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

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

July 2021

ABSTRACT

Governments often rely on schools to implement multiple programs targeting child outcomes. How to improve the implementation of these programs is an important, open question. As part of a randomized controlled trial in Odisha, India, we measured the impacts of a nutrition program and a monitoring intervention on the implementation of a pre-existing school-based nutrition program, specifically the Indian government's iron and folic acid supplementation (IFA) program. The new nutrition intervention distributed a micronutrient mix (MNM) to be added to school meals while the monitoring intervention varied the intensity of monitoring activities. We find that high intensity monitoring improved implementation of the government's IFA program, while the MNM intervention crowded it out. The net effect is that high intensity monitoring improved child health, while the MNM intervention did not. Both crowd-out of the IFA program and sensitivity to monitoring were predominantly found among schools that were resource or capacity constrained.

JEL Codes: O15, I18, H40

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

---

[1] We thank The International Initiative for Impact Evaluation (3ie), the Douglas B. Marshall, Jr. Family Foundation, the International Growth Center, Wellesley College Faculty Research Awards, and the Jerome A. Schiff Fellowship 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. We are grateful to Baafra Abeberese, Elizabeth Cascio, and seminar participants at Wellesley College, the Liberal Arts Colleges Development Economics Conference, Advances with Field Experiments Conference, the Eastern Economics Association meetings, the Indian Statistical Institute, and the North East Universities Development Consortium Conference for comments and suggestions.
James Berry: University of Delaware, jimberry@udel.edu; Saurabh Mehta: Cornell University, smehta@cornell.edu; Priya Mukherjee: University of Wisconsin-Madison, priya.mukherjee@wisc.edu; Hannah Ruebeck: MIT, hruebeck@mit.edu; Corresponding author: Gauri Kartini Shastry, Wellesley College, 106 Central Street, Wellesley, MA 02481, gshastry@wellesley.edu.

I. Introduction

Governments around the world rely on school infrastructure to implement programs aimed at improving child welfare, some of which are only indirectly related to education. For example, schools in many countries are required to provide students with nutritious meals through school feeding programs. Schools around the world also distribute micronutrient supplementation, deworming treatment, and immunization (Bundy et al. 2018). When schools are tasked with implementing additional programs, it is important to understand how these programs interact with one another. On the one hand, there may be complementarities across programs: for example, micronutrient supplementation may be easier to implement if done alongside school feeding programs (Best et al., 2011). On the other hand, these programs may crowd each other out: schools may not have sufficient capacity to implement additional programs, and new programs could inhibit the implementation of existing programs or interfere with school activities. Crowd-out is generally difficult to study because most evaluations focus on the intervention being evaluated, understandably, with less attention paid to other activities.

We examine interactions between school-implemented programs in the context of India's midday meals scheme and iron and folic acid (IFA) supplementation program, which provided students with weekly iron tablets and biannual deworming tablets. We take advantage of rich data on the implementation of the pre-existing IFA program during the randomized evaluation of a newly implemented program, conducted across 148 schools in Kendujhar district, in the state of Odisha. The new intervention provided schools with a micronutrient mix (MNM) to be added to school meals. Designed to complement the government's IFA program, the mix was intended to improve absorption of iron from the supplements. Implementation of both programs was to be supervised by school headmasters, making this an ideal context in which to study administrative constraints. We also varied the intensity with which school meals were monitored in order to further understand factors influencing program implementation. The MNM intervention was cross-randomized with high-intensity monitoring at the school level, allowing us to examine how monitoring affected existing program implementation on its own and when combined with the MNM intervention. Our analysis includes school-level outcomes such as nutrient content of meals and fidelity of IFA implementation, as well as child-level health outcomes for a sample of 1,920 children.

We first document relatively consistent implementation of the MNM intervention, as measured by the nutrient content of school meals. However, we find that the MNM intervention caused a reduction in IFA implementation – students in the MNM schools were *less* likely to report receiving IFA tablets regularly. On the other hand, while high intensity monitoring did not impact MNM take-up, it did improve IFA program implementation: students in the high intensity treatment arm were *more* likely to report receiving IFA tablets regularly. The combined high-intensity monitoring and MNM arm saw no significant change in IFA implementation, suggesting that the *crowd-in* effect of the additional monitoring cancelled out the *crowd-out* effect of the MNM program.

These results pose interesting puzzles, first regarding why the MNM program crowded out IFA implementation and second regarding why the high intensity monitoring crowded in IFA implementation but not MNM implementation. We first verify our results with an objective outcome, child hemoglobin levels, which is also the final outcome of interest. Despite serving meals with higher nutrient content, we find small, statistically insignificant, and often negative impacts of the MNM intervention on children's hemoglobin levels. This result can be explained by our finding that the MNM program crowded out IFA distribution. By contrast, increased monitoring of school meals did increase hemoglobin levels, again consistent with improved implementation of the IFA program. We also see a marginally significant improvement in child hemoglobin levels in schools that received both the MNM and the high intensity monitoring as compared to the *control* group, even though there was no significant difference in IFA implementation. This result suggests that the more nutritious meals in the MNM schools would have improved child health had it not been for the crowd-out of IFA distribution.[2]

Both crowd-out and crowd-in of IFA distribution suggest that there exist binding constraints affecting IFA implementation, for example, in time or effort exerted by school officials. To study this further and help reconcile these puzzles, we exploit the fact that IFA implementation becomes more challenging as the number of students rises, conditional on the number of staff. Each child

[2] It is possible that the MNM did not improve child health because of insufficient dosage. We had initially intended to provide 100% of the recommended daily allowance (RDA) for this age group, but the National Institute of Nutrition in India requested a dosage of approximately 50% of RDA, and that we add calcium to the mix (calcium may inhibit the absorption of iron, although usually at higher doses than in our mix). However, the improvement in child health in the schools that received both treatments suggests that even 50% of RDA can improve child health when combined with the IFA program.

must be given exactly one tablet once a week, and supervising the distribution involves ensuring that each child swallows the tablet with food and water to avoid discomfort (UNICEF 2017). We find that IFA tablet receipt responds more to the MNM program and high intensity monitoring in schools with higher than average students per staff member, that is, as IFA implementation gets more challenging. Meanwhile, adding MNM scoops to the meal does not get more challenging as the number of students rises. We find that the nutrient content of the meals, an indicator of MNM take-up, is not sensitive to the student to staff ratio, suggesting MNM program implementation is not subject to these constraints in the same way.

We conjecture that resource or administrative capacity constraints are more binding for IFA implementation than MNM implementation, and that the introduction of a new program exacerbates these constraints while the increased monitoring nudges headmasters to find ways to relax these constraints. We test this hypothesis by constructing an index of capacity to implement these programs and examining heterogeneity across schools. The capacity index includes measures of resource constraints such as staffing and meal funding, as well as organizational support from the community. We also include measures that indicate the administration's implementation of other health interventions such as ensuring that the school's water supply is treated to remove pathogens. As with staffing constraints, we find that both crowd-out of the IFA program by MNM and crowd-in of the IFA by monitoring are predominantly found among schools that appear more constrained, while MNM implementation is not responsive to these constraints. Taken together, these results highlight the need for policies to take into account capacity constraints when recommending how and whether to implement additional programs in schools.

Another question generated by our results is precisely how the extra monitoring visits affected headmaster behavior. High intensity schools received monthly monitoring visits starting from the beginning of the intervention, while other schools received these visits starting in the third month. Consistent with this timing, we show that the difference in headmaster behavior is greatest in the middle of the intervention when the difference in monitoring is greatest. Using additional variation in the number and timing of visits, we present evidence consistent with two possible explanations; the intervention may have a) served as more frequent reminders or b) induced compliance in anticipation of future monitoring visits.

This paper relates to a number of literatures. First, and most centrally, we contribute to the literature on how implementation of existing publicly-provided programs can be crowded out when new programs are introduced. Existing literature has uncovered a number of factors that lead to crowd-out. As our findings suggest, crowd-out can occur when those tasked with implementation face time or other capacity constraints. This is an important dimension of public service delivery: how those tasked with public good provision implement multiple programs simultaneously, i.e. how they multi-task (see Finan, et al., 2017, for a review). Within the context of school meals, Vermeersch and Kremer (2005) find that school meals reduced instruction time, which they attribute to poor time management by cooks and school staff. Alternatively, a publicly-provided program can crowd-out similar activities if agents believe the new program suffices. Das et al. (2013) find that the beneficial effects of increased school inputs were offset by reductions in educational spending at home. Muralidharan and Sundararaman (2013) find that teacher attendance and teaching activity fell when an extra teacher was introduced at the school. Crowd-out can also occur across organizations if they face a constrained supply of qualified workers (Deserranno et al. 2021).

