Can Iron-Fortified Salt Control Anemia?

Evidence from Two Experiments in Rural Bihar

Abhijit Banerjee, Sharon Barnhardt, Esther Duflo *

December 18, 2017

Abstract

Iron deficiency anemia is frequent among the poor worldwide. While it can be prevented with the appropriate supplement or food fortification, these programs struggle to reach the poorest, out-of-school children, men and the elderly. This paper reports on the impact of a potential strategy to address iron deficiency anemia in rural areas: salt fortified with iron and iodine (DFS). We conduct a large-scale experiment in 400 villages in Bihar to test the impacts of both selling DFS and giving it away for free. At baseline, 45% of the sample is anemic. In 200 randomly assigned villages, we introduce DFS for the first time at half the normal retail price and sell it for roughly 30 months. In 62 of those sales villages, we deliver DFS for free directly to 7 randomly assigned households over nearly 24 months. We find no evidence that either selling DFS or providing it for free has an economically meaningful or statistically significant impact on hemoglobin, anemia, physical health, cognition or mental health. For the sales experiment, we can reject a reduction of 2.4 percent in the fraction anemic in the entire sample, and 1 percent among those who were previously anemic. Using an IV strategy, we find a statistically significant, though relatively small, increase in hemoglobin and reduction in the fraction anemic for adolescents, a subgroup that responded well to supplements and fortification in earlier studies. These disappointing results are explained both by modest purchases and low impact of DFS for the majority of the population, even when consumed somewhat regularly.

*Banerjee and Duflo are from The Massachusetts Institute of Technology, Barnhardt is from the Centre for Experimental Social Sciences, CESS Nuffield – FLAME University. For financial support, we thank the U.K. Department for International Development, the International Initiative for Impact Evaluation, the U.S. National Institutes of Health (P01AG005842), the International Food Policy Research Institute, and the MIT Department of Economics. For dedicated research assistance we are grateful to the team at J-PAL, including Urmi Bhattacharya, Shruti Bhimsaria, Anna George, Dwijo Goswami, Radhika Jain, Seema Kacker, Sweta Kumari, Francine Loza, Bastien Michel, Krutika Ravishankar, Prianthi Roy, Achill Rudolph, Laura Stilwell, Niloufer Taber, Srinivasan Vasudevan, Micah Villareal, Mark Walsh, and dozens of survey specialists. Dr. Vandana Sharma provided medical expertise. Our partners at Tata Chemicals Limited made the distribution of DFS possible in rural Bihar. The National Institute of Nutrition (Hyderabad) provided ongoing assistance with the verification of iron and iodine content in salt samples. These studies can be found in the AEA RCT Registry (AEARCTR-0000010, AEARCTR-0000013). We thank Nel Druce for detailed comments on a DFID report that underlies the analysis in this paper. All errors are our own and the paper does not represent the views of DFID or any other organization.

1 Introduction

Iron deficiency is believed to be the most common nutrient deficiency in the world today. While quantifying the number of affected people is difficult, the WHO estimates that 1.6 billion people are anemic (de Benoist et al., 2008), and that about half of these cases can be traced to iron deficiency (World Health Organization, 2001). Iron deficiency anemia (IDA) is more common among populations with a diet low in animal proteins, and high in rice or in whole wheat with high phytate content, because phytates reduce absorption (Zijp et al., 2000). In Asia and in South Asia, where this type of diet is common, IDA is a particularly serious issue. In Indonesia, a large scale study of iron supplements found that 50 percent of women aged 15 and above and 40 percent of men sampled were anemic at baseline (Thomas et al., 2003).[1] Anemia rates are similar in our study villages where 45% of individuals had some form of anemia at baseline. The rate was higher than average for females (54%) and the elderly (55%), near the average for children under 15 years old (42%) and adults (41%) and below average - though still high - for males (35%) and adolescents aged 13 to 17 years (38%).

While we can find no national or Bihar-level statistics on the fraction of anemia estimated to be caused by iron deficiency, the reasons to focus on iron are clear. The causes of anemia can be grouped into four categories: genetic, environment, infectious and parasitic disease, and nutrition. The genetic condition *Beta Thalassemia* is responsible for anemia, but it has been measured in Bihar at only 3.4% of the population (Nagar et al., 2015). Poor sanitation, infectious diseases and parasites are also causes of anemia as iron absorption is reduced. In our sample, only 2.3% reported having parasites in the six months before our endline and deworming has now become standard for school children in Bihar. In our study sample, 24% have taken deworming medication, but only 57% of those had done so in the past year. Further, malaria is not endemic in Bihar with only 0.1 cases reported for every 1000 persons (Caravotta, 2009).

