Courant Research Centre

'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

Georg-August-Universität Göttingen
(founded in 1737)

Discussion Papers

No. 282

Iron-fortified school lunch: medium-run impacts of a school nutrition intervention in India

Liza von Grafenstein, Abhijeet Kumar, Santosh Kumar, Sebastian Vollmer

First version: July 2021
This version: March 2023

Platz der Göttinger Sieben 5 · 37073 Goettingen · Germany
Phone: +49-(0)551-3921660 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

Iron-fortified school lunch: medium-run impacts of a school nutrition intervention in India

Liza von Grafenstein[1], Abhijeet Kumar[2], Santosh Kumar[3], Sebastian Vollmer[4]

Abstract

Using a phase-in research design, we provide experimental evidence on the impacts of early versus late initiation of iron fortification in school lunch programs on children's health and cognitive outcomes in India. We find higher hemoglobin levels and a lower likelihood of anemia in the early treatment group that experienced four years of treatment, compared to the late treatment group that experienced only one and a half years of treatment. Despite significant health gains, we do not find evidence for treatment effects on cognitive and educational outcomes. Heterogeneity analyses show evidence of gendered effects—the anemia reduction is lower among females relative to males.

Keywords: Anemia, double-fortified salt, cognition, children, school lunch, India
JEL: C93, I15, I18, O10

[1]1Grafenstein: Department of Agricultural Economics and Rural Development, Georg August University of Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany (liza.grafenstein@uni-goettingen.de) & ID Insight India Private Limited, New Delhi, India
[2]Kumar: Centre for Modern Indian Studies, Georg August University of Göttingen, Germany (abhijeet.kumar@cemis.uni-goettingen.de)
[3]Kumar (corresponding author): Keough School of Global Affairs, University of Notre Dame, Notre Dame, Indiana, USA, 46556 (skumar23@nd.edu); Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA & IZA, Germany.
[4]Vollmer: Department of Economics & Centre for Modern Indian Studies, Georg August University of Göttingen, Germany (svollmer@uni-goettingen.de)

This study is registered in the AEA RCT Registry with the unique identifying number: "AEARCTR-0006800". It has received ethical approval from the Georg August University of Göttingen. We thank three anonymous referees and the editor for their insightful comments and suggestions. An earlier version of this manuscript was titled "Impacts of Double-Fortified Salt on Anemia and Cognition: Four-Year Follow up Evidence from a School-Based Nutrition Intervention in India". The current version supersedes all previous versions.

I. Introduction

Micronutrient deficiencies remain a health problem in developing countries, affecting over half of the world's population. More than one-half of preschool-aged children and two-thirds of non-pregnant women of reproductive age suffer from micronutrient deficiencies (Stevens *et al*., 2022). Growth impairment due to malnutrition in early childhood can lead to poor cognitive function and is often associated with lower human capital, lower household per capita expenditure, and a higher probability of living in poverty as an adult (Dasgupta and Ray, 1986; Strauss and Thomas, 1998; Alderman, 2006; Hoddinott *et al*., 2013). Iron deficiency anemia (IDA) is a common nutritional deficiency that can significantly impact children's physical and cognitive development. Addressing IDA is of interest to policymakers because it has a high prevalence in developing countries and a strong association with grades, attendance, and educational attainment (Halterman *et al*., 2001; Bobonis, Miguel, Puri-Sharma, 2006; Chong *et al*., 2016; Li *et al*., 2018).[2]

Efficacy studies conducted in a controlled environment show significant impacts of double-fortified salt (DFS) on anemia (Rajagopalan and Vinodkumar 2000; Sivakumar *et al*., 2001; Sivakumar and Nair 2002). Yet, a large-scale DFS program has shown limited impacts on anemia in developing countries (Banerjee, Barnhardt, and Duflo 2018). Additionally, a handful of studies evaluate the impacts of DFS and iron

[2] There are many types of anemia, but the most common type is iron-deficient anemia (IDA)—more than 50% of anemia cases worldwide are due to iron deficiency. Nutritional deficiencies and infectious disease burden are the leading causes of anemia among children globally but the predominant cause of IDA in India is a nutritional deficiency (WHO 2015).

supplementation pills and find limited short-term impacts of these interventions on anemia and cognition (Chong *et al*., 2016; Banerjee, Barnhardt, and Duflo 2018).[3]

Krämer, Kumar, and Vollmer (2021a), hereafter KKV, show that a school-based DFS intervention in India significantly reduced anemia but had no impact on cognitive outcomes after one year of treatment. It is of further interest whether the short-run health gains persist or wane over time and whether longer treatment duration increases the magnitude of the program impacts, relative to shorter treatment duration. Furthermore, several studies show program impacts in the medium- and long-run even when there are no impacts in the short run (Behrman *et al*., 2011; Baird *et al*., 2016; Parker and Todd, 2017; Bouguen *et al*., 2018; Ozier, 2018). Therefore, it is likely that despite null effects on cognition in the short term, continuous and sustained DFS treatment in the medium- to long-term may affect cognition. The key objective of this study is to estimate the medium-term impacts of school-based DFS intervention on anemia, cognition, and learning outcomes in rural parts of India.

This study is a follow-up to the KKV study, and causally identifies the effect of longer versus shorter treatment duration of a school-based DFS intervention on children's anemia status and cognition. We exploit the phase-in research design to estimate the persistence of short-run treatment impacts in the medium term. The original control group of schools was phased into treatment after 28 months, which created early and late treatment groups, a randomization design similar to Miguel and Kremer (2004).[4] Bouguen *et al*. (2018) highlight that measuring medium- or long-term

[3] DFS is fortified with iron and iodine.

[4] Students in the treated schools received fortified school lunch for 32-48 months and the students in the control schools received fortified school lunch for 4-20 months. Hence, in the absence of a pure control group, the treatment effect in this study is "the effect of more DFS or higher DFS program intensity".

program impacts is important for a better understanding of the underlying mechanisms of interventions. Additionally, measuring the short-term benefits of health interventions is valuable but does not offer much insight into the program effects if benefits persist in the long run. Thus, the long-term program impacts could be relevant to the analysis of the cost-effectiveness of the programs.

We conducted the DFS experiment in two administrative blocks of Jehanabad district in Bihar. Bihar is one of the poorest states in India, with a high prevalence of anemia among school children (World Bank, 2016; IIPS and ICF, 2017). The intervention delivered DFS to 54 public primary and middle schools starting in 2015, while the 53 control schools only started receiving DFS in December 2017. As the endline survey for this study was conducted in 2019, our phase-in experimental design provides about four years of treatment exposure in the early treatment schools and about one and a half years in the late treatment schools. The panel nature of data allows us to use the double differences (DD) method with child fixed effects to estimate the causal impacts of the DFS intervention on children's health, cognition, and education outcomes.

Our results show that children in the early treatment group benefited more than children in the late treatment group, who received the treatment for only one and a half years. Children in the early treatment group had on average a 0.267 g/dL higher hemoglobin (Hb) level than those in the late treatment group and also a lower likelihood of any anemia—less than 13.4 percentage points (pp). Further, the school-based DFS intervention decreased the incidence of mild anemia by 10.6 pp. The treatment effects on health outcomes are non-uniform and vary by gender and school attendance rate. However, despite significant health gains, we found no impact on the children's cognitive outcomes. The non-significant effect on cognition is similar to the findings in

Miguel and Kremer (2004)—the deworming health intervention found no cognitive gains after one to two years of treatment (Miguel and Kremer, 2004). We believe the null impact on cognitive outcomes is an important finding of this paper and we discuss the possible reasons for this in detail in the results section.

Our study makes several important contributions to the literature on the effects of nutritional interventions on the health and cognitive outcomes of school children. First, we provide evidence of the medium-run impact of a school-based nutritional intervention. KKV show that the iron-fortified school lunches program in primary school resulted in a significant decrease in anemia after one year of treatment. This study shows there is a cumulative effect of continuous treatment on anemia, as the magnitude of the treatment effect is larger in this four-year follow-up study compared to the short-term impacts. Our null effects on cognitive outcomes add to the evidence base on the relationship between DFS intervention and learning outcomes.

The exposure to DFS in the early treatment group started when children were on average eight years old in grade III (mid-childhood) while we collected health and other human development outcomes after the treatment period, which was a maximum of four years. At this time the children were on average twelve years old and in grade VI (early adolescence). The outcomes in mid-childhood years or early adolescence are important for long-term outcomes and are more amenable to policy interventions (Almond, Currie, and Duque 2018).

Second, our study also contributes to the policy debate on a channel for delivering micronutrients to children that will have high take-up and compliance. Other rigorous studies using the same DFS formula in India focus on another public channel: the Public Distribution System (PDS) (Banerjee, Barnhardt, and Duflo 2013, 2018). Banerjee,

Barnhardt, and Duflo (2013, 2018) found that despite the free delivery of DFS, only 61% to 75% of households used DFS, and impacts on general health or cognition for the pre-defined groups were not statistically significant.

Furthermore, our findings also speak to the effectiveness of fortified school lunches (mid-day meal (MDM) in our case) on the health and cognitive outcomes of school children.[5] The MDM has also been used to provide additional micronutrients via other fortification vehicles in India, as in the study by Berry *et al*. (2021), who assessed the impact of using a micronutrient mix to fortify the MDM in Odisha, India. The authors did not find any effects on Hb levels, child health, or human capital measures of cognitive or learning outcomes. Other studies using DFS found effects on Hb levels (Osei *et al*. 2010; Radhika *et al*. 2011; Pinkaew *et al*. 2013). However, these studies differ in sample size, diversity in the ages sampled, treatment duration, and treatment intensity—these studies had smaller sample sizes and short treatment durations, often less than one year. Our study provides novel insights into the increased effectiveness of the DFS in the MDM given an earlier treatment start.

Finally, this study provides evidence of DFS effectiveness for a group that is still somewhat understudied by health economists: adolescents. Two related studies find significant impacts of iron pills on anemia and cognition among adolescents in Peru (Chong *et al*., 2016), and somewhat modest impacts of DFS intervention on anemia in India (Banerjee *et al*., 2018). More recently Berry *et al*. (2020) examined the MDM program to provide iron-folic acid (IFA) supplementation for adolescent children in grades I-V in Odisha, India. They found that the IFA program had nonsignificant effects

[5] A few studies with school-level treatment used multiple fortified salts for meals and found increases in Hb levels as well as in memory and attention levels for children aged 5-18 (Sivakumar et al. 2001; Kumar and Rajagopalan 2007; Vinodkumar and Rajagopalan 2009).

on anemia on average, but the program had significantly large effects for moderately anemic students in schools that were administering the programs more recently, compared to schools with interrupted program implementation.

To the best of our knowledge, this is one of the handful of studies to estimate the medium-term impacts of DFS intervention after about four years of continued treatment among adolescents in India. Overall, we show that a longer period of nutrition interventions is more effective and there is a cumulative effect on health outcomes. Policymakers may consider school-based nutrition interventions as a means of addressing anemia among school children and ensuring that DFS is widely available and accessible to schools at a subsidized price. Another crucial finding of our study is the absence of significant effects on cognitive and educational outcomes despite the positive health outcomes. This could have various explanations, including that there is no causal effect of reduced anemia on cognition or educational outcomes, that higher doses of iron supplementation are required to affect cognitive outcomes, or that effect sizes are too small to be detected with the statistical power of this study.