Second, our paper relates specifically to prior work on monitoring service providers, reviewed in Finan et al. (2017). Despite not having explicit stakes, our monitoring most closely resembles top-down administrative monitoring. The literature generally finds positive impacts of such monitoring in diverse public services (see, for example, Duflo et al., 2012, and Muralidharan et al., 2017, in education, Dhaliwal and Hanna, 2017, in health care delivery, and Olken, 2007, in road construction); but, as Finan et al. (2017) note, various features of the monitoring scheme play an important role, such as the officials' ability to circumvent the system (Banerjee et al. 2008, Dhaliwal and Hanna 2017) and the incentive structures (and personality) of those in charge of the monitors (Callen et al. 2015, Duflo et al. 2013).[3] As we discuss in Section IV.D, the features of our monitoring design suggest that the visits may have acted as reminders to the headmasters to implement the IFA program or created accountability by making them expect further scrutiny. Consistent with the latter explanation, Muralidharan et al. (2021) use an at-scale experiment to

[3] Evidence on citizen monitoring is more mixed, likely due to varying ability of citizen groups to collect information and hold government officials accountable (Banerjee et al. 2010, Banerjee et al 2018, Pradhan et al. 2014, Björkman and Svensson 2009, Björkman, de Walque, and Svensson 2017).

show that simply announcing to officials that beneficiaries would be called and surveyed on program implementation improved service delivery.[4]

Finally, our paper is related to the literature on micronutrient supplementation and school-based nutrition programs. Like the IFA program, many of these interventions provide iron.[5] In the context of India's midday meal program, Krämer, Kumar, and Vollmer (2020) evaluate the provision of double-fortified salt — salt fortified with iron and iodine — to schools to be used in midday meals in the Indian state of Bihar. The authors find that the intervention increased hemoglobin levels. In Berry et al. (2020), we evaluate the impact of the IFA program in the year before our randomized evaluation using variation in tablet receipt across schools, finding that improvements in child health are limited by breaks in supplementation such as summer holidays. The MNM did not itself contain iron, but it was designed to help children absorb the iron supplements they were receiving in school through the IFA program and thereby lower iron-deficiency anemia. Our finding of crowd-out suggests that school-based interventions that complement one another in carefully controlled settings may not be successful if capacity constraints hinder the simultaneous implementation of all programs.

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

II. Description of nutrition programs

A. India's school-based health interventions

Context: The midday meal program

The programs we study take place in the context of India's midday meal program, which is designed to provide all public schoolchildren in grades 1 through 8 with nutritious cooked meals in schools. Our study took place in the eastern state of Odisha, in the rural district of Kendujhar. During the period of study, the midday meal program in Odisha was supervised by the state

[4] Closely related to our context and exploiting mobile-based technology, Debnath, Nilayamgode, and Sekhri (2020) find that monitoring reduces leakages in school meals in Bihar, India.

[5] See, for example, Hyder et al. (2007); Hirve et al. (2007); Ahmed et al. (2010); Tee et al. (1999); Gera et al. (2007); and Abrams et al. (2008).

Department of School and Mass Education, in coordination with district-, block- (administrative unit smaller than a district) and cluster- (administrative unit smaller than block) level officials. In most schools in our sample, either the headmaster or one or more of the teachers was responsible for purchasing food materials, obtaining cooking fuel and hiring and supervising cooks. In 43 percent of the schools in our sample, members of self-help groups (SHGs) assisted the school staff in acquiring ingredients and cooking the meals.[6] More members of the teaching staff typically helped during lunch to organize the seating of students, distribution of meals, and washing of utensils before classes resumed. While the district was supposed to train those responsible for providing the meals, in only 33 percent of our schools had anyone ever attended a training related to the midday meal program. Anecdotal evidence indicates that the headmasters and teachers considered running the midday meal program to be an administrative burden; this was one of the most common concerns reported by school officials during our field visits and focus group discussions.

Iron and folic acid supplementation program

In 2012, India's Ministry of Health and Family Welfare introduced a national iron supplementation program to reduce the prevalence and severity of anemia among school children. Beginning in January 2013, according to the guidelines distributed by the central government, iron and folic acid supplements, as well as deworming medication, were to be distributed free of charge to all students attending public schools. During the year of our intervention, children aged 5-10 were to receive 45 mg of elemental iron and 400 μg of folic acid once a week at school.[7] One tablet of deworming medication, Albendazole, was also to be administered to each child every six months.

In the first year of the IFA program (the year *before* this paper's study period), approximately 86% of schools we surveyed received iron and folic acid tablets, ranging from 49% of schools in one block to 99% in another. As described in detail in a companion paper, the process by which tablets were distributed to schools relied on government officials at many administrative levels (Berry et al. 2020). In that paper, we use variation in whether schools *received* the tablets to estimate the

[6] Self-help groups consist of local women (and sometimes men) who save and lend among members of the group.
[7] In the first year of the IFA program, the year *prior* to our intervention, the program was implemented differently. Children aged 5-10 were to be given 30 mg of elemental iron and 250 μg of folic acid daily for a duration of 100 days.

impact of the IFA program in its first year.[8] By the start of the MNM intervention, however, virtually all schools in our sample had received the tablets; in this paper, we study IFA implementation by schools, conditional on tablet receipt. Note that implementation of the IFA program affects the interpretation of the MNM intervention: the comparison is between schools that received both iron supplements and multi-micronutrient fortification and schools that only received iron supplements.

Within schools, implementation of the IFA program was the headmasters' responsibility. Headmasters received the iron tablets from block-level officials and supervised the provision at school. Official guidelines instructed headmasters to assign a teacher to be in charge of IFA program implementation; the teacher was to keep a ledger of supply and distribution and was responsible for distributing the IFA tablets to students with the assistance of student prefects. More than 90% of headmasters submitted reports on the IFA implementation to higher level officials. Recall that tablets were to be distributed once a week, requiring officials to maintain a weekly schedule, possibly disrupted by mid-week holidays. In addition, official guidelines emphasized swallowing the IFA tablets with food and water to minimize any side effects, implying that school officials were to oversee the actual ingestion of the tablets (UNICEF 2017; Government of Odisha 2012). Students were also supposed to receive tablets to take home for use over school vacations, and the program guidelines encouraged teachers to set an example by taking the tablets. While the central and state governments intended to monitor compliance with the IFA program intensively, monitoring activities generally involved comparing administrative records with school records.

B. Experimental interventions

The micronutrient mix program

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

[8] Using our survey data, we document that schools in the two blocks with more consistent receipt of IFA tablets serve a more advantaged population than schools in the remaining three blocks. However, the variation within each block appears random; we found no evidence of a relationship between tablet receipt and school-level variables or student demographics. We use this quasi-random variation to estimate the impact of the IFA program (Berry et al. 2020).

The district government provided a letter instructing headmasters to participate in the program and cooperate with the research team's surveys.

Note that the MNM we provided did not include iron or folic acid. Rather, the MNM was intended to help children absorb iron, complementing the IFA distribution. The intervention was motivated by the nutrition literature demonstrating that multi-micronutrient supplementation is more effective in combating anemia than iron and folic acid supplementation alone (Ramakrishnan et al. 2004, Best et al. 2011, Ahmed et al. 2010).[9] We restricted the dosage of the vitamins and minerals in the MNM to approximately 50% of the recommended daily allowance (RDA) at the request of the NIN.

In order to implement the intervention, we first trained headmasters, cooks, and other staff involved with meal preparation. During these trainings, we covered the health consequences of anemia and other forms of malnutrition, health benefits of consuming the various vitamins and minerals in the MNM, and directions for MNM use. We were very clear that the MNM was intended to help children absorb the iron from the IFA tablets and did not itself contain iron. Instructions for MNM use were provided to all participants, since teachers and headmasters were required to help cooks calculate the amount of the mix to be added each day. We distributed the mix, plastic sealable jars, and scoops that held 10 grams of the mix, which the headmaster kept in the school's storage room. The dosage approved by the National Institute of Nutrition meant that children were to receive 1.5 g of the MNM each day. Since the mix was to be added to the food before it was served to the children, it was necessary to multiply the number of children eating the meal by 1.5 to calculate the number of grams of the mix to add, and then divide this number by 10 to calculate the number of scoops to add. During the training it became clear that cooks were usually not confident about performing this calculation but headmasters helped them with the arithmetic. We also gave schools laminated fliers that clearly described the steps necessary to add the MNM to the food (see Appendix Figure 1). Every month, we contacted headmasters to enquire whether they needed more of the MNM and, if so, delivered additional packets to the school.