Alternatively, the indicators for nutrition suggest it as the leading cause of anemia. In rural Bihar, 31.8% of women have such poor general nutrition that their BMI is below normal (18.5 kg/m^2) and 26.9% of rural men fall in the same category (International Institute for Population Sciences (IIPS), 2017). Also according to National Family Health Survey-4 data, nearly half of children in rural Bihar under the age of 5 are likely to be stunted (49.3%) or

[1] Using the World Health Organization thresholds of 12g/dL for women and 13g/dL for men.

underweight (44.6%) for their age. Our study sample has a mostly vegetarian diet due to cost rather than preferences. The average household spends less than 1 USD per month per person on eggs, fish and meat combined. The main sources of iron are plant-based, non-haem and less readily absorbed. Diets are also high in components that bind to bioavailable iron and prevent its absorption, including tea and rice. Finally, our baseline analysis (Banerjee et al., 2013) established a correlation between Household Dietary Diversity Score (diet quality) and anemia within our study sample. Therefore, we not only focus on iron deficiency as the main cause of anemia in our study area, but also draw confidence when interpreting results that they are not due to anemia in the region being caused predominantly by genetic, environmental or disease factors.

IDA increases susceptibility to infection and increases the likelihood of experiencing weakness or fatigue (see Haas and Brownlie (2001) for a review of the medical evidence). It has been linked to low productivity in adults and slowing of cognitive and physical growth among children (Lozoff, 2007; Lozoff et al., 2006). Among pregnant women, severe anemia can result in low birth weight and child mortality (Stoltzfus, 2001). For the elderly, lower hemoglobin levels are associated with cognitive decline (Peters et al., 2008) and worse physical performance (Penninx et al., 2004).

While iron deficiency anemia has been recognized to be a serious public health problem in lower income countries for several years, not much progress has been made against it. The alternatives to control it are diet change, supplements and food fortification. The evidence on dietary diversification (and home gardening) interventions is insufficient (Bhutta et al., 2008). The WHO recommends that countries with endemic anemia adopt costly, large-scale, weekly iron and folic acid supplementation programs for children and adolescents, and several countries have embarked on such efforts over the past decades (Chong et al., 2015).

The largest and most comprehensive study to date of a supplementation program is Thomas et al. (2006) which found a large effect of iron supplements (where each individual was given 120 milligram of iron per week) on anemia and hemoglobin rates for those who were anemic at baseline. That study also found increases in the labor supply of males who were anemic at baseline, and an increase in the earnings of self-employed males, as well as improved health (including mental health). Chong et al. (2015) set up a randomized encouragement intervention among 219 adolescents in Peru to measure the impact of such programs on test scores. Iron pills

were made available free at a local health center, and randomly selected students were exposed to two versions of a short video that encouraged them to get the supplements. On average, students in the treatment groups received 80% more pills than students in the control group over 10 weeks and were about twice as likely to receive at least 600 mg over 10 weeks (about a third of the recommended dietary allowance).They found large reductions in anemia among treatment students who were previously anemic, and improvements in both cognitive tests and performance at school.

These interventions suggest that iron supplements, if regularly taken, can be effective. However, Thomas et al. (2006) spent considerable resources insuring consistent take up. In Chong et al. (2015) few adolescents took supplements, which were available free in the local clinic, without additional information (the control group).[2] While systematic supplementation seems like a feasible strategy in health centers or schools, this method faces issues with non-compliance, as well as supply, procurement, and distribution, since many countries lack the public health infrastructure to handle the logistics (Gillespie, 1998). Outside of these settings, systematic distribution of iron supplements may not be a practical policy in resource-poor settings, where the public health systems do not have the capacity to distribute these supplements reliably on a large scale.

The reasoning behind the potential for food fortification is simple. The first premise is that, contrary to a change in diet or regular consumption of supplements, consuming fortified food does not require a major behavioral change, and therefore should be easy for individuals. Hence, the take up of fortified food could easily be very high. The second premise is that regular intake would compensate for the fact that individuals would get less of a micronutrient from a fortified foods than from supplements.