II. Related Literature

Iron supplementation and cognition link

The biomedical literature highlights the need of adequate iron intake for optimal child development. There is high physiological iron demand during childhood to support physical, neural, and cognitive development. Prior studies have shown that iron deficiency (ID) or IDA is an important risk factor for the development of cognition and brain cells (Beard, 1995; Halterman, *et al*. 2001). The foundations of the central nervous system are laid down in early childhood. ID impairs the development of the central

nervous system and thus reduces children's capacity for cognitive development and can lead to development delays (Halterman, *et al*. 2001). Children with iron deficiency are likely to have problems with attention span, cognition, and sensory perception, as well as difficulties with emotions and behavior.

Despite a strong theoretical link between IDA and cognitive development, the empirical evidence supporting this link is scant and ambiguous. The evidence on the impact of the iron intervention on Hb and cognition depends on the dosage and type of the treatment, duration of the treatment, inflammation due to other micronutrient deficiencies, age group of the children, take-up of the treatment, and type of cognitive tests. Therefore, it is difficult to draw conclusions from it. Iron supplementation pills seem to be more effective than DFS due to their higher iron content. On the one hand, several studies have found beneficial impacts of iron interventions on cognitive development. For example, a study found that 30 mg iron supplementation for 4 months improved cognition scores in children in India aged 8-15 years old (Seshadri and Gopaldas, 1989). In a systematic review, iron supplementation improved global cognitive scores by 0.5 (standardized mean difference) and there was an overall benefit to IQ from iron supplementation in anemic primary school children (Low *et al*., 2013). A systematic review found beneficial effects on the IQ of iron interventions in children aged 8 years or older (Sachdev *et al*., 2005). A study by East *et al*. (2021) found that iron deficiency in infancy was associated with poor executive control, and lower verbal IQ at age 21 in Chile.

In contrast, in an RCT conducted in Thailand, a 100 mg ferrous sulfate intervention for 14 weeks had no significant effect on cognition scores, suggesting that cognitive deficits due to ID are irreversible (Pollitt *et al*., 1989). Kayshap and Gopaldas (1987) found that 4 months of 60 mg iron supplements had no impact on cognition, but the

cognitive effects appeared after 8 months. At 8 months, iron-supplemented IDA children had improved Hb levels and cognitive tasks compared to control children. In an experimental study conducted in China, Luo *et al.* (2012) found that a daily dose of 5 mg of iron for 5 months increased Hb level by 0.2 standard deviations, reduced anemia by about 10 percentage points, and raised the test scores of anemic students by about 0.2 standard deviations. Another study conducted by Chong *et al.* (2016) in Peru found that the treatment of 500 mg of iron pills for 10 weeks reduced anemia by 34% and increased cognitive test scores by 21% among anemic adolescents. These effects were found only in the sample of baseline anemic children. Non-anemic children were unaffected by the treatment. Two evaluation studies of DFS found mixed impacts on anemia reduction but there were no impacts on cognitive scores (Banerjee, Barnhardt, and Duflo, 2018; Krämer *et al*., 2021a). The ambiguous effects of iron interventions on cognitive development may be due to differences in the dose and duration of the interventions, timing of commencement of iron supplementation, population, epidemiology of anemia, and compliance with the interventions (Larson, Phiri, and Pasricha, 2017).

Medium to long-run impacts of health interventions

Baird *et al.* (2016) show that ten years after the school-based deworming program in Kenya, treated boys had better educational and labor market outcomes. These effects were sustained after 20 years—treated groups of children experienced a meaningful improvement in adult living standards and earnings (Hamory *et al*., 2021). Vikram and Chindarkar (2020) examine the impacts of the Integrated Child Development Services (ICDS) scheme in India that provided food supplementation, immunization, health check-ups, and early childhood education to children under the age of six years, and

they find a positive impact on cognitive achievement, primarily for girls and children in low-income families. An iodine supplementation program that provided iodine pills to pregnant women in Tanzania increased years of schooling (Field, Robles, and Torero, 2009); however, a replication study of the same intervention found the results to be sensitive to the specification choice and sample restriction (Bengtsson, Sävje, and Swartling Peterson, 2019). Tafesse (2022) shows that in-utero and early-life exposure to the Universal Salt Iodization program improved children's test scores in rural India. The study by Adhvaryu *et al.* (2020) found that in-utero access to iodized salt had positive effects on income, labor force participation, and full-time employment, particularly for women. This is likely because iodine deficiency can have negative impacts on cognitive development and physical health, which can, in turn, affect an individual's capacity for paid employment.

Mid-Day Meal program

The mid-day meal program in India is a large-scale school lunch program that was launched in 1995 with the goal of increasing primary school enrollment and attendance and reducing hunger and malnutrition among primary school children in public schools. The program provides a daily lunch with a predefined menu and content – a minimum of 450 calories and 8-12 grams of protein - on at least 200 days per year to all primary and upper primary school children in public schools throughout India. The Food Corporation of India provides grains directly to the schools and school staff purchase other ingredients at the local market. The program is funded by the government and administered through public schools.

The MDM program covered an estimated 120 million school children in 1.2 million schools in 2017-18. The MDM program in Bihar served about 10 million lunches every

day and covered all children from grades I-VIII. The program had an average cost of approximately US¢6.4 per meal for primary school children and US¢9.6 per meal for upper primary school children, which amounts to an annual cost of $13 for primary and $19 for upper primary for the mandated 200-day school year. The school lunch is served every day except Sundays and holidays.

We use the MDM program as the delivery platform for distributing iron to children. The widespread coverage of the MDM program ensures higher compliance, as children who are attending schools will be consuming fortified meals regularly. The fortified MDM also allows for a regular, steady, and nearly daily provision of iron to children, which is more effective than intermittent intake of iron pills. Additionally, the program has the potential to reach a large fraction of the high-risk population at a low marginal cost, due to its wide coverage and the fact that many children from low-income households attend public schools. Therefore, we believe that the MDM program could be a cost-effective distribution channel because it already has established infrastructure and logistics.

III. Data and experimental design

The intervention was implemented in two administrative blocks (Kako and Modanganj) in the Jehanabad district of Bihar, India. Out of the 228 government-funded schools in the two blocks, 107 schools were randomly selected to participate in the study. 54 schools were randomly chosen to receive the DFS intervention starting in 2015, while the remaining 53 schools started receiving DFS in 2017. At baseline children were 7-9 years old, which is a critical period for brain development. In this period the frontal lobes develop, and this in turn affects cognitive development (Thatcher 1991; Anderson 2002). The baseline survey was conducted between November 2014 and January 2015

and included around 2,000 children — about 20 children from grade II from each school. The treatment started in August 2015 and the endline survey of this study was conducted from February to July 2019.

Intervention

Our intervention implemented a school-based delivery system for providing DFS to treatment schools each month, to be used in MDM preparation. The DFS was provided to the treatment schools at a subsidized price and the control schools continued to use the conventional iodized salt—the subsidy amount was equal to the price difference between the DFS and the conventional iodized salt used by the schools in pre-program years. DFS serves as a complement to other foods, not a substitute. Only the headmaster and the cook were informed of the treatment and cooks were advised to use the DFS in the MDM preparation instead of the usual iodized salt. Our field staff conducted unannounced visits to treatment schools to monitor the availability and use of DFS and the quality of meals.

The rollout of the intervention was staggered. The treatment schools received the intervention from August 2015 to December 2019. The one-year evaluation in KKV showed positive health effects of the DFS intervention on Hb levels and anemia, therefore, for ethical reasons the study team did not want to deny the health benefits of DFS to control schools. The control schools started receiving DFS in December 2017. Appendix Table A1 displays the timeline of the intervention. It should be noted that we do not have a pure control group anymore; rather the identification comes from the variation in the treatment duration between early and late treatment groups.

The types of schools that students attend generate additional variation in the exposure to treatment (for an overview see Table A2). Primary schools (PS) provide

education from grades I-V while middle schools (MS) can either have grades from I-VIII or VI-VIII. The children in our sample in the baseline year attended grade II, and several of them transitioned to another type of MS during the course of the study.[6] This could bias our findings if differences in children's characteristics of children correlated with the types of schools they moved to, or if there were differences in the quality of education provided by the different types of schools. Since parents and children were not informed about the intervention, the selection of schools after grade V is unlikely to be influenced by the treatment status of the schools. This helps to minimize the potential for selection bias, as the assignment of children to different types of schools was orthogonal to the intervention. It is worth noting, however, that other factors may have influenced the selection of schools by children and their parents. For example, if children and their families have certain characteristics (such as higher socioeconomic status), they may be more likely to attend certain types of schools. It is important to consider these potential confounders while interpreting our main findings.

Our treatment design led to treatment exposure of 32-48 months for children who attended a treatment school at baseline. The children who moved to MS who are not in our sample after grade V would have received fortified meals for 32 months (from August 2015 to March 2018), while children who stayed in one of our sampled schools consumed fortified meals for up to 48 months (from August 2015 to July 2019). The children in the control group who attended one of our sampled schools at least until grade V received the DFS in their lunch only briefly (four months from December 2017

[6] Starting in April 2018, children who were enrolled in PS had to transition to MS to attend grade VI. Some of the children joined MS with grades I-VIII that had been selected as a school in our sample. To continue with their education, other children went to schools (MS with grades I-VIII or only VI- VIII) that were not included in our sample. We were able to survey children who had attended at baseline either one of 43 treatment schools or one of 42 control schools.

to March 2018). Other children in the control group who continued going to a school in our sample had fortified school lunches for at most 20 months (from December 2017 to July 2019). This staggered treatment design gives us a panel sample of 1,058 children for health outcomes and 808 children for cognition and education outcomes.

In our study we are interested in the intention-to-treat (ITT) effects of a longer-lasting school-lunch fortification program, so we focus on the original treatment and control group. This means that the measured effects are potentially downward biased because a group of children in the treatment group did not receive the treatment for the full four years, and there isn't a pure control group. Thus, the ITT estimates would be a conservative estimate of the DFS intervention.

DFS, Iron Content, and Compliance

The DFS was purchased from private firms, delivered to district headquarters, and either distributed to schools by field staff or collected by the headmaster. The DFS was fortified with 0.86 mg of iron per gram of salt and was intended to provide up to 50% of the daily dietary iron requirement for children. The intervention was designed to provide, on average, 4 g of DFS, providing 3.5 mg (4*0.86) of iron per meal—delivering about 29-44% of the daily iron requirement for children between 4 and 12 years of age (Bihar Government, 2017; Krämer *et al.*, 2021a; WHO, 1959; see appendix A for more details). However, the actual iron intake would depend on several factors, including compliance with the use of the DFS in the MDM program at the school level and children's attendance rate and consumption of school meals. If children attended schools five days a week and the DFS was used consistently in the preparation of school lunches, this would result in an estimated intake of 17.5 mg of iron per week for each

student. The unannounced visits by our field staff and robust monitoring mechanism ensured that students were receiving the intended amount of iron in the treated schools.