[9] Fawzi et al. (2007) and Mehta et al. (2011) find that multi-micronutrient supplementation even without iron and folic acid can improve hemoglobin levels, although we should note that these studies were conducted in otherwise sick populations with dosages that were greater than the RDA.

High intensity monitoring

The second intervention involved earlier and more consistent monitoring of school meals during the study period. Schools in the high intensity monitoring treatment group were visited during meal time on a random day once per month starting from the first month of the five-month intervention; schools in the low intensity monitoring group received these visits starting in month three. These monitoring visits were fairly intrusive, involving detailed observations of meal quality, child attendance, the distribution of food items and quantities to the children, and the amount of food consumed. In addition, enumerators 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.

III. Experimental design

A. Timeline

Figure 1 presents the chronology of key activities for the study. The original design was to fortify the school meals with iron. Three hundred seventy-five schools were selected for the study, and an initial baseline survey (Baseline 1) was conducted in these schools between September 2012 and January 2013. However, the plan was halted when the government's IFA program was announced, and the study was revised to evaluate MNM, monitoring, and their interactions with the IFA program. Changing the intervention plan required securing approvals for the new design from a number of government agencies and took approximately 16 months, with final approvals received at the end of September 2014. While waiting for final approval, we conducted a survey measuring the intensity of IFA implementation. We then conducted a second baseline survey (Baseline 2) in a subset of sample schools during August and September of 2014, early in the 2014-2015 school year. Baseline 2 focused on the three administrative blocks (157 schools) with variation in IFA implementation in the first year in order to evaluate the impact of the government's IFA program on child health (see Berry et al. 2020).

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

B. School sample

The sample schools in Kendujhar district were selected for the study based on whether they satisfied the following conditions: (i) the school was located within 50 kms from the town of Kendujhar, the capital of the district, and (ii) the school was located in one of five blocks: Banspal, Ghatagaon, Jhumpura, Sadar, or Patna. This minimized the fixed costs of dealing with government officials in charge of schools in each block. We began with a sample of 377 primary schools that satisfied these conditions and randomly selected 150 schools in which to conduct the MNM and high intensity monitoring treatments.[10] These schools are primarily rural with a high fraction of students from tribal or scheduled caste communities (approximately 95%). Households are relatively poor – 50% have electricity, 30% own a phone, and 50% of household heads are literate – and children are relatively unhealthy – 44% are underweight and 60% are anemic. In terms of child health, the sample is fairly representative of the state of Odisha, in which 41 percent of children under the age of five are underweight and 65 percent are anemic (International Institute for Population Sciences 2007).

C. Treatment assignment

Out of the 150 schools in the sample, 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.

Table 1 provides the number of schools and students in each group. While the original sample contained 150 schools, 2 schools refused to participate from the beginning of the study, before their treatment status was revealed. Thus, we were left with 75 MNM treatment schools and 73 comparison schools. Out of the 75 schools in the MNM treatment group, 37 were monitored

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

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.

D. Data

We collected data on a number of outcome variables at various points during the study. Our school-level data includes information on the quality of midday meals, take-up of the micronutrient mix program (including the quantity of vitamin A and zinc in food samples), and the implementation of the IFA program. We also collected child-level data including household demographics; hemoglobin levels; anthropometric measures such as height, weight, and mid-upper arm circumference; cognition; school attendance; and test scores.[11] We describe each survey and the variables of interest below. Figure 1 indicates the timing of these survey modules.

School-level data

Our school-level data collection began with a baseline survey conducted in late 2012 and early 2013. This survey measured school characteristics and teacher demographic details and qualifications. This baseline information was updated during the first two months of the intervention in late 2014 and early 2015. We carefully monitored take-up of the MNM program throughout the intervention period. One measure of take-up is the amount of MNM each school used. We calculate this as the amount of MNM received minus the amount that remained at the end of April 2015, relative to the amount we estimated they would need to serve their students.

In addition, trained enumerators made surprise visits to the study schools to observe the quantity and quality of school meals. Schools in the high intensity monitoring arm received these visits every month of the intervention (5 visits total). Schools in the low intensity treatment arm received these visits during the third, fourth and fifth months. As described above, these visits were fairly intensive. During the third and fifth months of the intervention, enumerators also collected samples of the meal and sent these samples to a laboratory for nutritional analysis. We have data on the amount of vitamin A and zinc in the food sample.[12]

---

[11] We did not directly measure children's blood levels of the nutrients in the MNM for budgetary reasons.
[12] We chose to test only vitamin A and zinc for budgetary reasons and because pilot tests of samples of fortified food cooked by our research team and sent to the lab provided the most consistent results for these micronutrients.

We also carefully monitored implementation of the IFA program. During March-April 2014 (the school year before our intervention) 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 they were being distributed to children. After speaking to the headmaster at each school about the IFA program, our enumerators randomly selected three children to answer additional questions on whether they had received IFA tablets. One student was randomly chosen from each of grades 2, 4, and 5. For each school, we calculate the fraction of those three children that reported receiving tablets regularly.[13]

Child-level data

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

As described in Section III.A, we conducted a second baseline (Baseline 2) at the beginning of the 2014-2015 school year, in about half of the schools in the sample. This survey was conducted with 9 children per school. Appendix A describes the sampling procedure for students in Baseline 2.[15]

[13] Students were also asked if they swallowed the tablets they received. Almost all students responded that they had. In focus group discussions, headmasters and teachers mentioned that children do sometimes discard the tablets, but none of the children surveyed reported doing that.

[14] Only 19 schools out of the 148 in our sample have hostels. We confirm that all our results are robust to excluding these schools entirely.

[15] At Baseline 2, we only surveyed children in the 3 administrative blocks with variation in IFA implementation. Data from Baseline 2 is primarily used in the companion paper evaluating the IFA program (Berry et al. 2020), but we control for these updated baseline hemoglobin measures in some specifications below.

After obtaining parental consent, enumerators visited schools to measure the selected children's height, weight, and hemoglobin levels during the Baseline 1, 2, and Endline surveys. School attendance data were also collected once per month in each of the last three months of the intervention through random, unannounced visits.

Finally, we conducted household surveys at Baseline 1 and Endline to collect demographic information, household assets, knowledge of anemia, and perceptions of the school's midday meal.

E. Summary statistics and balance

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

Table 3 checks balance on child health and demographics across each of the experimental treatment groups. Panels A and B focus on child health before the intervention; Panel A includes children who were in the sample at Baseline 1, while Panel B includes children surveyed at Baseline 2. Panels C-F focus on demographics; for children added to the sample during the 2014-15 school year, we fill in demographic information collected at the endline survey if the variable is most likely time-invariant or unrelated to treatment (for example, we fill in the variable for mother's education but not whether the child takes any supplements). Appendix Figure 2 plots the distribution of baseline hemoglobin levels for each of the four treatment groups. The groups are well balanced on most variables, including child health outcomes, household characteristics and demographic information on children, mothers, and heads of household, with a slight imbalance on a few of the 35 variables in the table across the four groups. The number of significant differences across all four groups (3 out of 35 at the 10 percent level, 2 at the 5 percent level) is about what would be expected through random chance.

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

IV. Results

A. MNM take-up

Our first outcome of interest is take-up of the MNM by schools in the MNM treatment group. We begin with intervention 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. This data is available only for the schools in the MNM treatment arm. We estimate

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

where y_{sb} is a measure of take-up in school s in block b and $High_{sb}$ is a dummy variable for schools that received higher frequency monitoring visits. The results are reported in Table 4. In addition to block fixed effects, α_b, these regressions control for the number of children enrolled in the school at the start of the intervention. Schools assigned to the MNM treatment did take up the mix. The schools that were not monitored intensely received 2.8 deliveries during the study period (the dependent variable mean, presented at the bottom of Columns 1-2). As shown in Columns 3-6, the average school received approximately 0.6 kg of the mix per child enrolled and used almost all of it. This represents more than 58 percent of the amount we estimated they would need based on enrollment. Since schools should be cooking for the number of students present, not enrolled, high absenteeism among children suggests that this measure of take-up is a lower bound.[16] Ninety percent of the schools used at least 25 percent of the amount we estimated they would need. The

[16] Attendance rates are between 70 and 80% which would mean that schools used more than 70% of the mix that they needed, higher than the 58% we estimate using enrollment data.

high intensity monitoring did not affect these measures of take-up. For some specifications, we also control for whether the school received IFA tablets during the previous school year but find that additional experience with nutrition supplements did not improve MNM implementation.