Iron fortification of foods has generated considerable policy excitement. It requires no additional effort on the part of the consumer and can be done relatively cheaply in centralized locations. Foods that can be fortified with iron include flour, milk products, fish sauce, and salt. In particular, salt seems to be an ideal product to fortify: it is ubiquitous, cheap, and generally purchased from stores. Adding iron to packaged iodized salt thus seems to be a promising way to increase iron intake and reduce IDA.

[2] Even with the encouragement of several reminder videos, only 57% of the adolescents in the treatment groups took a dose considered to be sufficient to affect hemoglobin levels.

Double-fortified salt (DFS) was not commercially available in India until recently, due to technical difficulties ensuring the stability of both the iron and the iodine. In the mid 1990s, India's National Institute of Nutrition (NIN, Hyderabad) developed DFS, containing both iron and iodine. Fortified with 1 mg iron per gram of salt, DFS is estimated to provide about 30% of the RDA of iron (National Institute of Nutrition (India), 2005) when consuming 10 g salt per day regularly (Ranganathan and Sesikeran, 2008). A recent systematic review found that the mean adult in India consumes about 11 grams of salt per day (Johnson et al., 2017). NIN scientists first demonstrated the long-term safety of DFS in animal studies (Nair et al., 1998). They also established the stability and bioavailability of iron in DFS as well as the acceptability and effectiveness of DFS with school children and with tribal populations through small-scale trials (Nair et al., 1998; Brahmam et al., 2000; Sivakumar et al., 2001).

In the last five years, NIN and the Indian Government have sought to encourage wider adoption of DFS. Since 2011, the NIN formulation of DFS can be manufactured by private companies through a license agreement requiring a certain percentage of production to be donated to charities, such as school meal programs. Several manufacturers produce and market DFS, including Tata Chemicals Limited (TCL), one of the leading manufacturers of salt in India.

In 2012, India's Department of Women and Child Development directed states to use DFS in the national mid-day meal scheme (school lunches) and the Human Resource Development Ministry did the same for the Integrated Child Development Scheme (Mudur, 2013). A nationwide policy to use DFS in school meals was put in place despite the lack of any large-scale trials of DFS.

There are two existing small-scale trials of NIN's DFS in India (Sivakumar et al., 2001) conducted by NIN researchers plus one by external researchers. The first NIN study finds no impact on hemoglobin across age groups in a tribal area of Andhra Pradesh.[3] The second NIN study gives DFS or iodized salt to residential schoolchildren, using a school-level randomization. It finds a decrease of .042 g/dL for the DFS group, a smaller decline than the 1.13 g/dL decrease in the iodized group. However, there are only four schools randomized into treatment and control and the study suffers from 78.6% attrition. Reddy and Nair (2014) gives DFS for free to the families of 6 to 15-year-olds. They find hemoglobin increases of 0.21 g/dL (DFS alone) and

[3] It does find an increase in hemoglobin of 1.3 g/dL more for males age 14-17 in the DFS group versus the control, but starts with only 22 control individuals and loses nearly two thirds of them by the end of two years.

0.60 g/dL (DFS plus deworming) versus hemoglobin declines among children not given DFS. However, the randomization is clustered at the school level, with only two schools randomized into DFS and the analysis does not adjust the standard errors for this design choice.[4]

Another three trials in India study other formulations of iron-fortified salt. Andersson et al. (2008) uses two types of iron-fortified salt. They are higher dose than NIN's DFS at 2 mg of iron per gram of salt, and one has encapsulated iron while the other uses a stabilizer technology.[5] The encapsulated, fortified salt increases hemoglobin by 0.2 g/dL over 10 months in 5-15 year old schoolchildren who are also given Vitamin A supplements and whose households are directly reminded every two months to use the fortified salt. The stabilized, fortified salt does not product a statistically significant effect on hemoglobin under the same conditions.

Haas et al. (2014) studies female teapickers and Rajagopalan and Vinodkumar (2000) studies male and female teapickers in India and find statistically significant increases in hemoglobin of 0.24 to 0.72 g/dL. However, the subjects' location on remote tea estates means no other packaged salt was available. There were simply no competing brands or lower quality (less expensive) salt to use. Both the teapicker studies and the residential school trials look at captive audiences, which is not the desirable sample for policy questions.