Furthermore, the analysis of monitoring data shows that treatment schools had a higher probability of having DFS in the kitchen and receiving a DFS delivery from the study team, compared to control schools. Additionally, a higher percentage of headmasters of the treated schools reported that the use of DFS changed the color of the food, which was further evidence that DFS was being used in the MDM preparation in the treatment schools. Hb levels are another indicator of strong program compliance. The Hb distribution shifted right in the post-treatment period, indicating a positive effect of the program on Hb levels (Krämer *et al*., 2021a). Taken together, this evidence suggests that the program was successful in increasing the availability and use of DFS in the MDM program, resulting in improved Hb levels in the treated schools.

Variables

The main health outcomes of interest are Hb levels in g/dL, any anemia status, mild anemia status, and moderate or severe anemia status. Depending on the Hb level, we classify the children's anemia status following the WHO (2011) and adjusting for age. Any anemia is defined as a Hb value < 11.5 g/dL for children aged 5-11 years, < 12g/dL for children aged 12-14 years and girls aged 15 and above, and < 12.9 g/dL for boys aged 15 and above. Mild anemia is defined as a Hb value $\geq$ 11 & < 11.5 g/dL for children aged 5-11 years, value $\geq$ 11 & < 12g/dL for children aged 12-14 years and girls aged 15 and above, and $\geq$ 11 & < 12.9 g/dL for boys aged 15 and above. Moderate or severe anemia is defined as a Hb value < 11 g/dL. We group moderate and severe anemia status because there were very severely anemic children in our baseline sample.

Effects on cognitive development were assessed by the following five tests: block design, forward digit-span, backward digit span, Raven's Colored Progressive Matrices, and Stroop tests (Malin, 1969; Gerstadt, Hong, and Diamond 1994; Raven, Raven, and Court 1998, Röthlisberger et al. 2010)).[7] Based on these five tests we constructed a cognitive index using principal component analysis. We adapted the tests to capture the temporal and secular increase in cognitive ability. We derived math and reading test scores using the survey tool developed by the Indian Governmental Organization Pratham (ASER Centre, 2014). All cognitive outcomes were normalized by subtracting the baseline mean and dividing it by the baseline standard deviation. The unit of the outcome is interpreted as standard deviations from the baseline mean. The control variables include time-variant characteristics such as household size, the mother's and father's years of schooling, and an asset index. The asset index was generated using the first component of a principal component analysis consisting of several household assets.

Balance check and attrition

Tables 1 and A5 describe the baseline characteristics and outcome variables across the treated and control groups for the health and education samples, respectively. The baseline characteristics are mostly balanced across the treatment arms apart from a few exceptions in Table 1. The Hb level is not balanced across the treatment arms and that led to an imbalance in anemia outcomes as well, since the anemia variable is based on the Hb cutoff. The pre-attrition data in columns (5)-(8) have already been reported in the KKV study since the baseline sample is the same in both studies. All the control

[7] Table A3 summarizes the test details.

variables in Panel B and C are balanced across treatment and control at baseline, apart from gender and student-teacher ratio in column 4 and caste in column 8. We attribute these imbalances to chance since randomization was carried out carefully and correctly. Moreover, our empirical model includes child fixed effects that are likely to account for these baseline imbalances in health outcomes and covariates.[8] Except for health outcomes, covariates are balanced in Table A5, indicating that children were successfully randomized across interventions in the education sample.

(Table 1 about here)

Systematic differences between children who stayed in the study and those who dropped out could bias the estimates. For example, severely anemic children may miss school more frequently, or students with better cognition may have transferred to high-quality private schools. After four years of treatment, our study suffers from an attrition rate of about 40% compared to the baseline sample.[9] The main reason for attrition is the relocation of households, transfer to private or boarding schools, and migration to urban areas for work.

To check for selective attrition, Table 2 examines the correlation between the probability of attrition and baseline characteristics. The model regresses the attrition dummy on treatment, observed baseline characteristics, and the interaction of treatment and baseline characteristics. Results show that some of the baseline characteristics are statistically significantly associated with attrition. For the health outcomes sample

[8] Since anemia prevalence is higher in the treated groups relative to the control groups, the parameter 2 in equation (1) will potentially be biased in the downward direction.

[9]The baseline sample available for analysis is 1,789 students for the health sample. Of these 1,051 students have been successfully re-interviewed in 2016 and 2019 with all available covariates for our estimation sample for health outcomes. The attrition rates are 41.3 % ((1789-1051)/1789) and 53.7% ((1770-820)/1770) in the health and cognitive & education sample, respectively.

presented in column (1), we find that gender and mother's primary schooling predict attrition. The interactions of the treatment dummy with anemia and gender are also statistically significant. Column (2) shows the results for the cognition and education samples. We find that total school enrollment and the student-teacher ratio predict attrition. None of the interactions of the treatment dummy with baseline characteristics are statistically significant in column (2). Overall, the significant coefficients in Table 2 show evidence of non-random attrition. We address these concerns related to selective and differential attrition by weighting the regression model with inverse probability weights (IPW).

(Table 2 about here)

IV. Empirical specification

We employ the IPW-weighted Double Difference (IPW-DD) method to estimate the causal impact of the DFS intervention on children's health and cognitive outcomes. The following model would estimate the ITT effects of the intervention:

$$Y_{ist} = \alpha_i + \beta_1 Post_t + \beta_2 Post_t \times Treat_s + \delta_1 X_{it} + \varepsilon_{ist} \qquad (1)$$

where, Y_{ist} represents outcome variables for child i attending school s at time t. α_i constitutes the intercept and captures child fixed effects. $Post_t$ is a dummy variable that takes the value of one for the post-treatment period and zero for the pre-treatment period. $Treat_s$ is a dummy indicator of assignment to the intervention arm of longer exposure to the DFS; otherwise, it is zero. The variable $Treat_s$ is not included separately because the main effect of treatment is constant within the child and is absorbed by the child fixed effects. X_{it} denotes time-variant controls at the child level. ε_{ist} is the independent and identically distributed error term across clusters and children

within clusters. Standard errors are clustered at the level of randomization, the school level.

To address concerns related to non-random attrition, we assigned weights to observations based on the inverse probability of their attrition status. We generated the weights with a probit model including all variables used in Table 2. The parameter β_2 is the ITT estimate of the effect of the DFS intervention on children's outcomes. Additionally, we report corrected standard errors for multiple hypothesis testing using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

V. Results

Health and Cognitive Impacts

The main results on Hb, anemia, cognitive scores, and test scores are presented in Table 3. Results are from the estimation of the IPW-DD model which accounts for attrition between the baseline and endline surveys.[10] We find that DFS treatment had substantial positive impacts on Hb level which in turn led to a reduction in anemia prevalence. Compared to the late treatment group that consumed fortified MDM only for one and a half years, children in the early treatment group who were exposed to treatment for four years had on average a 0.267 g/dL higher Hb level. We also find statistically significant effects on any anemia and mild anemia. On average, children in the early treatment group have a 13.4 pp and 10.6 pp lower prevalence of any and mild anemia compared to children in the late treatment group. Considering the baseline mean prevalence of any and mild anemia of 40.2% and 16.1%, the estimated treatment effects translate to

[10] Results for the unweighted model are shown in Table 5 as a comparison. The difference between the treatment effect in the unweighted model and the IPW model is primarily driven by the variables for which the interactions of baseline characteristics with the treatment dummy are statistically significant in Table 2 (baseline anemia status and gender).

a 33.3% and 66.6% reduction, respectively.[11] The sign for moderately and severely anemic children is negative but the estimates are not statistically significant.

These medium-term impacts are substantially higher than the short-term impact estimates (mean Hb level increase of 0.19 g/dL and anemia reductions of 22% and 27% after one year of treatment in the KKV study). This points to an increasing and cumulative reduction in anemia from longer term DFS supplementation in school lunches. The biomedical literature suggests that the cumulative effects may be due to higher iron absorption supported by steady and continuous use of DFS (Hurrell, 2021). A severely anemic child has a lower iron absorption capacity and as the iron level in the body increases, the iron absorption capacity is expected to increase. A steady and continuous use of DFS increases the iron reserves in the body in a cumulative way.

Furthermore, there could be behavioral changes in response to the school meal program that may lead to inadvertent reductions in home food. Our experiment was designed in such a way that only the headmaster and cook of the school were aware of the treatment and households/parents/villagers were unaware that school meals were fortified with iron. Therefore, it is unlikely that parents would have responded to the interventions in a way that may bias our findings. In a companion paper, we found no impact of the nutrition information experiment (parents were informed of the anemia status of their children and were encouraged to feed iron-rich food) on parental response to a change in diet quality. This indicates that even when parents knew the anemia status of their children and received information about iron-rich diets, they did not change their behavior (Krämer, Kumar, and Vollmer, 2021b). Furthermore, there is some

[11]This is the ratio of the coefficient divided by the baseline mean of the estimation sample: 0.134/0.402 or 0.106/0.159.

evidence of partial intrahousehold redistribution of food resources away from the school lunch recipients to other children in the household (Chakraborty and Jayaraman, 2019), but this would affect children in both treatment and control schools the same way. For the reasons mentioned above, it is very unlikely that the substitution of resources away from the children who are consuming fortified meals would be related to the treatment with DFS.

(Table 3 about here)

Panels B and C show the effects on cognitive outcomes and test scores, respectively. We do not find statistically significant effects of the DFS intervention on cognitive outcomes. The coefficients are imprecisely estimated and, unexpectedly, change signs across outcomes. The coefficients for math and reading test scores are positive but are also imprecisely estimated. These results show that despite substantial improvements in the Hb levels and a statistically significant and substantial reduction in anemia, overall neither cognitive outcomes nor test scores were affected. We are likely underpowered to estimate treatment effects on cognitive outcomes and test scores (see Table A4 for power calculation).

Heterogeneity

The intensity of treatment probably increases with school attendance as students would consume the fortified MDM more frequently, thus leading to greater exposure to the DFS intervention. Table 4 explores heterogeneity by baseline school attendance. We calculated school attendance as the ratio of the total number of days a child was present in school and the total number of days the school was open in the twelve months before the baseline survey. We took school attendance information from the school register. We conducted the heterogeneity analyses for the sub-groups of 70, 80, and 90%

attendance in Panel A and attendance terciles in Panel B. There is evidence of heterogeneous impacts on health outcomes by attendance rate. For any anemia and mild anemia, the effect sizes are larger for high-attendance sub-groups. For example, any anemia probability is 17.0 pp at 70% attendance, over 18.1 pp at 80%, and 21.5 pp at 90% attendance. All coefficients are statistically significant at least at the 5% level of significance. For Hb levels and moderate or severe anemia, such a pattern is less obvious for both attendance levels and terciles. Yet, the effects of DFS intervention on any anemia or mild anemia are higher in the middle and top attendance terciles compared to the bottom attendance tercile.