Next, we study laboratory reports of the amount of vitamin A and zinc present in meal samples collected at each school (Table 5). These measures allow us to assess take-up between the MNM treatment schools and the non-MNM treatment schools as well as across high and low intensity treatment schools. We estimate:

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

where MNM_{sb} is a dummy variable for schools that received the MNM treatment, and the other variables are as defined above. Meals can contain vitamin A and zinc even if they do not contain the mix, so these measures could be considered a measure of meal quality, and not simply take-up of the intervention. However, dependent variable means from the control group (presented at the bottom of Table 5) indicate very little vitamin A and zinc present in the benchmark samples.

We present these estimates with and without the interaction term – the interaction term accounts for any differential impact that high intensity monitoring may have had on the MNM treatment, but can also reduce statistical power. The results are consistent across both sets of models, but it is worth noting that the results from the "short model" without the interaction term should be interpreted carefully; they represent a weighted average of the treatment effect with and without the cross-cutting treatment. Specifically, the impact of the MNM treatment should be interpreted as the weighted average of the impact of the MNM treatment with low intensity monitoring and the impact of the MNM treatment with high intensity monitoring (Muralidharan, Romero, and Wuthrich 2021).

We find large, significant increases in the amount of vitamin A and zinc for schools in the MNM treatment. The increase in zinc persists through April, the last month of the intervention; the increase in vitamin A is still significant in April, but it is smaller than in February. We suspect this is due to higher stability of zinc than of vitamin A during storage (Kuong et al. 2016). Appendix C describes additional measures of take-up from the midday meal monitoring visits that support our conclusion that take-up did not decline over time, as shown in Appendix Table 1. Consistent with Table 4, high intensity monitoring also does not affect the amount of vitamin A or zinc found in the samples – the coefficients are small and of inconsistent sign. The low levels of vitamin A and zinc in the control group samples suggest that spillovers between treatment arms were very unlikely – headmasters in the control schools did not obtain a similar mix to fortify meals.

As noted above, the dosage agreed upon for our intervention would give children approximately 50% of RDA for the micronutrients listed, including vitamin A and zinc. We conduct back-of-the-envelope calculations based on these measures of take-up to get a better sense of how much of the RDA children in MNM treatment schools received on average. Our midday meal observations document that children receive (and eat) approximately 130 mL of dal or vegetable curry each day, a little bit more than half a cup. We estimate that this weighs about 110 g. Using the range of estimates in Table 5, this means the MNM treatment increased vitamin A intake by 190-376 μg, roughly 30-60% of RDA for this age group, and zinc intake by 2 mg, roughly 20% of RDA for this age group, on the days children ate a meal in school.[17]

B. IFA Implementation

We next report how MNM provision and high intensity monitoring influenced implementation of the government's IFA program. Table 6 estimates the impact of MNM treatment and high intensity monitoring on measures of how well the IFA program was implemented. We use specification (2) above, but also include month fixed effects and cluster standard errors by school since IFA implementation was measured 4 times during the intervention. We focus on three measures of IFA implementation quality: (i) whether the headmaster is able to show the enumerator an IFA tablet (Columns 1-3), (ii) the number of tablets distributed per child in the past week (as reported by the headmaster, Columns 4-6), and (iii) the percent of students who say they get the tablets weekly or more frequently (out of three randomly chosen students spanning different grades, Columns 7-9). The results indicate that neither of our activities affects whether the headmaster shows the enumerator a tablet. Headmasters report distributing more tablets in both treatment arms, but the coefficients are not statistically significant. However, both of our interventions significantly affect whether *students* report getting IFA tablets regularly: students in schools that received the MNM treatment are less likely to report getting IFA tablets regularly, while students in the more intensely

[17] The two measures of take-up are remarkably consistent. One dose of the MNM contained 300 mcg of vitamin A (50-75% of RDA depending on child age) and 5 g of zinc (55-71% of RDA). If schools used about 60-70% of the amount we intended (as calculated from the amount remaining, in Table 4), children would receive 30-53% RDA of vitamin A and 33-50% RDA of zinc, close to the estimates we calculate using the laboratory tests of the meals.

monitored schools are more likely to report getting IFA tablets regularly. Children were randomly chosen each month, making this outcome difficult for school officials to manipulate.

As before, we present these estimates with and without the interaction between the two treatments. The results are consistent across both models and the coefficients on the interaction term are statistically insignificant. Consider the fully saturated model in Column 8. It is worth noting that the magnitudes of the treatment effects are quite similar; along with the small and statistically insignificant interaction term, this suggests that the crowd-out effect of the MNM treatment on IFA implementation effectively cancelled out the crowd-in effect of the monitoring on IFA implementation.[18] We discuss this further in Section IV.C. below.

Panels B and C of Table 6 demonstrate that the monitoring results are stronger later in the school year. The effects are insignificant at the first IFA visit during the intervention (usually in December 2014), when the MNM intervention had just started and many schools in the high intensity treatment arm had yet to receive a monitoring visit. By February 2015, however, the effects start to appear – most high intensity schools had received at least 2 and sometimes 3 midday meal visits while low intensity monitoring schools had received at most 1 visit. In fact, in the later part of the year, even headmasters are more likely to report distributing tablets in the highly monitored schools ($p < 0.1$ in Table 6, Columns 4 and 6). We further explore variation in visit timing in Section IV.D. to understand the mechanisms behind the impact of monitoring.

While headmaster reports confirm the impact of monitoring on IFA implementation, it is not surprising that headmaster reports do not reveal the reduction in IFA tablet distribution in MNM treatment schools; headmasters have an incentive to report fully implementing government programs. In the next section we present evidence consistent with the results using students' reports but not the results using headmasters' reports – by looking at child health as an objective outcome. We also note that during the first IFA implementation survey of the intervention year, students in schools that had received IFA tablets in the *prior* year are more likely to report receiving regular medications in school (Table 6, Panel B, Column 9). This difference goes away

[18] We also examined crowd-out on measures of midday meal implementation, such as whether a meal was served and the contents of the meal. However, because our sample had near-universal implementation in the control group by our measures, this analysis is not informative.

over time, since all the schools received IFA tablets by the start of the intervention year. This finding provides another check on the reliability of student reports of tablet receipt.

C. Child health

We next examine impacts on child health, as measured through hemoglobin levels. Recall that the nutritional motivation for the MNM intervention was that the vitamins and minerals in the mix would complement the iron from the IFA program (making the iron easier to absorb). However, crowd-out of IFA implementation could mute these effects. High-intensity monitoring could also influence child health through its effects on program implementation.

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

$$y_{1isb} = \beta_0 + \beta_1 MNM_{sb} + \beta_2 High_{sb} + \beta_3 (MNM_{sb} * High_{sb}) + \beta_4 y_{0isb} + \alpha_b + age_{isb} + \varepsilon_{isb} \quad (3)$$

where y_{1isb} is the health outcome of child i in school s in block b at endline, and y_{0isb} is a baseline measure of the outcome variable. We include fixed effects for both block, α_b, and age, age_{isb}.

Table 7 presents the treatment effects on a continuous measure of anemia status, hemoglobin levels (in g/dl), in Panel A, while Panel B focuses on a dummy variable for whether a child is anemic.[19] Column 1 includes no lagged controls while Column 2 includes the lagged dependent variable from Baseline 1. Columns 3-5 include the lagged dependent variables from both Baseline 1 and 2 and dummies for missing observations to allow for the inclusion of all children surveyed at endline. Recall that some children were included in the sample only at endline because they were too young to be enrolled in school during the Baseline 1 survey, two years prior to the intervention.