Finally, to our knowledge, four small trials (with less than 400 individuals each) have been conducted with other formulas of iron-fortified salts in other low or middle-income countries.[6] Their results are mixed, with one trial producing a null result for schoolchildren (Wegmüller et al., 2006), one trial having a null result for children but an increase (0.4 g/dL) in hemoglobin for women (Asibey-Berko et al., 2008), and two trials producing positive findings (+0.9 g/dL and

[4] Further, the average child in Reddy and Nair (2014) started with a weight of 20 kg, BMI of 13.5 and hemoglobin level of about 8.5 in the DFS plus deworming school and 8.9 in the DFS school. This means the average child was on the low end of the moderately anemic range (8 to 10.9 g/dL). Our baseline sample was much better nourished; the mean girl had a BMI above 15 and hemoglobin level of 11.8 g/dL while the mean boy had a BMI of 15.9 and hemoglobin level of 12.2.

[5] The use of a stabilizer or encapsulation is to prevent iodine decay in salt containing both iodine and iron. NIN's DFS uses a stabilizer.

[6] The more obvious difference between our study and most well designed, small experiments is the formulation of iron-fortified salt that was used. Bioavailability of iron in fortified salt is expected to differ along with the level of fortification (and the stabilizer used). At 1 mg iron per gram of salt with sodium-hexametaphosphate as a stabilizer, NIN's DFS contains less iron than some formulas, but the same amount used in previous experiments in India (Haas et al., 2014; Rajagopalan and Vinodkumar, 2000; Sivakumar et al., 2001) and two successful experiments in other countries (Asibey-Berko et al., 2008; Zimmermann et al., 2003).The salt used in the trial with the largest statistically significant increase in hemoglobin (+1.6 g/dL among children in Morocco, over a relatively short 10 months) was fortified at 2 mg of iron per gram of salt (Zimmermann et al., 2004). However, the trial with the highest level of fortification at 3 mg of iron per gram of salt (in Cote d'Ivoire, over 6 months) did not lead to a statistically significant change in hemoglobin among children (Wegmüller et al., 2006).

1.6 g/dL) for schoolchildren (Zimmermann et al., 2003, 2004). In addition to formulation and geographic variation, however, these studies also conflate fortification with significant monitoring of usage (Zimmermann et al., 2003, 2004; Wegmüller et al., 2006; Asibey-Berko et al., 2008) and deworming (Wegmüller et al., 2006).

In sum, there are many good reasons to believe that fortified foods can have a meaningful impact at scale, but no clear evidence that salt fortification itself could. To fill this important gap between policy and evidence, this paper reports on two experiments that test two distribution strategies. First, to experiment on a large scale with sales, we partnered with Tata Chemicals Ltd., a major manufacturer of salt with a well recognized brand name. We make double fortified salt available at the reduced price of ₹9 in shops in 200 randomly selected villages in Bihar where neither Tata Salt Plus nor other iron-fortified salt is available.[7]

In all of these villages, Tata conducted information campaigns, and in some of them, we used additional randomized interventions to study how to increase sales.[8] Take up was reasonably high: on average 42.5% of households in sales villages ever tried DFS, and 14.5% were using it at the time of the endline survey, approximately 23 months after the product had been launched.

The second experiment was embedded in the sales experiment: in 62 villages where we were selling DFS in shops, we distributed a regular supply of DFS for free at home to a subset of households, randomly assigned within each of the villages. Take up was higher among these households, although not near perfect: 61% of households were using it at the time of the survey, and 75% of households had been using it (many of the others had just recently run out). This allows us to test the impact of higher DFS availability on hemoglobin and anemia as well as downstream outcomes. In what is, to our knowledge, the first large-scale trial of double fortified salt, we compare individuals in households that were offered DFS for free at home to individuals in two control groups. The first control group comprises households in villages where DFS was not available at all. The second control is made up of households in villages where DFS was given for free to some other households, but who were not getting it free themselves. This

[7] In Banerjee et al. (2013), we set up a smaller-scale, household-level, randomized pricing experiment to determine the demand curve for DFS. We found that demand falls sharply at a price of ₹10 per kilogram, the price of the cheapest alternative branded salt. Just under a third of the households tried it at just below that price (₹9).