Turning to the cognitive and education outcomes, the heterogeneous pattern is even less consistent as none of the estimates are precisely estimated, and they even alternate signs without becoming statistically significant (Table A6). This is different from KKV, in which there were sizeable effects of the treatment on math and reading outcomes for children with high school attendance. This could imply that for these outcomes the shorter treatment duration was sufficient for children in the late treatment group to catch up with children in the early treatment group. Another potential explanation for the absence of statistically significant effects on the education outcomes could be the low variation in these outcomes due to ceiling effects (Wang *et al*. 2008). More than 53% and 67% of the children received the highest or almost highest possible scores in math and reading tests, respectively. The median math score is 15 (the highest possible score is 16) and the median reading test score is 4 (the highest possible score is 4). In the reading test, the median child in grade VI reached the highest level (read a story fluently) while the median child in the previous waves reached the level of being able

to read letters (see figure A1 in appendix). This issue of ceiling effects was not yet prevalent in KKV when math and reading tests were conducted in lower grades.

We further explore heterogeneity by gender, baseline anemic status, and school quality (Table A7-A10).[12] The DFS treatment seems to have had a stronger impact on boys' health than on girls' health, on non-anemic children than on anemic children at baseline, and on children in high-quality schools than on children in low-quality schools.[13] For adolescent girls, iron needs double after menarche. These increased iron requirements for adolescent girls may explain the null impacts for female students.

(Table 4 about here)

Robustness and Sensitivity Analysis

We conducted several robustness checks to test the validity and stability of our main findings. The results are in Table 5. We estimated equation (1) without the time-variant control variables (panel A); with the inclusion of IFA supplementation as control (panel B), and with the inclusion of school-level controls to ward off the confounding effects of school quality (panel C). The model without control variables has a larger sample size but the results are qualitatively similar (panel A).

(Table 5 about here)

Panel B shows the results after accounting for the implementation of the weekly iron-folic acid supplementation (WIFS) program. The WIFS program provides a weekly IFA tablet containing 100 mg elemental iron and 500 ug folic acid for each child attending grades VI-XII of government, government-aided, and municipal

[12] We are not able to use differences in meal quality for our heterogeneity analysis because all public schools in the state serve the same menu with fixed contents.

[13] We define high quality schools as schools that have attendance higher than the mean of school attendance rate at baseline.

schools (Ministry of Health & Family Welfare, 2016). The WIFS program is administered by the schools. The WIFS program started in late 2017 in Bihar and in 2018 only reached four schools that the children in our sample attended. In 2019, all but one school in our sample was participating in the WIFS program. The provision of IFA supplements might have affected our estimates if the IFA supplementation program were correlated with the treatment, which is highly unlikely. Regardless, we added it as an additional control to the covariates in this robustness check. The effects do not change either in statistical significance or sign. They are similar to the main findings in Table 3. Moreover, the inclusion of time-variant school-level controls (enrollment, class size, and student-teacher ratio) in Panel C also does not affect the main findings for anemia. However, the magnitude of Hb shrinks and it is no longer statistically significant.

Furthermore, we used a non-parametric attrition correction approach in Panel D. Lee (2009) bounds create worst-case scenarios to achieve equal attrition between treatment and control groups assuming that participants who select into the sample rank at the top or the bottom of the outcome distribution (Tauchmann, 2014). The group with less attrition is trimmed from above (dropping observations with high values) for lower bounds or from below (dropping observations with low values) for upper bounds. Though for Hb the upper bound is positive and statistically significant, the lower bound has a negative sign but does not deviate significantly from zero. For any anemia and mild anemia, the bounds are negative in sign. Only the lower bounds are statistically significant at the 5% significance level. This assures us that our results still hold in the worst case, in which students with lower Hb levels and mild anemia attrited. However,

in a scenario where only students with high Hb levels and non-anemia students left the sample, we cannot exclude the possibility of a null effect.

Finally, we corrected standard errors for multiple hypothesis testing using the Benjamini-Hochberg method as a robustness check (Benjamini and Hochberg, 1995; Benjamini, Krieger, and Yekutieli, 2006; Anderson, 2008) because testing for multiple hypotheses increases the probability of false rejection of at least one null hypothesis. The sharpened q-values for the main findings are reported in Table A11 and the findings of the heterogeneity analysis are in Table A12. Results reported in Table A11 and A12 are qualitatively similar to the main findings in Table 3 and the findings for the heterogeneous treatment effects (Table 4, A7, A9) respectively. For example, the significance level of all the coefficients in the main findings in Table 3 remains, with a few exceptions in which the significance level reduces from the 1% level to the 5% level.

VI. Discussion and Conclusions

A prior evaluation of the DFS program through a school lunch program found improvement in Hb levels and a reduction in mild anemia among school children in a resource-constraint setting like Bihar, India (Krämer *et al.*, 2021a). However, whether the short-run effects of DFS persist or dissipate in the medium- to long-run remained an open question. We fill this gap by conducting a four-year follow-up study of the DFS experiment conducted in KKV. We find that after about four years of continued treatment, treated children, on average, have higher Hb levels and a lower likelihood of being anemic than children who were treated for only one and a half years. The point estimates for the health outcomes in our study are larger than the magnitude of the one-year treatment effects in the KKV study. According to our findings, higher treatment

intensity—measured by school attendance—increases the effectiveness of the DFS intervention. However, although health status improved, we did not find an improvement in children's cognitive and educational outcomes. Our paper contributes to the limited literature on the medium-term impacts of fortified school lunches on health and cognitive gains among children. Our paper also provides novel insights into how a childhood intervention (in grade III) affects outcomes for young adolescents (in grade VI), which has implications for outcomes in adulthood.

In contrast to the heterogeneous effects on test scores in the one-year evaluation of our intervention, we find no such evidence after about four years of continued treatment. Ceiling effects, adapted tests, lack of power, and the increasing influence of school quality are all possible explanations for the absence of significant effects (Wang *et al*., 2008). As we do not have a pure control group, our results should be interpreted as effects of earlier and longer treatment relative to a control group that also received the DFS for a year and a half prior to the follow-up survey.

From a policy perspective, our results are important. We show the potential of the use of fortified foods in school feeding programs to increase the health of adolescents, without crowding out other interventions. Unlike the labor-intensive school-based food fortification intervention in Berry *et al*. (2021), using DFS in the MDM does not require additional labor or resources because salt is an important ingredient for school lunch and school staff were already purchasing and using the non-DFS salt before the intervention. So, our DFS intervention does not impose an additional burden on school staff or cooks, and therefore, is easy to implement and scale up. However, we note two bottlenecks that may constrain the scaling up of the DFS program at the school level—price and availability of the DFS in the local market. DFS is not easily available in the

local market and is also two times more expensive than non-DFS salt. These two factors may affect the scaling up of this program, otherwise, it is relatively easy to implement and scale up. Given the substantial impacts on anemia reduction due to the DFS intervention, an important policy implication could be that the government should make DFS available through the public distribution shop (or fair price shop) at a subsidized price to schools and households.

The results highlight that anemia reduction induced by a treatment that starts two-and-a-half-years earlier and lasts longer can be retained even when all children received the DFS in the MDM for about the past one and a half years. This suggests that early treatment initiation and longer treatment duration are important. An important policy implication is that even a continuous use of the DFS in one meal provided at school for about four years is not sufficient to affect the cognitive outcomes of anemic children. It is likely that mild anemia is not a barrier to learning and school attendance, therefore, a reduction in mild anemia may have little or no effect on cognition and learning. Future studies should examine a larger sample of severely anemic children to explore the effects of DFS or iron supplementation programs on cognitive and learning outcomes. Apart from a treatment arm with only a DFS supply, future studies coould complement the DFS intervention with measures improving the quality of education in schools. Future work could also include preschool children, different types of fortified food products, appropriate tests to overcome the ceiling effects and a larger sample size of severely anemic children to understand the nuanced relationship between micronutrient deficiency and cognitive development.

Acknowledgments

We thank the editor Climent Quintana-Domeque and three anonymous referees for constructive suggestions. We are grateful to John Hoddinott, late Prof. Stephan Klasen, Christina Martini, and Meike Wollni for feedbacks on earlier versions of this paper. We thank Bianca Dülken, Thea Eyting, Lena Gempke, Dara Krolpfeifer, Maximilian Köster, Tabea Recksiek, and Carmen Steinmetz for excellent research assistance. We thank conference participants at NEUDC 2020, NCDE 2021, DGGÖ 2021, GLAD 2021, ACEGD 2021, and ANH Academy Week 2022 as well as participants and discussants in the GlobalFood Research Colloquium at Georg August University of Göttingen, the International Nutrition Seminar at Cornell University, and the Kyoto Environment & Development Seminar at Kyoto University for valuable feedback. We thank Professor Dr. Claudia Mähler for her collaboration regarding cognitive ability tests.

Funding Sources

This work was supported by the German Research Foundation (DFG) for funding via RTG 1666 and RTG 1723 and by SWAGATA - Erasmus+ Consortium of Göttingen, Heidelberg, and Köln Universities. Santosh Kumar acknowledges the funding received from the Sam Houston State University Office of Research and Sponsored Programs.

References

Adhvaryu, A., Bednar, S., Molina, T., Nguyen, Q., Nyshadham, A. (2020). 'When It Rains It Pours: The Long-Run Economic Impacts of Salt Iodization in the United States', *The Review of Economics and Statistics*, Vol. 102, pp. 395–407.

Alderman, H. (2006). 'Long term consequences of early childhood malnutrition', *Oxford Economic Papers*, Vol. 58, pp. 450–474.

Almond, D., Currie, J. and Duque, V. (2018). 'Childhood Circumstances and Adult Outcomes: Act II', *Journal of Economic Literature*, Vol. 56, pp. 1360–1446.

Anderson, P. (2002). 'Assessment and development of executive function (EF) during childhood', *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, Vol. 8, pp. 71–82.

ASER Centre (2014). 'Annual Status of Education Report', (ASER).

Baird, S., Hicks, J.H., Kremer, M. and Miguel, E. (2016). 'Worms at work: long-run impacts of a child health investment', *The quarterly journal of economics*, Vol. 131, pp. 1637–1680.

Banerjee, A., Barnhardt, S. and Duflo, E. (2013). 'Nutrition, Iron Deficiency Anemia, and the Demand for Iron-Fortified Salt: Evidence from an Experiment in Rural Bihar', *Discoveries in the Economics of Aging*, pp. 343–384.

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

Behrman, J.R., Parker, S.W.,Todd, P.E. (2011). 'Do conditional cash transfers for schooling generate lasting benefits? A five-year followup of PROGRESA/Oportunidades', *J. Hum. Resourc*, Vol. 46, pp. 93–122.

Bengtsson, N., Sävje, F. and Peterson, S.S. (2020). 'Fetal Iodine Deficiency and Schooling: A Replication of Field, Robles, and Torero (2009)', *The Scandinavian Journal of Economics*, Vol. 122, pp. 582–621.

Berry, J., Mehta, S., Mukherjee, P., Ruebeck, H. and Shastry, G. K. (2020). 'Implementation and effects of India's national school-based iron supplementation program', *Journal of Development Economics*, Vol. 144, 102428.

Berry, J., Mehta, S., Mukherjee, P., Ruebeck, H. and Shastry, G. K. (2021). 'Crowd-out in school-based health interventions: Evidence from India's midday meals program', *Journal of Public Economics*, Vol. 204, 104552.

