the null hypothesis that the three means are the same.

Table A2: Comparison of Deworming, non-Deworming schools in high-variation blocks

		Blocks with high IFA variation		
	Panel A: Demographic Characteristics	Got Deworming	No Deworming	P-Value
School-reported (SCH)	Distance to the block headquarters (km)	20.96	21.98	0.5166
	Primary enrollment	72.60	80.34	0.3015
	Secondary enrollment	29.29	31.66	0.7406
	Number of teachers	2.65	2.28	0.1923
	Percent of schools have a kitchen	0.78	0.63	**0.0472**
	Percent of schools have at least one latrine	0.84	0.87	0.7339
	Percent of schools have sufficient water	0.79	0.65	**0.0526**
Household-reported, aggregated to school level (HH)	Mean % of students are female	0.50	0.50	0.7213
	Mean % of families in a non-disadvantaged caste	0.04	0.05	0.3280
	Mean % of village adults in agricultural work	0.20	0.17	**0.0361**
	Mean % of village adults work in own home	0.27	0.22	**0.0012**
	Mean % of village adults work in others' homes	0.24	0.30	**0.0021**
	Mean % of village adults work as laborers	0.15	0.20	**0.0064**
	Mean % of village adults with no formal schooling	0.55	0.62	**0.0293**
	Mean % of village adults who own a phone	0.32	0.30	0.7103
	Mean % of families that live in high-quality housing	0.09	0.10	0.8925
	Mean % of families with electricity	0.53	0.50	0.5995
	Panel B: Implementer Variables			
SCH	Percent with parent group for MDM	0.10	0.16	0.3031
	Percent with MDM training	0.63	0.48	**0.0747**
	Percent receiving MDM rice on a regular schedule	0.35	0.43	0.3116
HH	Mean number of MDM per week	4.80	4.68	0.2889
	Mean % of parents satisfied with MDM	0.90	0.90	0.6906
	Panel C: Anthropometric Measures at Baseline			
Child-level measures aggregated to school level and averaged	Mean % of students with anemia	0.56	0.59	0.2095
	Mean % with mild anemia	0.22	0.24	0.2567
	Mean % with moderate anemia	0.32	0.34	0.5464
	Mean % with severe anemia	0.01	0.01	0.7747
	Mean child Hb level	11.23	11.15	0.2898
	Mean student BMI	13.70	13.64	0.7182
	Mean student weight	18.31	18.29	0.9313
	Mean BMI, girls	13.56	13.50	0.7534
	Mean BMI, boys	13.80	13.81	0.9343

Note: P-value tests the difference in the two means, unconditional on block. Bolded p-values are significant at the 10% level.

Table A3: Overall effect of the IFA on height and weight

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:	Dependent variable: Children's height					
IFA*Post	0.705	0.620	-0.622	-0.868	-1.902	-2.190
	(0.813)	(0.832)	(1.465)	(1.461)	(2.444)	(2.529)
Deworming*Post	--	--	1.860	2.152	0.010	0.106
	--	--	(1.626)	(1.751)	(0.811)	(1.195)
IFA*Deworming*Post	--	--	--	--	2.921	3.178
	--	--	--	--	(2.605)	(2.966)
N	1460	1414	1460	1414	1460	1414
Panel B:	Dependent variable: Children's weight					
IFA*Post	0.564***	0.528***	0.987***	0.980***	1.359***	1.427***
	(0.174)	(0.201)	(0.310)	(0.331)	(0.497)	(0.530)
Deworming*Post	--	--	-0.594*	-0.655*	-0.046	0.041
	--	--	(0.328)	(0.351)	(0.281)	(0.299)
IFA*Deworming*Post	--	--	--	--	-0.859	-1.078*
	--	--	--	--	(0.556)	(0.619)
N	1462	1416	1462	1416	1462	1416
School fixed effects?	No	Yes	No	Yes	No	Yes
Added controls?	No	Yes	No	Yes	No	Yes

*Note: The dependent variable is child's height measured in cm (panel A) and child's weight measured in kg (panel B). IFA is a dummy variable that is one if a school reported recieiving IFA tablets and zero otherwise. All regressions include an indicator for whether hemoglobin measurement was taken after IFA implementation and the other relevant main effects of each interaction term. "Added controls" include the following variables interacted with "post": distance to block headquarters, whether or not a school has a kitchen, the percent of parents satisfied with MDM implementation, the percent of families employed in housework outside the home, and the percent of families in a non-disadvantaged caste. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*