The results in Table 7 indicate that the MNM treatment alone had no effect on child health; in fact, the coefficients are often negative although always small and never significant. By contrast, high intensity monitoring increased hemoglobin levels by 0.17-0.24 g/dL and reduced the probability of being anemic by about 6-9 percentage points after 5 months — a 10-15% decrease relative to the control mean.[20] Note that studies have typically detected changes in hemoglobin levels after

[19] Hemoglobin level cutoffs used to classify children as anemic are those defined by the WHO at sea level by age group (WHO 2011). For the majority of the sample (ages 5-11), children with hemoglobin below 11.5 g/dL are anemic.
[20] We look to other school-based iron supplementation programs to put the magnitude of this result in context. Krämer, Kumar and Vollmer (2018), using double-fortified salt in school midday meals in Bihar, find that hemoglobin

2-3 months of consistent supplementation (Gera et al. 2007) and our programs were implemented over 5 months.

These results are consistent with our evidence on program implementation described above. In Table 7, we present the results with and without the interaction term between the two treatments (the results are consistent across both models), but it is also helpful to view the treatment effects a slightly different way. Table 8 presents a summary of our results but compares each of the treatment groups to the control group. Column 1 replicates the regression from Table 6, Column 8, as a measure of IFA take-up. Column 2 replicates the result from Table 5 using nutrient content of the meal samples as a measure of MNM take-up, except that in Table 8 the observations are stacked by nutrient (vitamin A and zinc) and month (February and April). Column 3 replicates the regression from Table 7, Column 4, estimating the impact on anemia.[21] Considering schools that only received the MNM treatment (row 1), we see a statistically significant increase in the nutritional content of the meals, but substantial crowd-out of the IFA program (p = 0.11), explaining the small and statistically insignificant difference in children's health. For the schools that only received high intensity monitoring (row 2), we see no difference in the nutritional content of the meals and a marginally significant improvement in IFA implementation (p = 0.065), explaining the significant improvement in hemoglobin status. In row 3, we compare the schools that received both the MNM and the intense monitoring and see significant take-up of the MNM (resulting in more nutritious meals), but no evidence of crowd-out of the IFA program. The crowd-in from the monitoring intervention essentially cancelled out the crowd-out from the MNM program. Children are healthier in this treatment arm, likely because of the MNM.[22] These results

increased by about 0.14 g/dL and the probability of being anemic fell by 9.3 percentage points. Luo et al. (2012) provided iron supplements in school to 4th graders in rural China and find that, on average, hemoglobin increased by 0.23 g/dL after one year. For an alternative comparison, a meta-analysis of 55 efficacy trials concludes that consistent iron supplementation increases children's hemoglobin levels by 0.74 g/dL — 1.1 g/dL for children with baseline hemoglobin levels below 11 g/dL and 0.49 g/dL for children with baseline hemoglobin above 11 g/dL (Gera et al. 2007). One might consider these highly-monitored, randomized placebo-controlled trials an upper bound on the potential effect of school-based distribution.

[21] The rest of the outcome variables in Tables 5-7 are presented in this format in Appendix Tables 2-4.

[22] In Appendix Table 5, we verify that the MNM and monitoring treatments had no impact on anthropometric outcomes as we would expect given the short duration of the intervention. Specifically, we estimate the impact on BMI, height-for-age, weight-for-age, and mid-upper-arm circumference. If the improvement in child health in the monitored schools was due to differences unrelated to the IFA or MNM distribution, we might have expected to see improvements in these other measures of health. We also show, in Appendix Table 6, that using a differences-in-differences specification does not affect these conclusions. In results available upon request, we replicate Table 7 for school attendance, cognitive ability, and proficiency in reading and mathematics (see Appendix D for descriptions of

suggest that the MNM program would have improved child health if implementing it had not crowded out the IFA distribution.[23]

D. Crowd-out and crowd-in

Table 8 helps explain how our interventions affected hemoglobin levels via changes in the implementation of the two nutrition programs, but the crowd-out and crowd-in of the pre-existing IFA program warrants further exploration. Specifically, these responses draw attention to the role played by school officials in program implementation, especially when multiple programs compete for administrative time and effort. In this section, we provide evidence of heterogeneous effects that help address three related questions: 1) Why did the MNM program crowd-out IFA implementation? 2) Why did high intensity monitoring crowd-in IFA implementation but not MNM implementation? and 3) What is the mechanism behind the impact of monitoring on headmaster behavior?

To study the first two questions—crowd-out of IFA by MNM and crowd-in of IFA by monitoring—we more closely examine how the IFA and MNM were to be implemented. The time cost of distributing IFA tablets has a variable element that increases with the number of children in the school. Every child must be given exactly one IFA tablet exactly once a week and children must swallow their tablet with food and water in order to minimize side effects (UNICEF 2017). In contrast, the work involved with implementing the MNM does not scale with the number of students. The headmaster or a teacher had to be involved but simply had to calculate the amount of the mix needed (number of students to eat the meal multiplied by 1.5 and divided by 10) and then added the correct number of scoops of the mix to the food before it was served.

---

the data). Neither intervention has statistically significant effects on these outcomes. Again, this is not surprising given the short time horizon and the fact that no other school characteristics likely changed (such as teacher motivation).

[23]Appendix E presents heterogeneity in the child health impacts by baseline hemoglobin status. We find that the positive effect of monitoring on hemoglobin levels is driven by children around the threshold of anemic (around 11.5-12 g/dL), rather than by the children with moderate or severe anemia. This result is surprising, since the nutritional literature on iron supplementation consistently finds that those who are most anemic are more likely to respond to treatment (see, e.g., Abrams et al. 2008, Tee et al. 1999). This unexpected result could be because of lower attendance rates for moderately anemic children (which we also find), but it could also be because of the low levels of micronutrients distributed in one-size-fits-all programs like the IFA program. This conclusion relates closely to that of Banerjee, Barnhardt and Duflo (2018), who study the viability of double-fortified salt as a means to improve anemia levels. Even when provided for free, they find minimal effects on hemoglobin and attribute it to the low levels of iron that can be safely added to food intended for large-scale distribution.

Table 9 re-estimates our main regressions but for subsets of schools to explore heterogeneous effects by school characteristics measured at the beginning of the intervention. Since we have limited power, we omit the interaction term between the two treatments, but present comparisons of each of the three treatment arms with the control group (as in Table 8) in Appendix Table 7. In Columns 1-3 of Table 9 we test whether the impact of our interventions differs by students per staff member since IFA implementation gets more challenging (but MNM implementation does not) as this ratio rises. The dependent variable measures IFA implementation (student reports of tablet receipt) in Panel A and MNM take-up (nutrient content of the meals) in Panel B. We find that the reduction in IFA tablet distribution caused by the MNM program is predominantly found in schools with an above average number of students per staff member ($p = 0.116$ for the difference across types of schools). High intensity monitoring crowds in IFA implementation more in these schools as well (but not significantly more). We find the opposite pattern for MNM take-up: schools with more students per staff member have higher MNM take-up (again, not statistically significant).

This pattern suggests that implementation of these nutrition programs is constrained by limited manpower on the part of headmasters and other school personnel. We next examine heterogeneity by administrative capacity more broadly defined. To implement this analysis, we create an index to proxy for a school's capacity to implement school-based nutrition programs. The index includes having more than four teachers administer the midday meal (the median number), having an external self-help group administer the midday meal, having anyone from the school or self-help group attend a midday meal training, having a record of the most recent school management committee meeting, and reporting sufficient funds to administer the midday meal program. We also include an indicator for whether the school's water supply is treated since treating a schools' water source requires administrative effort on the part of the headmaster and could act as a proxy.[24] Consistent with the results using the student-to-staff ratio, we find that both the crowd-out and crowd-in are driven by schools with below-median scores on the index (Columns 4-6, Panel A). The crowd-out effect is statistically significantly different between above- and below-median

[24] For each of these measures, a positive response indicates high administrative capacity and a negative response indicates low administrative capacity. Failing to answer any of these questions is recorded as low capacity. The index is calculated by standardizing each variable with respect to the control group distribution and summing the standardized variables. We obtain qualitatively similar results if we do not standardize the variables and simply sum the binary indicators, though the p-value on the difference between the two groups increases.

schools. As before, MNM take-up does not differ significantly between these schools (Column 4-6, Panel B), but point estimates suggest meals in MNM schools with below-median capacity are more nutritious.[25]

Taken together, Panels A and B of Table 9 show that crowd-out of IFA by MNM and crowd-in of IFA by high intensity monitoring are concentrated in schools with lower administrative capacity. One possible explanation for these results is that headmasters who feel particularly administratively constrained prioritize implementing the easier nutrition program. In other words, it could be that the MNM program crowds out the IFA program not because the MNM program itself is challenging to implement but because the IFA program is challenging. In addition, the two programs have the same objective: improving child health. It is possible that the introduction of the second program with the same objective enabled headmasters to shirk on the harder one, even though our training program for headmasters highlighted that the two programs were complements, not substitutes.[26] This can also explain why the high intensity monitoring only affected IFA implementation, since administrative constraints were binding only for that program. It is important to note, however, that our experiment was not designed to rigorously test this hypothesis or rule out other explanations.