Bobonis, G.J., Miguel, E. and Puri-Sharma, C. (2006). 'Anemia and School Participation', *J. Hum. Resourc*, Vol. 41, pp. 692 – 721.

Bouguen, A., Huang, Y., Kremer, M. and Miguel, E. (2019). 'Using RCTs to Estimate Long-Run Impacts in Development Economics', *Annual Review of Economics*, 11:1, 523-561.

Chong, A., Cohen, I., Field, E., Nakasone, E. and Torero, M. (2016). 'Iron Deficiency and Schooling Attainment in Peru', *American Economic Journal: Applied Economics*, Vol. 8, pp. 222-55.

Chakraborty, T. and Jayaraman, R. (2019). 'School feeding and learning achievement: Evidence from India's midday meal program', *Journal of Development Economics*, Vol. 139, pp. 249-265.

Dasgupta, P. and Ray, D. (1986). 'Inequality as a Determinant of Malnutrition and Unemployment: Theory', *The Economic Journal*, Vol. 96, pp. 1011–1034.

East, P., Doom, J. R., Blanco, E., Burrows, R., Lozoff, B., & Gahagan, S. (2021). 'Iron deficiency in infancy and neurocognitive and educational outcomes in young adulthood', *Developmental psychology*, Vol. *57*, pp. 962–975.

Field, E., Robles, O. and Torero, M. (2009). 'Iodine deficiency and schooling attainment in Tanzania', *American Economic Journal: Applied Economics*, Vol. 1, pp. 140–169.

Gerstadt, C. L., Hong, Y. J. and Diamond, A. (1994). 'The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test', *Cognition*, Vol. 53, pp. 129–153.

Hamory, J., Miguel, E., Walker, M., Kremer, M. and Baird, S. (2021). 'Twenty-year economic impacts of deworming', *Proceedings of the National Academy of Sciences of the United States of America*, *118*(14), e2023185118.

Halterman, J. S., Kaczorowski, J. M., Aligne, C. A., Auinger, P., & Szilagyi, P. G. (2001). 'Iron deficiency and cognitive achievement among school-aged children and adolescents in the United States', *Pediatrics*, Vol. *107*, pp. 1381–1386. https://doi.org/10.1542/peds.107.6.1381

Hoddinott, J., Behrman, J. R., Maluccio, J. A., Melgar, P., Quisumbing, A. R., Ramirez-Zea, M., Stein, A. D., Yount, K. M. and Martorell, R. (2013). 'Adult consequences of growth failure in early childhood', *American Journal of Clinical Nutrition*, Vol. 98, pp. 1170–1178.

Hurrell R. F. (2021). 'The Potential of Iodine and Iron Double-Fortified Salt Compared with Iron-Fortified Staple Foods to Increase Population Iron Status', *The Journal of nutrition*, *151*(Suppl 1), 47S–63S.

International Institute for Population Sciences (IIPS) and ICF. 2017. 'National Family Health Survey (NFHS-4)', 2015-16: India. Mumbai: IIPS.

Krämer, M., Kumar, S. and Vollmer, S. (2021a). 'Improving Child Health and Cognition: Evidence from a School-Based Nutrition Intervention in India', *The Review of Economics and Statistics*, Vol. 103, pp. 818-834.

Krämer, M., Kumar, S. and Vollmer, S. (2021b). 'Anemia, diet, and cognitive development: Impact of health information on diet quality and child nutrition in rural India', *Journal of Economic Behavior & Organization*, Vol. 190, pp. 495-523.

Kumar, M. V., & Rajagopalan, S. (2007). 'Multiple micronutrient fortification of salt and its effect on cognition in Chennai school children', *Asia Pacific journal of clinical nutrition*, Vol. 16, pp. 505–511.

Larson, E. L., Wong-McLoughlin, J. and Ferng, Y. H. (2009). 'Preferences among immigrant Hispanic women for written educational materials regarding upper respiratory infections', *Journal of community health*, Vol. 34, pp. 202–209.

Larson, L. M., Phiri, K. S., and Pasricha, S. R. (2017). 'Iron and Cognitive Development: What Is the Evidence?', *Annals of nutrition & metabolism*, 71 Suppl 3, 25–38.

Lee, D.S. (2009). 'Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects', *The Review of Economic Studies*, Vol. 76, pp. 1071–1102.

Li, L., Huang, L., Shi, Y., Luo, R., Yang, M. and Rozelle, S. (2018). 'Anemia and student's educational performance in rural Central China: Prevalence, correlates and impacts', *China Economic Review*, Vol. 51, pp. 283–293.

Low, M., Farrell, A., Biggs, B. A., and Pasricha, S. R. (2013). 'Effects of daily iron supplementation in primary-school-aged children: systematic review and meta-analysis of randomized controlled trials', *CMAJ : Canadian Medical Association journal*, Vol. 185, pp. E791–E802. https://doi.org/10.1503/cmaj.130628

Malin, A. J. (1969). 'Malin's intelligence scale for Indian children', (Nagpu (India), Child Guidance Center).

Miguel, E., and Kremer, M. (2004). 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities', *Econometrica*, Vol. 72, pp. 159–217.

Ministry of Health & Family Welfare. (2016). 'Operational Guidelines for Weekly IFA Supplementation Programme for School Based Adolescents: Guidelines for Teachers and Principals', Government of India.

Osei, A. K., Rosenberg, I. H., Houser, R. F., Bulusu, S., Mathews, M. and Hamer, D. H. (2010). 'Community-level micronutrient fortification of school lunch meals improved vitamin A, folate, and iron status of schoolchildren in Himalayan villages of India', *The Journal of nutrition*, Vol. 140, pp. 1146–1154. https://doi.org/10.3945/jn.109.114751.

Ozier, O. 2018. 'Exploiting externalities to estimate the long-term effects of early childhood deworming', *Am. Econ. J. Appl. Econ*. Vol. 10, pp. 235–62.

Parker, S.W. and Todd, P.E. (2017). 'Conditional cash transfers: the case of PROGRESA/Oportunidades', *J. Econ. Lit*., Vol. 55, pp. 866–915.

Pinkaew, S., Winichagoon, P., Hurrell, R. F. and Wegmuller, R. (2013). 'Extruded rice grains fortified with zinc, iron, and vitamin A increase zinc status of Thai school children when incorporated into a school lunch program', *The Journal of nutrition*, Vol. 143, pp. 362–368. https://doi.org/10.3945/jn.112.166058.

Radhika, M. S., Nair, K. M., Kumar, R. H., Rao, M. V., Ravinder, P., Reddy, C. G. and Brahmam, G. N. (2011). 'Micronized ferric pyrophosphate supplied through extruded rice kernels improves body iron stores in children: a double-blind, randomized, placebo-controlled midday meal feeding trial in Indian schoolchildren', *The American journal of clinical nutrition*, Vol. 94, pp. 1202–1210. https://doi.org/10.3945/ajcn.110.007179.

Raven, J. C. and Court, J. H. (1998). 'Coloured Progressive Matrices', (Oxford, Oxford Psychologists Press).

Röthlisberger, M., Neuenschwander, R., Michel, E., & Roebers, C. M. (2010). 'Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten Vorschulalter'. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, Vol. 42 Suppl 2, pp. 99–110. doi: 10.1026/0049-8637/a000010

Sivakumar, B., Brahmam, G. N., Madhavan Nair, K., Ranganathan, S., Vishnuvardhan Rao, M., Vijayaraghavan, K. and Krishnaswamy, K. (2001). 'Prospects of fortification of salt with iron and iodine', *The British journal of nutrition*, Vol. 85 Suppl 2, pp. 167–173.

Stevens, G. A., Beal, T., Mbuya, M. N. N., Luo, H., Neufeld, L. M. and Global Micronutrient Deficiencies Research Group. (2022). 'Micronutrient deficiencies among preschool-aged children and women of reproductive age worldwide: a pooled analysis of individual-level data from population-representative surveys', *The Lancet. Global health*, 10(11), e1590–e1599.

Strauss, J. and Thomas, D. (1998). 'Health, Nutrition, and Economic Development', *Journal of Economic Literature*, Vol. 36, pp. 766–817.

Tauchmann, H. (2014). 'Lee (2009) treatment-effect bounds for nonrandom sample selection', *Stata Journal*, Vol. 14, pp. 884–894.

Thatcher, R. W. (1991). 'Maturation of the human frontal lobes: Physiological evidence for staging', *Developmental Neuropsychology*, Vol. 7, pp. 397–419.

Wiktoria, T. (2022). 'The effect of Universal Salt Iodization on cognitive test scores in rural India', *World Development*, Vol. 152, 105796.

Vikram, K. and Chindarkar, N. (2020). 'Bridging the gaps in cognitive achievement in India: The crucial role of the integrated child development services in early childhood', *World Development*, Vol. 127, 104697.

Vinodkumar, M. and Rajagopalan, S. (2009). 'Efficacy of fortification of school meals with ferrous glycine phosphate and riboflavin against anemia and angular stomatitis in schoolchildren', *Food and nutrition bulletin*, Vol. 30, pp. 260–264.

Wang, L., Zhang, Z., McArdle, J. J. and Salthouse, T. A. (2009). 'Investigating Ceiling Effects in Longitudinal Data Analysis', *Multivariate behavioral research*, Vol. 43, pp. 476–496.

WHO. 2011. 'Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity', Vitamin and Mineral Nutrition Information System, Geneva, World Health Organization).

WHO. 2015, 'The global prevalence of anaemia in 2011', Geneva, World Health Organization.

World Bank. 2016. 'Bihar - Poverty, Growth, and Inequality', India States Briefs, Washington, D.C. :World Bank Group.

Table 1: Balance in samples before and after attrition using baseline data for the health sample

Baseline sample	After attrition				Before attrition			
	N	Control means	Treatment means	p-values	N	Control means	Treatment means	p-values
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Child level outcome variables								
Health outcomes								
Hemoglobin (g/dL)	1,058	11.632 (1.059)	11.457 (1.043)	0.025**	1,789	11.587 (1.095)	11.445 (1.097)	0.024**
Any anemia	1,058	0.402	0.497	0.002***	1,789	0.421	0.485	0.022**
Mild anemia	1,058	0.159	0.232	0.005***	1,789	0.175	0.207	0.082*
Moderate/ severe anemia	1,058	0.243	0.265	0.425	1,789	0.246	0.278	0.197
Panel B. Child and household level covariates								
Hindu	1,058	0.964	0.975	0.686	1,789	0.970	0.971	0.963
Caste (SC/ST)	1,058	0.237	0.317	0.146	1,789	0.219	0.311	0.068*
Rural	1,058	0.974	0.980	0.750	1,789	0.974	0.986	0.410
Family size	1,058	7.801 (3.315)	7.681 (3.395)	0.622	1,789	7.845 (3.530)	7.662 (3.352)	0.390
Father's years of schooling	1,058	5.316 (4.756)	5.526 (4.855)	0.644	1,789	5.429 (4.848)	5.555 (4.891)	0.731
Mother's years of schooling	1,058	1.634 (3.098)	1.595 (2.980)	0.863	1,789	1.810 (3.275)	1.798 (3.218)	0.956
Asset index	1,058	-0.100 (0.801)	-0.033 (0.989)	0.429	1,789	-0.030 (0.950)	-0.021 (0.999)	0.903
Female	1,058	0.594	0.535	0.062*	1,789	0.542	0.521	0.446
Panel C: School level covariates								
Total enrollment	106	224.231 (169.239)	222.278 (149.302)	0.950	108	220.537 (167.112)	222.278 (149.302)	0.995
Class size	106	29.288 (20.375)	27.519 (12.626)	0.594	108	28.648 (20.280)	27.519 (12.626)	0.729
Student-teacher ratio	106	37.695 (12.608)	33.866 (10.457)	0.092*	108	37.139 (12.744)	33.866 (10.457)	0.147

Notes: Columns 2, 3, 6, and 7 report baseline means by intervention arm for outcomes (Panel A), child and household level covariates (Panel B), and school level covariates (Panel C) in the study analysis. N stands for the number of observations and standard deviations are reported in parentheses. Columns 5 and 8 report p-values from tests on the equality of means for each variable. SC/ST denote Scheduled Caste/Scheduled Tribe. *, ** and *** denote statistical significance at the 10%, 5%, 1% levels. This table is an adaptation of KKV Table 1.

Table 2: Correlation between attrition and pre-treatment characteristics

	Health sample	Cognition & education sample
	(1)	(2)
Treatment	0.546	-0.231
	[0.452]	[0.467]
Hemoglobin (Hb)	-0.018	0.007
	[0.027]	[0.025]
Anemic (Hb < 11.5)	0.025	0.052
	[0.060]	[0.058]
Female	-0.124***	-0.039
	[0.031]	[0.031]
Mother is primary schooled	0.078	0.061
	[0.051]	[0.050]
Hindu	0.089	-0.206
	[0.074]	[0.125]
Above median family size	0.005	0.011
	[0.031]	[0.033]
Total enrollment in school	-0.000	-0.001***
	[0.000]	[0.000]
Student-teacher ratio	0.002	0.007***
	[0.002]	[0.003]
Treat * Anemic (Hb < 11.5)	-0.143*	-0.054
	[0.079]	[0.074]
Treat * Hb	-0.041	-0.014
	[0.036]	[0.036]
Treat * Female	0.093*	0.066
	[0.051]	[0.051]
Treat * Mother is primary schooled	-0.020	-0.044
	[0.074]	[0.070]
Treat * Hindu	-0.159	0.173
	[0.109]	[0.168]
Treat * Above median family size	-0.012	-0.026
	[0.048]	[0.048]
Treat * Total enrollment in school	-0.000	-0.000
	[0.000]	[0.000]
Treat * Student-teacher ratio	0.003	0.006
	[0.003]	[0.004]
Observations	1,789	1,727
p-value from joint F-statistics on the interaction	0.124	0.572

Notes: Coefficients are from the Linear Probability Model for a dummy indicating attrition. Robust standard errors clustered at school levels are in parentheses. *, ** and *** denote statistical significance at the 10%, 5%, 1% levels. Baseline anemic is a dummy variable and coded as one for children with less than 11.5 Hb levels. All models control for asset tercile, class size, and block fixed effects. Coefficients for these control variables are not shown in the table but none of them are significant and are available upon request.

Table 3: Effects of the DFS on anemia and cognitive outcomes (IPW-DD estimates)

Outcomes	Treat*post	Mean of the dependent variable, baseline	Observations
	(1)	(2)	(3)
Panel A: Health outcomes			
Hemoglobin (g/dL)	0.267***	11.632	2,116
	(0.097)		
Any anemia	-0.134***	0.402	2,116
	(0.040)		
Mild Anemia	-0.106***	0.159	2,116
	(0.033)		
Moderate or severe anemia	-0.028	0.243	2,116
	(0.031)		
Panel B: Cognitive outcomes			
Block design	-0.096	3.730	1,616
	(0.116)		
Forward digit-span	-0.002	4.071	1,616
	(0.079)		
Backward digit-span	-0.043	1.076	1,616
	(0.093)		
Raven's Colored Progressive Matrices	0.019	4.766	1,616
	(0.118)		
Stroop test	-0.070	5.116	1,616
	(0.124)		
Cognitive index	-0.027	-0.036	1,616
	(0.092)		
Panel C: Education outcomes			
Math test score	0.083	4.812	1,616
	(0.111)		
Reading test score	0.118	0.918	1,616
	(0.108)		

Notes: Estimated coefficients are based on an inverse probability of attrition weighted double differences (IPW-DD) model estimated separately in each row. All rows include child fixed effects and time-variant household controls (household size, mother's and father's years of schooling, and asset index). The asset index was generated using the first component of a principal component analysis consisting of several household assets. Any anemia is defined as a hemoglobin value < 11.5 g/dL for children aged 5-11 years, < 12g/dL for children aged 12-14 years and girls aged 15 and above, and < 12.9 g/dL for boys aged 15 and above. Mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dL for children aged 5-11 years, value ≥ 11 & < 12g/dL for children aged 12-14 years and girls aged 15 and above, and ≥ 11 & < 12.9 g/dL for boys aged 15 and above. Moderate or severe anemia is defined as a hemoglobin value < 11 g/dL. Outcomes in Panel B and C are normalized with reference to the baseline mean, however, the mean at baseline is reported without normalization. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively. Standard errors, clustered at the school level, are reported in parentheses.

Table 4: Heterogeneous treatment effects on health outcomes, by baseline attendance rate

	Hemoglobin (g/dL)	Any anemia	Mild anemia	Moderate or severe anemia
	(1)	(2)	(3)	(4)
Panel A: High attendance levels				
70% attendance				
Treat*post	0.290***	-0.170***	-0.115***	-0.055
	(0.105)	(0.042)	(0.033)	(0.035)
Observations	1642	1642	1642	1642
80% attendance				
Treat*post	0.255**	-0.181***	-0.136***	-0.045
	(0.124)	(0.049)	(0.040)	(0.039)
Observations	1226	1226	1226	1226
90% attendance				
Treat*post	0.271	-0.215***	-0.146**	-0.068
	(0.171)	(0.073)	(0.059)	(0.058)
Observations	594	594	594	594
Panel B: Attendance terciles				
Bottom tercile				
Treat*post	0.292	-0.071	-0.080	0.010
	(0.187)	(0.074)	(0.060)	(0.066)
Observations	602	602	602	602
Middle tercile				
Treat*post	0.349***	-0.181***	-0.106**	-0.074*
	(0.120)	(0.057)	(0.049)	(0.043)
Observations	714	714	714	714
Top tercile				
Treat*post	0.250	-0.190***	-0.129**	-0.061
	(0.157)	(0.066)	(0.055)	(0.052)
Observations	702	702	702	702

Notes: Each cell reports the DD coefficients from a separate regression. Standard errors, clustered at the school level, are reported in parentheses. Any anemia is defined as a hemoglobin value < 11.5 g/dL for children aged 5-11 years, < 12g/dL for children aged 12-14 years and girls aged 15 and above, and < 12.9 g/dL for boys aged 15 and above. Mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dL for children aged 5-11 years, value ≥ 11 & < 12g/dL for children aged 12-14 years and girls aged 15 and above, and ≥ 11 & < 12.9 g/dL for boys aged 15 and above. Moderate or severe anemia is defined as a hemoglobin value < 11 g/dL. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table 5: Robustness to exclusion or inclusion of control variables

	Hemoglobin (g/dL)	Any anemia	Mild anemia	Moderate or severe anemia
	(1)	(2)	(3)	(4)
Panel A: Without control variables				
Treat*post	0.284***	-0.129***	-0.091***	-0.038
	(0.104)	(0.039)	(0.030)	(0.031)
Mean of dependent variable, baseline	11.638	0.403	0.166	0.237
Observations	2,468	2,468	2,468	2,468
Panel B: With IFA control				
Treat*post	0.242**	-0.162***	-0.139***	-0.022
	(0.117)	(0.046)	(0.043)	(0.036)
Mean of dependent variable, baseline	11.658	0.380	0.143	0.237
Observations	1,414	1,414	1,414	1,414
Panel C: With time-variant school-level controls				
Treat*post	0.195	-0.120**	-0.113**	-0.007
	(0.119)	(0.048)	(0.044)	(0.036)
Mean of dependent variable, baseline	11.640	0.386	0.145	0.240
Observations	1,502	1,502	1,502	1,502
Panel D: Lee bounds				
Treat	0.032	-0.019	-0.021	0.001
	(0.056)	(0.022)	(0.019)	(0.014)
Lower bound	-0.022	-0.068*	-0.075**	-0.045
	(0.085)	(0.036)	(0.036)	(0.035)
Upper bound	0.216**	-0.016	-0.023	0.007
	(0.089)	(0.026)	(0.022)	(0.016)
Observations	1,234	1,234	1,234	1,234

Notes: Any anemia is defined as a hemoglobin value < 11.5 g/dL for children aged 5-11 years, < 12g/dL for children aged 12-14 years and girls aged 15 and above, and < 12.9 g/dL for boys aged 15 and above. Mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dL for children aged 5-11 years, value ≥ 11 & < 12g/dL for children aged 12-14 years and girls aged 15 and above, and ≥ 11 & < 12.9 g/dL for boys aged 15 and above. Moderate or severe anemia is defined as a hemoglobin value < 11 g/dL. Panel A to C use an inverse probability of attrition weighted double differences (IPW-DD) model including child fixed effects and clustered standard errors at the school level. Panel B and C include time-variant household controls reported in Table 3. Panel C additionally includes time-variant school level controls: number of children enrolled in school, class size, and student-teacher ratio. Panel D shows results of an OLS regression with endline data only and Lee (2009) bounds estimated using the STATA command by Tauchmann (2014). *, **, *** denote significance at the 10%, 5% and 1% levels, respectively. Standard errors are reported in parentheses.

Appendix

A: Dosage of iron in the DFS

Children aged 4 to 6 years require 8 mg of daily iron intake, children aged 7 to 9 years 10 mg, children aged 10 to 12 years 12 mg, and children aged 13 to 19 years 15 mg (WHO, 1959). One 450 kcal meal portion for primary school children (first to fifth grade) should contain 4 g of DFS per meal, i.e. 3.5 mg of iron (Krämer et al.,2020; Bihar Government, 2017). Upper primary school children (from sixth grade onward) should receive a larger portion 700 kcal, so 6.222 g of DFS, i.e. 5.444 mg of iron (Bihar Government, 2017). Thus, the iron intake due to DFS accounts for 43.75 percent of the required daily iron intake for children between age 4 and 6 years, 35 percent for children between age 7 to 9 years, and 29.167 percent for children between age 10 and 12 years considering the size of a meal for the first to the fifth grade. Considering the portion size at sixth grade, the iron intake due to DFS accounts for 45.367 percent of the required daily iron intake for children between age 10 and 12.