In Panel C of Table 9 we use student-level data to show that the schools where high-intensity monitoring increased IFA implementation the most—those with the highest student-to-staff ratio or lowest managerial capacity—saw greater reductions in anemia. In MNM schools with low managerial capacity, the crowd out of IFA appears to have been offset by high MNM take-up, resulting in no change in levels of anemia. Although the differences in impacts between below and above median schools are not statistically significant, the analysis in Panel C again suggests that our interventions had downstream effects on child health in schools with fewer administrative resources.

[25] In Appendix Table 8, we estimate heterogeneous effects with respect to each component of the index and find similar results for most of the individual components.

[26] A related possibility is that headmasters sold the iron tablets since the students are now receiving other micronutrients. The fact that these tablets have very little market value in the region, and that we see no significant difference in the ability of headmasters to produce a tablet to show the enumerator (Columns 1-3 in Table 6) provides some evidence against this hypothesis.

Finally, we consider the mechanism for how the two additional monitoring visits affected program implementation. As shown in Panel B of Figure 1, we conducted a number of survey activities in the schools, including a facilities survey, IFA surveys, food sample collection, and attendance surveys. However, as the figure illustrates, the extra two visits represent a substantial increase in monitoring early in the intervention. At the time of the first IFA visit, in December 2014, some high intensity schools had received one meal monitoring visit but others had not; none of the low intensity schools had received a meal monitoring visit. Most schools had not received any other visits. In subsequent IFA surveys (months 3-5 of the intervention), the difference in visits between high and low intensity schools averages about 2 visits, varying slightly due to differences in timing, but proportionally the biggest difference is in the third month of the intervention (February). When IFA take-up was measured in February, high intensity schools had received 2 or 3 midday meal monitoring visits while low intensity schools had received 0 or 1. High intensity schools had received these two additional visits on top of a base of 1-2 total visits. By March, high intensity schools had received two more visits on top of a base of 3-4 and by April, high intensity schools had received two more visits on top of a base of 5-6. When we estimate the impact of high intensity monitoring by month of IFA survey, we find that the increase in IFA take-up reported by students is largest in February when the proportional increase in monitoring is greatest (Appendix Table 9). Appendix Table 9 also confirms that around this time *headmasters* are also more likely to report distributing tablets in high intensity schools; this impact peaks in March. Recall that in Table 6, there was marginally significant evidence that headmasters in high intensity schools were also reporting higher IFA compliance by the end of the school year.

We further exploit the variation in the number of visits and the timing of the most recent visit in Appendix Table 10, pooling all months. First, we confirm that the number of previous meal monitoring visits has a positive and statistically significant effect on IFA implementation compliance reported by both the students *and* headmasters. We also find that the number of previous other (non-meal monitoring) visits did not affect IFA take-up in the same way – we have limited variation within month, but the coefficients are actually of the wrong sign – suggesting that the meal monitoring visits were special.

One significant difference between the meal monitoring and non-meal monitoring visits was timing within the day. The school facilities survey and attendance checks were intentionally not during meal time, to avoid disrupting a busy time and to measure attendance during class time. By contrast, meal monitoring naturally took place during meal time, which is also when headmasters report distributing IFA tablets.[27] The midday meal visits were likely more salient to headmasters than the other visits for other reasons as well. They were 25-71% longer by various measures, such as the number of minutes or pieces of information recorded. Enumerators also interacted with three students chosen randomly from among those present, making detailed observations about the quantity of food being served to these children, in addition to measuring their heights.

In Appendix Table 10, we also provide suggestive evidence for mechanisms by examining nonparametrically which visits most influence IFA implementation, as well as differential impacts by the number of days since the most recent visit. We find that the effect of monitoring grows with the first few visits, suggesting that repeated visits are important, but plateaus after 3 visits. It could be that the visits acted as reminders and that repeated reminders affected behavior. It could also be that headmasters took the accountability suggested by monitoring visits seriously only once they believed the visits would be repeated. We find some suggestive evidence that a visit in the past week improves compliance relative to no previous visit or one more than 5 weeks ago, but the impact of a visit exactly four weeks ago is as big or bigger. One possible explanation is that a visit in the past week increased compliance by acting as a reminder, while a visit one month ago increased compliance by increasing the expectation of a repeat visit. However, we emphasize that the experiment was not designed to differentiate these effects, and we leave the question of precisely how this type of monitoring can impact compliance to future work.

V. Conclusion

We evaluate two interventions aimed at improving implementation and impacts of India's school-based nutrition programs. We show that a program providing MNM to schools actually crowded out implementation of the government's existing IFA program. Consistent with this crowd-out in implementation of the IFA program, we find no effects of the MNM program on child health. We also find that frequent monitoring visits improved implementation of the IFA program and child health, reaffirming that top-down monitoring may be a promising strategy to improve

[27] While our intent was for the monitors to gather information on the quality of midday meals and take-up of the MNM, this intent was not conveyed to the schools. Since schools almost uniformly reported that they distributed the IFA tablets during 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.

implementation of public health programs. In fact, we find that the crowd-in effect of frequent monitoring cancelled out the crowd-out effect of having to implement a second nutrition program: the MNM program *with* frequent monitoring marginally improved child health. Exploring heterogeneity across schools, we provide suggestive evidence that the crowd-out and crowd-in of IFA tablet distribution depends on the administrative capacity of the school.

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

While schools are a natural setting for implementing social programs for children, it is unclear what the optimal number and types of programs should be, and how to hire and motivate school officials to implement the programs effectively. Our results, combined with efficacy trials of multi-micronutrient supplementation for children, suggest that school-based multi-micronutrient distribution remains a promising area; in fact, implementation of the MNM program was less vulnerable to administrative capacity constraints than IFA tablet distribution. Further research is needed to better understand the conditions under which programs are likely to crowd each other out, for example whether crowd-out is limited to programs with similar objectives. Nonetheless, our results highlight the importance of designing programs with administrative capacity constraints in mind. In addition, combining programs with an eye towards easy implementation – for example, a micronutrient mix that also includes iron and folic acid (our initial design before the announcement of the government's IFA program) – may be more effective at improving child health than multiple programs that must be implemented separately.

Further research is also needed to better understand the impact of monitoring visits. While our results suggest that the timing of monitoring visits and who the auditors speak to matters, these hypotheses could be rigorously tested in future work. Our findings also point to the possibility of positive externalities to monitoring one school-based activity, if school officials respond by improving implementation of other school programs, as occurred with the MNM monitoring and

IFA implementation. Finally, our results suggest that monitoring not just new programs, but also pre-existing programs, could minimize crowd-out when administrative capacity is limited.

REFERENCES

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

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

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2 (1): 1-30.

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

Banerjee, Abhijit V., Rachel Glennerster, and Esther Duflo. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6 (2/3): 487-500.

Banerjee, A., Hanna, R., Kyle, J. C., Olken, B. A., & Sumarto, S. 2018. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia." *Journal of Political Economy* 126 (2): 451-491.

Berry, J., S. Mehta, P. Mukherjee, H. Ruebeck, and G.K. Shastry. 2020. "Implementation and Effects of India's National School-Based Iron Supplementation Program." *Journal of Development Economics*, 144(2020).

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

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

Björkman Nyqvist, Martina, Damien De Walque, and Jakob Svensson. "Experimental evidence on the long-run impact of community-based monitoring." *American Economic Journal: Applied Economics* 9.1 (2017): 33-69.

Bundy, D. A. P., N. de Silva, S. Horton, D. T. Jamison, and G. C. Patton 2018. *Optimizing Education Outcomes: High-Return Investments in School Health for Increased Participation and Learning*. Washington, DC: World Bank.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee. 2015. "Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan." NBER Working Paper 21180.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman. 2013. “School Inputs, Household Substitution, and Test Scores.” *American Economic Journal: Applied Economics* 5 (2): 29-57.

Debnath, S., Nilayamgode, M. and S. Sekhri. 2020. “Information Bypass: Using Low-Cost Technological Innovations to Curb Leakages in Welfare Programs,” Working paper, University of Virginia.