Table A1: Project Timeline

Year	2014		2015		2016		2017			2018		2019	
Children in grade	I	II	II	III	III	IV	IV	V	V	VI	VI		VII
Intervention													
Delivery to treatment schools				■	■	■	■	■	■	■	■	■	■
Delivery to control schools								■	■	■	■	■	■
Data collection													
Survey I		■	■										
Survey II						■							
Survey III												■	

Table A2: Variation in treatment exposure of children

Year	2014		2015		2016		2017			2018		2019	
Children in grade	I	II	II	III	III	IV	IV	V	V	VI	VI		VII
Treatment group													
Stayed at sample school				■	■	■	■	■	■	■	■	■	■
Left sample school				■	■	■	■	■	■				
Control group													
Stayed at sample school								■	■	■	■	■	■
Left sample school								■					

Notes: Dark-shaded cells indicate treatment exposure.

Table A3: Cognitive tests

Test	Original Source	Cognitive ability	Executive function
Digit span forward	Malin's Intelligence Scale for Indian Children (Malin, 1969)	- Short-term auditory memory - Simple verbal expression	
Digit span backward	Malin's Intelligence Scale for Indian Children (Malin, 1969)	- Store, use and manipulate new information - Attention - Impulse control - Shifting	x
Block design	Malin's Intelligence Scale for Indian Children (Malin, 1969)	- Planning and organizing	x
Stroop-like day-and-night test for wave 1 & 2	Gerstadt *et al*. (1994)	- Inhibition - Memorizing two rules simultaneously	x
Fruit/vegetable Stroop for wave 3	Röthlisberger *et al*. (2010)		
Raven's colored progressive matrices	Raven *et al*. (1998)	- Abstract reasoning - Capacity to simultaneously solve several problems involving new information	x

Table A4: MDE for different outcomes and different % of take-up for wave 3 data

	SD	Baseline ICC	MDE (c = 0.6)	MDE (c = 0.7)	MDE (c= 0.8)	Values from simple reg
Block design	0.62	0.06	0.21	0.17	0.15	0.03
Digit span forwards	0.93	0.01	0.25	0.21	0.18	0.01
Digit span backwards	0.77	0.02	0.21	0.18	0.16	0.01
Progressive Matrices	1.11	0.04	0.35	0.30	0.26	0.03
Stroop	1.08	0.02	0.30	0.25	0.22	-0.05
Cognition Index	0.99	0.06	0.33	0.28	0.24	0.02
Math	1.22	0.12	0.51	0.43	0.37	0.06
Reading	1.35	0.13	0.56	0.48	0.41	0.05

Notes: SD: Standard deviation. ICC: Intercluster correlation. MDE: Minimal detectable effect. C: Take up rate. Assumptions: Sample size: 1712, number of clusters: 107 schools, deviation of observations between treatment and control: 50:50. Hemoglobin is expressed in g/dl and the different forms of anemia represent percentage points). Cognition and education outcomes are normalized with respect to the control group mean and standard deviation.

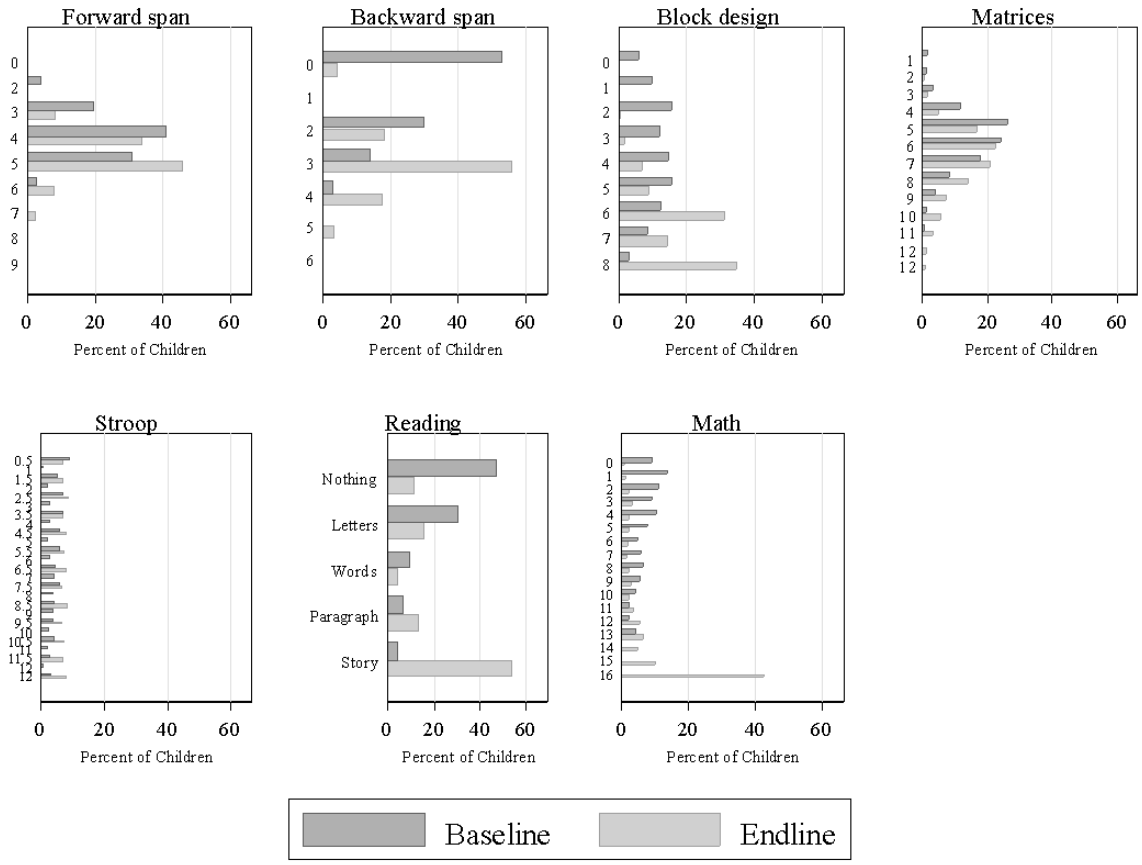

Figure A1: Distribution of cognition and education scores of school children by wave

Notes: The sample is the analytical sample with 820 observations per wave. In the following, we report the variance for baseline (I) and endline (III) for all cognition and education scores: forward span 0.886 (I) and 0.834 (III), backward span 1.712 (I) and 0.973 (III), block design 4.676 (I) and 1.850 (III), matrices 2.649 (I) and 3.642 (III), Stroop 11.814 (I) and 13.942 (III), reading 1.316 (I) and 2.267 (III), and math 14.141 (I) and 20.409 (III).

Table A5: Balance in samples before and after attrition using baseline data for cognition and education sample

Baseline sample	After attrition				Before attrition			
	N	Control means	Treatment means	p-values	N	Control means	Treatment means	p-values
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Child level outcome variables								
Health outcomes								
Hemoglobin (g/dL)	808	11.613	11.455	0.072*	1,727	11.597	11.439	0.012**
		(1.046)	(1.117)			(1.096)	(1.104)	
Any anemia	808	0.398	0.485	0.021**	1,727	0.421	0.486	0.019**
Mild anemia	808	0.147	0.204	0.044**	1,727	0.178	0.206	0.122
Moderate/severe anemia	808	0.251	0.281	0.368	1,727	0.243	0.280	0.138
Cognition outcomes								
Block design	808	3.730	3.921	0.416	1,770	3.690	3.801	0.566
		(2.178)	(2.148)			(2.254)	(2.188)	
Digit-span forward	808	4.071	4.120	0.541	1.770	4.068	4.094	0.719
		(0.909)	(0.968)			(1.016)	(0.995)	
Digit-span backward	808	1.076	1.245	0.200	1,770	1.105	1.142	0.720
		(1.330)	(1.288)			(1.294)	(1.304)	
Progressive matrices	808	4.766	4.864	0.573	1,770	4.815	4.687	0.371
		(1.582)	(1.665)			(1.655)	(1.711)	
Stroop test	808	5.116	5.372	0.437	1,770	5.462	5.271	0.475
		(3.524)	(3.363)			(3.488)	(3.370)	
Cognitive index	808	-0.036	0.082	0.256	1,770	0.002	-0.003	0.954
		(0.982)	(0.945)			(1.022)	(0.979)	
Education outcomes								
Math score	808	4.812	5.052	0.615	1,770	4.910	4.749	0.687
		(3.762)	(3.760)			(3.860)	(3.798)	
Reading score	808	0.918	0.941	0.852	1,770	0.947	0.871	0.446
		(1.161)	(1.137)			(1.157)	(1.102)	
Panel B. Child and household level covariates								
Hindu	808	0.967	0.968	0.975	1,770	0.968	0.971	0.913

Scheduled Caste/Scheduled Tribe	808	0.267	0.320	0.417	1,770	0.250	0.315	0.221
Rural	808	0.970	0.982	0.539	1,770	0.974	0.983	0.538
Family size	808	7.796	7.662	0.620	1,770	7.787	7.678	0.595
		(3.191)	(3.260)			(3.405)	(3.307)	
Father's years of schooling	808	5.272	5.569	0.570	1,770	5.377	5.521	0.701
		(4.758)	(4.800)			(4.816)	(4.870)	
Mother's years of schooling	808	1.599	1.658	0.830	1,770	1.800	1.779	0.920
		(3.084)	(3.037)			(3.264)	(3.222)	
Asset index	808	-0.036	-0.049	0.897	1,770	-0.008	-0.037	0.710
		(0.860)	(0.962)			(0.962)	(0.987)	
Female	808	0.578	0.533	0.224	1,770	0.550	0.540	0.726
Panel C: School level covariates								
Number of children enrolled in school	92	243.341	234.292	0.794	108	220.537	222.278	0.955
		(176.597)	(152.709)			(167.112)	(149.302)	
Class size	92	29.432	28.208	0.744	108	28.648	27.519	0.729
		(21.641)	(12.647)			(20.280)	(12.626)	
Student-teacher ratio	92	36.960	33.655	0.144	108	37.139	33.866	0.147
		(11.589)	(9.716)			(12.744)	(10.457)	

Notes: Columns 2,3,6, and 7 report baseline means by intervention arm for outcomes (Panel A), child and household level covariates (Panel B), and school level covariates (Panel C) in the study analysis. N stands for the number of observations and standard deviations are reported in parentheses. Columns 5 and 8 report p-values from tests on the equality of means for each variable. SC/ST denote Scheduled Caste/Scheduled Tribe. *, ** and *** denote statistical significance at the 10%, 5%, 1% levels.