Deserranno, E., Nansamba, A., & Qian, N. 2021. “The Unintended Consequences of NGO-Provided Aid on Government Services in Uganda.” NBER Working Paper 26928.

Dhaliwal, Iqbal, and Rema Hanna. 2017. "The devil is in the details: The successes and limitations of bureaucratic reform in India." *Journal of Development Economics* 124: 1-21.

Duflo, E., R. Hanna, and S.P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102 (4): 1241-1278.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. “Truth Telling by Third-Party Audits and the Response of Polluting Firms: Experimental Evidence from India.” *Quarterly Journal of Economics* 128 (4): 1499-1545.

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

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2017. "The personnel economics of the developing state." Handbook of economic field experiments 2: 467-514.

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

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

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

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

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

Krämer, M., S. Kumar and S. Vollmer. 2020. “Improving child health and cognition: Evidence from a school-based nutrition intervention in India.” *Review of Economics and Statistics* 2020:1-51.

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

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

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

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

Miller, G., R. Luo, L. Zhang, S. Sylvia, Y. Shi, P. Foo, Q. Zhao, R. Martorell, A. Medina, S. Rozelle and H.F. Farnsworth. 2012. "Effectiveness of provider incentives for anaemia reduction in rural China: a cluster randomised trial." *British Medical Journal* 345:e4809 doi: 10.1136/bmj.e4809

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

Muralidharan, K. and P. Niehaus, 2017. "Experimentation at Scale." *Journal of Economic Perspectives* 31(4): 103-124.

Muralidharan, K., P. Niehaus, and S. Sukhtankar, and J. Weaver. 2018. "Improving last-mile service delivery using phone-based monitoring." *American Economic Journal: Applied* Economics 13(2) 52-82.

Muralidharan, K., M. Romero, and K. Wüthrich. 2021. "Factorial designs, model selection, and (incorrect) inference in randomized experiments." Working paper. UC San Diego.

Muralidharan, K., and V. Sundararaman, 2013. "Contract Teachers: Experimental Evidence from India." NBER Working Paper 19440.

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

Pradhan, Menno P., Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Armida Alisjahbana, and Rima Prama Artha. 2014. "Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Economic Journal: Applied Economics* 6 (2): 105-126.

Ramakrishnan U., N. Aburto, G. McCabe, and R. Martorell. 2004. "Multi-micronutrient interventions but not vitamin A or iron interventions alone improve child growth: results of 3 meta-analyses." *Journal of Nutrition* 134: 2592–2602.

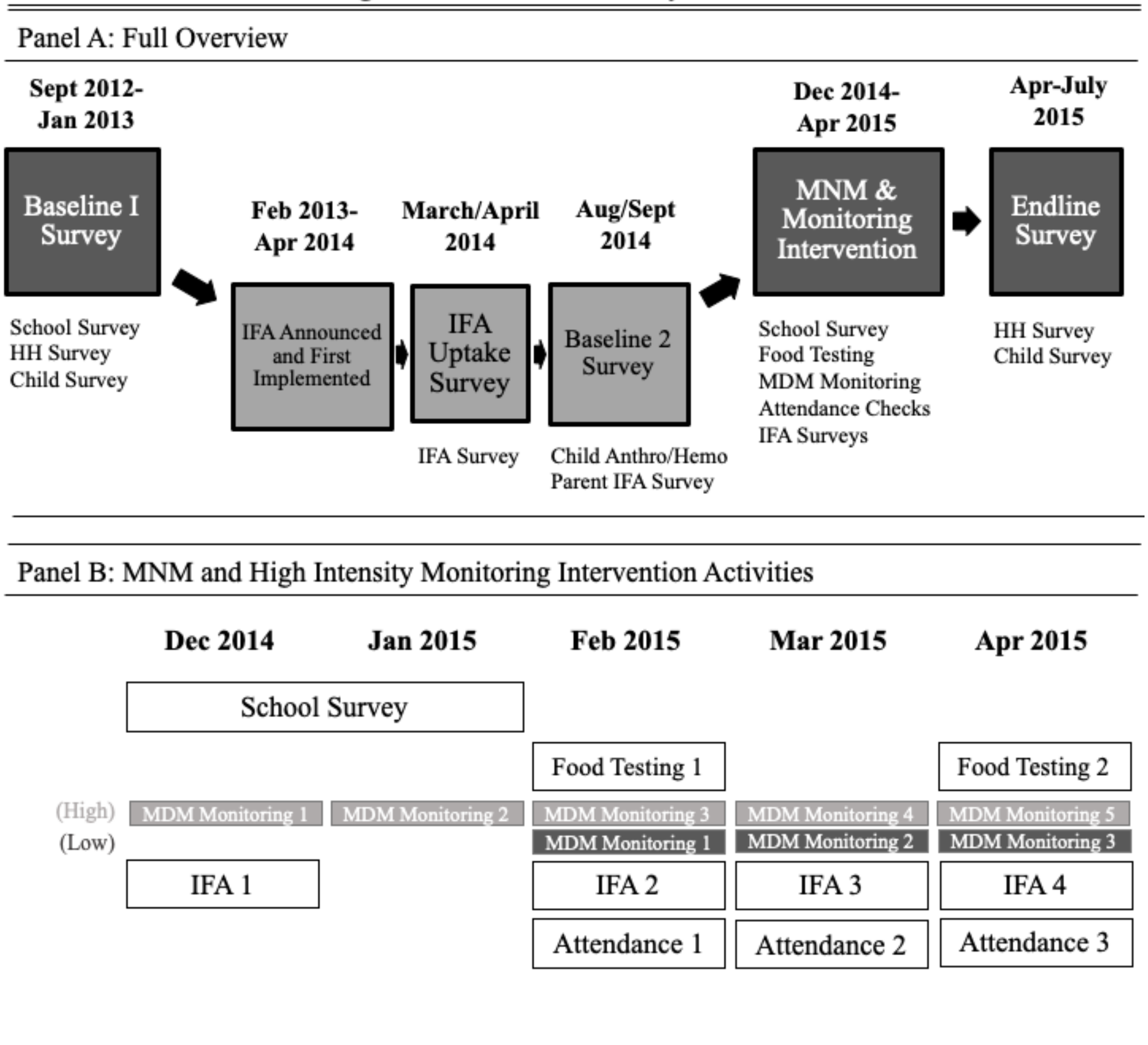
Tee, E.S., M. Kandiah, N. Awin, S. Chong, N. Satgunasingam, L. Kamarudin, S. Milani, A.E. Dugdale, and F.E. Viteri. 1999. "School-administered weekly iron-folate supplements improve hemoglobin and ferritin concentrations in Malaysian adolescent girls." *The American Journal of Clinical Nutrition* 69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

UNICEF. 2017. Process Documentation on National Iron Plus Initiative (NIPI) in Odisha. Government of Odisha. Accessed online at https://www.unicef.org/india/reports/process-documentation-national-iron-plus-initiative-nipi-odisha.

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

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

Figure 1: Timeline of key activities



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

Table 1: Treatment arms

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

Table 2: Balance across treatments at baseline: School characteristics

	Control	Only MNM	Only high intensity	Both	P-value of all 3 differences
Panel A: At Baseline 1 or earlier					
Distance to the block headquarters (km)	22.973	22.789	24.861	24.889	0.815
Primary enrollment	85.351	84.868	74.028	74.378	0.686
Secondary enrollment	13.270	10.132	11.361	16.378	0.814
Number of teachers	2.514	2.421	2.472	2.486	0.994
Number of female teachers	2.757	2.868	2.528	2.676	0.641
Number of rooms	4.455	4.444	4.057	3.778	0.516
Percent of schools have a kitchen	0.784	0.833	0.800	0.676	0.462
Percent of schools have at least one latrine	0.838	0.789	0.889	0.865	0.693
Percent of schools have sufficient water	0.778	0.667	0.735	0.622	0.474
Percent of schools with treated water	0.324	0.263	0.286	0.243	0.887
Percent with parent group for MDM	0.394	0.444	0.471	0.343	0.713
Percent with MDM training	0.389	0.324	0.314	0.333	0.918
Percent receiving MDM rice on a regular schedule	0.472	0.486	0.400	0.278 *	0.225
Received IFA during previous year	0.811	0.842	0.917	0.892	0.540
Panel B: First month of the intervention					
Primary enrollment	86.838	83.526	68.333	73.432	0.466
Secondary enrollment	10.595	10.395	11.278	15.622	0.823
Number of teachers	3.162	3.105	3.250	3.108	0.986
Number of female teachers	3.108	3.711	2.972	3.216	0.248
Number of rooms	4.611	4.514	4.556	4.622	0.995
Percent of schools have a kitchen	0.784	0.886	0.818	0.833	0.688
Percent of schools have at least one latrine	0.917	0.947	0.944	0.944	0.957
Percent of schools have sufficient water	0.706	0.556	0.667	0.706	0.531
Headmaster has bachelor's degree or higher	0.412	0.216 *	0.344	0.441	0.156
Percent of schools with 4+ teachers helping out with midday meal	0.595	0.658	0.583	0.595	0.908
Percent of schools with parent group for MDM	0.333	0.342	0.343	0.243	0.732
Percent of schools with treated water	0.514	0.263 **	0.382	0.270 **	0.091
Percent with MDM training	0.559	0.343 *	0.441	0.545	0.232
Percent of schools with school management committee records	0.216	0.342	0.278	0.351	0.529
Percent of schools reporting sufficient funds to administer the midday meal program	0.735	0.781	0.594	0.647	0.356
Percent receiving MDM rice on a regular schedule	0.657	0.684	0.588	0.514	0.448
Received IFA tablets this year	1.000	1.000	1.000	0.973	0.319
Number of tablets distributed per child past week (school report)	0.838	1.105	0.889	0.889	0.621
Percent of students who say they get meds weekly or more frequently (out of 3)	0.417	0.356	0.455	0.480	0.713
Number of schools	37	38	38	37	

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

Table 3: Balance across treatments at baseline: Child characteristics

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

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

Table 4: Take-up of MNM by schools

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

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

Table 5: Take-up of MNM, as seen in micronutrient levels from lab tests of food samples

	February						April					
	Vitamin A			Zinc			Vitamin A			Zinc		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MNM treatment	351.9***	347.4***	351.6***	16.6***	14.5***	16.6***	165.8***	181.9***	165.7***	15.6***	16.6***	15.6***
	(44.8)	(65.4)	(44.8)	(2.8)	(4.2)	(2.8)	(33.4)	(52.7)	(33.4)	(4.4)	(5.8)	(4.4)
High intensity	-5.3	-10.0	-9.8	1.3	-0.8	1.4	-5.6	10.3	-6.3	5.7	6.7	5.6
	(44.6)	(25.4)	(44.9)	(2.8)	(2.1)	(3.0)	(31.7)	(31.2)	(32.5)	(4.5)	(6.1)	(4.5)
MNM treatment * high intensity		9.3			4.1			-32.3			-2.1	
		(90.8)			(5.6)			(67.2)			(8.9)	
Received IFA during previous year			66.6			-2.7			12.9			1.8
			(77.0)			(5.8)			(57.8)			(5.7)
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	0.216	0.154	0.156	0.155	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 shows treatment effects on the micronutrients (namely, vitamin A and zinc) present in school meals, as measured in the laboratory using samples collected by enumerators during February and April of the treatment year. All columns include block fixed effects. Robust standard errors are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 6: Treatment effects on IFA program implementation

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

*Notes: This table shows treatment effects on how well the government's IFA program was implemented. We use three measures of IFA implementation quality: (i) whether the headmaster shows the enumerator an IFA tablet (Columns 1-3), (ii) the number of tablets distributed per child in the past week, as reported by the headmaster (Columns 4-6), and (iii) the percent of students who say they get the tablets weekly or more frequently, out of three randomly selected students that were asked the question (Columns 7-9). All columns include block fixed effects and survey month fixed effects. Standard errors clustered by school are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

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

Lagged dependent variable from survey:	None	Just Baseline 1	Baseline 1 and Baseline 2 with dummies for missing		
	(1)	(2)	(3)	(4)	(5)
Panel A: Dep var: Hemogloblin (g/dL)					
MNM treatment	-0.044	-0.012	-0.018	0.032	-0.022
	(0.057)	(0.067)	(0.057)	(0.072)	(0.057)
High intensity	0.174***	0.244***	0.179***	0.229***	0.168***
	(0.058)	(0.067)	(0.058)	(0.079)	(0.059)
MNM treatment * high intensity				-0.101	
				(0.114)	
Received IFA during previous year					0.129
					(0.089)
N	1920	1118	1920	1920	1920
R-squared	0.024	0.173	0.129	0.130	0.131
Panel B: Dep var: Anemic					
MNM treatment	-0.000	-0.024	-0.009	-0.023	-0.009
	(0.026)	(0.031)	(0.026)	(0.035)	(0.026)
High intensity	-0.066**	-0.089***	-0.064**	-0.077**	-0.062**
	(0.027)	(0.030)	(0.027)	(0.035)	(0.027)
MNM treatment * high intensity				0.027	
				(0.052)	
Received IFA during previous year					-0.022
					(0.045)
N	1920	1113	1920	1920	1920
R-squared	0.017	0.136	0.089	0.089	0.089

*Notes: The dependent variable in each specification is child's hemoglobin in g/dl (Panel A) and an indicator for whether a child is anemic (Panel B). All columns include block and age fixed effects, in addition to the lagged dependent variable as described in the headers. Standard errors, clustered by school, are in parentheses. Significance at the 0.10, 0.05, and 0.01 levels indicated by *, **, and ***, respectively.*

Table 8: Treatment effects by treatment group

Measure of	IFA take-up	MNM take-up	Child health
Dependent Variable	Percent of students who say they get meds weekly or more frequently (out of 3)	Nutrient content in meal sample	Anemic
	(1)	(2)	(3)
MNM treatment only	-0.073	140.8***	-0.023
	(0.045)	(22.4)	(0.035)
High intensity only	0.072*	1.6	-0.077**
	(0.039)	(12.9)	(0.035)
Both treatments	0.021	136.5***	-0.073*
	(0.046)	(19.4)	(0.039)
N	538	586	1920
R-squared	0.128	0.281	0.089
Dep. var mean, control group	0.608	30.4	0.589

*Notes: This table presents the difference between each of the treatment arms and the control group for the percent of students who report receiving IFA tablets frequently (Column 1), the amount of nutrients present in the meal sample (Column 2), and child anemia status (Column 3). The specifications are as in Tables 6, 5 and 7 except that we estimate separate treatment effects for the three intervention arms. Column (2) also includes fixed effects for survey month interacted with nutrient. 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 by school characteristics

	Above-median students to staff ratio			Above-median capacity index		
	No	Yes	P-value of diff	No	Yes	P-value of diff
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dep var: Percent of students who say they get meds weekly or more frequently (out of 3)						
MNM treatment	-0.010	-0.112***	0.116	-0.110**	0.012	0.057
	(0.049)	(0.042)		(0.044)	(0.046)	
High intensity	0.052	0.117***	0.309	0.117***	0.029	0.165
	(0.049)	(0.040)		(0.042)	(0.047)	
N	262	276		304	234	
R-squared	0.188	0.112		0.148	0.130	
Panel B: Dep var: Nutrient content in meal sample						
MNM treatment	119.110***	150.206***	0.268	146.749***	120.749***	0.361
	(21.003)	(18.611)		(18.482)	(21.649)	
High intensity	10.165	-22.551	0.248	-10.911	6.789	0.527
	(21.390)	(18.517)		(17.798)	(21.671)	
N	292	294		330	256	
R-squared	0.281	0.315		0.302	0.266	
Panel C: Dep var: Anemic						
MNM treatment	-0.011	-0.015	0.934	0.001	-0.024	0.641
	(0.036)	(0.038)		(0.034)	(0.043)	
High intensity	-0.054	-0.088**	0.508	-0.083**	-0.045	0.498
	(0.038)	(0.036)		(0.035)	(0.043)	
N	951	969		1081	839	
R-squared	0.102	0.097		0.107	0.084	

*Notes: This table presents heterogeneous treatment effects by school characteristics on the percent of students who report receiving IFA tablets frequently (Panel A), the amount of nutrients present in the meal sample (Panel B) and child anemia status (Panel C). The specifications are similar to those in Table 6, Table 5, and Table 7 (respectively), except that the sample is split according to the students to staff ratio (Columns 1-2) and the capacity index (Columns 4-5). All columns include block fixed effects and survey month fixed effects. Panel B also includes fixed effects for survey month interacted with nutrient. Column 3 (6) presents p-values testing the hypothesis that the coefficients in Columns 1 and 2 (4 and 5) are equal; these are estimated from a fully saturated regression with interactions between the school characteristic and the treatments as well as the 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.*