Table A6: Heterogeneous treatment effects on cognition and education outcomes, by baseline attendance rate

	Block design	Digit-span forward	Digit-span backward	Progressive matrices	Stroop tests	Cognitive index	Math test score	Reading test score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: High attendance levels								
Treat*post (70% attendance)	-0.055	-0.060	-0.091	0.091	0.044	-0.028	0.111	0.062
	(0.133)	(0.089)	(0.105)	(0.131)	(0.142)	(0.105)	(0.120)	(0.111)
Observations	1266	1266	1266	1266	1266	1266	1266	1266
Treat*post (80% attendance)	0.011	-0.010	-0.057	0.084	0.055	0.015	0.132	0.052
	(0.151)	(0.096)	(0.119)	(0.131)	(0.135)	(0.121)	(0.133)	(0.126)
Observations	966	966	966	966	966	966	966	966
Treat*post (90% attendance)	0.051	0.031	-0.069	0.289	0.153	0.157	0.152	0.130
	(0.184)	(0.137)	(0.146)	(0.192)	(0.194)	(0.155)	(0.191)	(0.172)
Observations	482	482	482	482	482	482	482	482
Panel B: Attendance terciles								
Treat*post (Bottom tercile)	-0.298	-0.070	0.070	-0.076	-0.253	-0.103	-0.160	0.188
	(0.186)	(0.165)	(0.188)	(0.212)	(0.193)	(0.158)	(0.190)	(0.170)
Observations	470	470	470	470	470	470	470	470
Treat*post (Middle tercile)	-0.127	-0.018	-0.225	-0.019	-0.073	-0.165	0.108	-0.075
	(0.165)	(0.128)	(0.175)	(0.180)	(0.170)	(0.138)	(0.141)	(0.173)
Observations	516	516	516	516	516	516	516	516
Treat*post (Top tercile)	0.086	0.014	-0.001	0.201	0.142	0.143	0.178	0.144
	(0.169)	(0.124)	(0.138)	(0.175)	(0.190)	(0.134)	(0.17)	(0.154)
Observations	568	568	568	568	568	568	568	568

Notes: Each cell DD coefficients from a separate regression report based on an inverse probability of attrition weighted double differences (IPW-DD) model. Standard errors, clustered at the school level, are reported in parentheses. All outcomes are normalized with reference to the baseline mean. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A7: Heterogeneous treatment effects on health outcomes, by gender and baseline anemic status

	Hemoglobin (g/dL)	Any anemia	Mild anemia	Moderate or severe anemia
	(1)	(2)	(3)	(4)
Panel A: Gender				
Female				
Treat*post	0.097	-0.080*	-0.079*	-0.001
	(0.120)	(0.047)	(0.042)	(0.040)
Observations	1192	1192	1192	1192
Male				
Treat*post	0.445***	-0.197***	-0.139***	-0.059
	(0.136)	(0.058)	(0.052)	(0.044)
Observations	910	910	910	910
Panel B: Baseline anemic status				
Anemic (Hb < 11.5 g/dL)				
Treat*post	-0.242	0.071	-0.083	0.154**
	(0.179)	(0.071)	(0.080)	(0.077)
Observations	674	674	674	674
Moderately anemic (Hb < 11 g/dL)				
Treat*post	-0.050	0.133	0.225**	-0.092
	(0.288)	(0.113)	(0.098)	(0.082)
Observations	346	346	346	346
Non-anemic (Hb > 11.5 g/dL)				
Treat*post	0.257*	-0.098**	-0.071**	-0.027
	(0.133)	(0.039)	(0.030)	(0.023)
Observations	1051	1051	1051	1051

Notes: Each cell reports the DD coefficients from a separate regression. Standard errors, clustered at the school level, are reported in parentheses. The definitions for the used outcomes are reported in Table 4. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A8: Heterogeneous treatment effects on cognition and education outcomes, by gender and baseline anemic status

	Block design	Digit-span forward	Digit-span backward	Progressive matrices	Stroop tests	Cognitive index	Math test score	Reading test score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Gender								
Treat*post (Female)	-0.178	-0.015	-0.162	-0.127	-0.051	-0.152	0.119	0.047
	(0.147)	(0.110)	(0.119)	(0.138)	(0.137)	(0.119)	(0.139)	(0.140)
Observations	894	894	894	894	894	894	894	894
Treat*post (Male)	-0.005	0.027	0.108	0.178	-0.082	0.123	0.040	0.194
	(0.129)	(0.108)	(0.141)	(0.172)	(0.179)	(0.113)	(0.138)	(0.136)
Observations	722	722	722	722	722	722	722	722
Panel B: Baseline anemic status								
Treat*post (Anemic (Hb < 11.5 g/dL))	-0.274	0.063	-0.389**	-0.013	0.122	-0.189	0.226	0.093
	(0.236)	(0.208)	(0.180)	(0.246)	(0.257)	(0.176)	(0.220)	(0.217)
Observations	496	496	496	496	496	496	496	496
Treat*post (Moderately anemic (Hb < 11 g/dL))	-0.437	0.093	-0.622**	-0.030	-0.486	-0.335	-0.249	-0.058
	(0.267)	(0.407)	(0.308)	(0.309)	(0.333)	(0.264)	(0.296)	(0.332)
Observations	273	273	273	273	273	273	273	273
Treat*post (Non-anemic (Hb > 11.5 g/dL))	-0.001	0.002	0.029	0.186	0.165	0.079	0.059	0.117
	(0.160)	(0.127)	(0.135)	(0.158)	(0.165)	(0.145)	(0.152)	(0.154)
Observations	820	820	820	820	820	820	820	820

Notes: Each cell DD coefficients from a separate regression report based on an inverse probability of attrition weighted double differences (IPW-DD) model. Standard errors, clustered at the school level, are reported in parentheses. All outcomes are normalized with reference to the baseline mean. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A9: School-level heterogeneous treatment effects on health outcomes, by baseline school attendance rate

	Hemoglobin (g/dL)	Any anemia	Mild anemia	Moderate or severe anemia
	(1)	(2)	(3)	(4)
Panel A: Above mean school attendance				
Treat*post	0.427***	-0.208***	-0.138***	-0.070
	(0.155)	(0.053)	(0.037)	(0.048)
Observations	960	960	960	960
Panel B: Below mean school attendance				
Treat*post	0.195	-0.079	-0.073	-0.006
	(0.138)	(0.058)	(0.052)	(0.040)
Observations	1156	1156	1156	1156

Notes: Each cell reports the DD coefficients from a separate regression. Standard errors, clustered at the school level, are reported in parentheses. The definitions for the used outcomes are reported in Table 4. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A10: School-level heterogeneous treatment effects on cognition and education outcomes, by baseline school attendance rate

	Block design	Digit-span forward	Digit-span backward	Progressive matrices	Stroop tests	Cognitive index	Math test score	Reading test score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Above mean school attendance								
Treat*post	-0.100	0.016	-0.123	0.110	0.245	0.046	0.020	0.122
	(0.186)	(0.111)	(0.135)	(0.155)	(0.181)	(0.140)	(0.152)	(0.147)
Observations	738	738	738	738	738	738	738	738
Panel B: Below mean school attendance								
Treat*post	-0.069	-0.063	0.036	-0.072	-0.335**	-0.091	0.097	0.093
	(0.141)	(0.112)	(0.128)	(0.184)	(0.163)	(0.132)	(0.158)	(0.154)
Observations	878	878	878	878	878	878	878	878

Notes: Each cell DD coefficients from a separate regression report based on an inverse probability of attrition weighted double differences (IPW-DD) model. Standard errors, clustered at the school level, are reported in parentheses. All outcomes are normalized with reference to the baseline mean. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A11: Multiple hypothesis corrected p-values of the DFS on anemia and cognitive outcomes (IPW-DD estimates)

Outcomes	Treat*post	Naïve p-value	Multiple hypothesis corrected sharpened q-values
	(1)	(2)	(3)
Panel A: Health outcomes			
Hemoglobin (g/dL)	0.245**	0.007	0.005
	(0.094)		
Any anemia	-0.134***	0.001	0.005
	(0.039)		
Mild Anemia	-0.106***	0.002	0.005
	(0.033)		
Moderate or severe anemia	-0.027	0.361	0.100
	(0.029)		
Panel B: Cognitive outcomes			
Block design	-0.096	0.411	1.000
	(0.116)		
Forward digit-span	-0.002	0.984	1.000
	(0.079)		
Backward digit-span	-0.043	0.648	1.000
	(0.093)		
Raven's Colored Progressive Matrices	0.019	0.873	1.000
	(0.118)		
Stroop test	-0.070	0.576	1.000
	(0.124)		
Panel C: Education outcomes			
Math test score	0.083	0.458	1.000
	(0.111)		
Reading test score	0.118	0.278	1.000
	(0.108)		

Notes: Estimated coefficients are based on an inverse probability of attrition weighted double differences (IPW-DD) model estimated separately in each row. All rows include child fixed effects and time-variant household controls (household size, mother's and father's years of schooling, and asset index). The asset index was generated using the first component of a principal component analysis consisting of several household assets. Outcomes in Panel B and C are normalized with reference to the baseline mean. *, **, *** denote significance at the 10%, 5% and 1% levels for naïve p-values, respectively. Standard errors, clustered at the school level, are reported in parentheses. We omit the cognitive index because it is already a way of aggregating data to prevent multiple hypotheses testing.

Table A12: Multiple hypothesis corrected p-values of heterogeneous treatment effects on health outcomes, by baseline attendance rate

	Hemoglobin (g/dL)	Any anemia	Mild anemia	Moderate or severe anemia
	(1)	(2)	(3)	(4)
Panel A: High attendance levels				
70% attendance				
Treat*post	0.290***	-0.170***	-0.115***	-0.055
	(0.105)	(0.042)	(0.033)	(0.035)
	[0.007]	[0.000]	[0.001]	[0.122]
	{0.016}	{0.003}	{0.006}	{0.085}
Observations	1642	1642	1642	1642
80% attendance				
Treat*post	0.255**	-0.181***	-0.136***	-0.045
	(0.124)	(0.049)	(0.040)	(0.039)
	[0.042]	[0.000]	[0.001]	[0.250]
	{0.048}	{0.004}	{0.006}	{0.143}
Observations	1226	1226	1226	1226
90% attendance				
Treat*post	0.271	-0.215***	-0.146**	-0.068
	(0.171)	(0.073)	(0.059)	(0.058)
	[0.117]	[0.004]	[0.015]	[0.241]
	{0.085}	{0.012}	{0.024}	{0.143}
Observations	594	594	594	594
Panel B: Attendance terciles				
Bottom tercile				
Treat*post	0.292	-0.071	-0.080	0.010
	(0.187)	(0.074)	(0.060)	(0.066)
	[0.123]	[0.345]	[0.186]	[0.885]
	{0.085}	{0.181}	{0.121}	{0.285}
Observations	602	602	602	602
Middle tercile				
Treat*post	0.349***	-0.181***	-0.106**	-0.074*
	(0.120)	(0.057)	(0.049)	(0.043)
	[0.005]	[0.002]	[0.033]	[0.089]
	{0.012}	{0.009}	{0.004}	{0.083}
Observations	714	714	714	714
Top tercile				
Treat*post	0.250	-0.190***	-0.129**	-0.061
	(0.157)	(0.066)	(0.055)	(0.052)
	[0.114]	[0.005]	[0.021]	[0.248]
	{0.085}	{0.012}	{0.031}	{0.143}
Observations	702	702	702	702

Notes: Each cell reports the DD coefficients from a separate regression. Standard errors, clustered at the school level, are reported in parentheses. Naïve p-Values and multiple hypothesis corrected sharpened q-values are reported in brackets and curly brackets respectively. We correct the standard errors across all outcomes for the following bundles: subgroups for endline school attendance (70,80, and 90 attendance level as well as bottom, middle, and top tercile), child characteristics (female and male and anemic, moderately anemic, and non-anemic), and school quality (above and below mean school attendance). The definitions of the health outcomes are reported in Table 3. All regressions include child fixed effects and time-variant household controls reported in Table 3. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.