[8] We screened an edutainment movie in some villages and provided shopkeepers incentives to sell the salt in others(Banerjee et al., 2017).

simulates, for example, the policy of making free DFS available to households eligible for food rations when it is otherwise available. In the latter case, both treatment and control households were subjected to all the information campaigns randomized to their villages. We describe our experiments in more detail along with our analytical strategy in Section 2.

In Section 3, we present several results. We first focus on the effects of the sales treatment and the free DFS treatments on our main outcomes of hemoglobin concentration and anemia incidence. Next we estimate the downstream outcomes of cognition, physical health, and mental health. We also create 2SLS estimates on our main and downstream outcomes. Our two instruments, availability of DFS via sales and free DFS, allow us to construct instrumental variables estimates of consumption of DFS. We estimate these results across the full sample and several subsamples of interest – females, males, children, adults and the elderly – as the effects of anemia are dependent on these characteristics.

We find no significant impact of the sales experiment in the whole sample on hemoglobin (the point estimate is 0.033 g/dL, with a 95% confidence interval of [-0.024; +0.09]) or on anemia, including on those who were anemic at baseline. The free DFS experiment also has no impact on our main outcomes overall using both comparison groups.

Unsurprisingly, we also find little improvement in our downstream outcomes of cognition, physical fitness, and mental health. The sales treatment appears to improve cognition overall, but the effect is not robust to an adjustment made for multiple hypothesis testing.

Overall, in the IV regressions, we also fail to reject a zero effect of DFS consumption for the entire sample or subgroups defined by baseline anemia status, gender, and broad age groups. For the anemic at baseline subgroup overall, the 95% confidence interval for hemoglobin is (-0.198; +0.312 g/dL). However, among the narrow subgroup of adolescents (13-17 year olds) we do find a sizable and significantly positive impact on Hb concentration (and, correspondingly, a reduction in anemia) in the IV regressions. For this subgroup, hemoglobin is 0.401 g/dL higher and anemia lower by 12.1 percentage points.

In Section 4, we compare our results to other recent RCTs to control anemia. First, we briefly discuss the results of another experiment where we attempted to provide a simple technology to fortify flour at the local level, with equally disappointing results. Second, we look at two supplementation experiments Chong et al. (2015); Thomas et al. (2006) and compare the impact of our experiments on adolescent males and females, groups that have responded well to

supplementation in other studies (Chong et al., 2015; Sivakumar et al., 2001). It appears that DFS had some impact on adolescents by preventing declines in hemoglobin common in this age group (rather than helping those already suffering from anemia), similar to the pattern found in Chong et al. (2015).

Although these results are only obtained in one setting and probably would need to be replicated, they are not encouraging for the prospect of DFS as a way to control anemia in rural areas for the general population. The challenge yet to overcome may be that to make a noticeable difference for most people (perhaps to sustain their interest in the product), the iron dose must be large. But to be safe (and avoid poisoning due to over-consumption of iron), the concentration of added iron in any fortified food faces limits. With a single source of fortified food, and a diet that continues to be low in iron, the fortification is perhaps insufficient to make enough of a difference for individuals to continue purchasing the product. This pattern would further reduce impact and ultimately make DFS a non-viable strategy for the general population.

2 Research design

This paper reports the results of two experiments conducted concurrently in Bihar. The main experiment studies sales, while a smaller experiment investigates the impact of distributing free DFS to homes. Our primary outcome measurements are hemoglobin level and whether or not an individual is classified as being anemic according to WHO guidelines. Physical fitness, cognition and mental health are important downstream outcomes.

2.1 Double Fortified Salt

The double-fortified salt used in this experiment was manufactured and distributed by TCL under the brand name "Tata Salt Plus." It follows the formula established by the NIN and provides 40 micrograms of iodine and 1 mg of iron per gram of salt. The maximum retail price of Tata Salt Plus is ₹20 (rupees) per kg, making it a relatively low-cost iron source, but around twice the price of regular iodized salt.[9]

Consumption of 10 grams of salt is needed to obtain 10 mg of iron per day, which is approximately 30% of the recommended daily value for Indian populations, though the exact

[9] Regulations in India already require salt to be iodized. Regular Tata Salt, which is the highest quality iodized salt and one of the most expensive available, normally sells for ₹15 per kg.

recommendation is dependent on age and gender (Indian Council of Medical Research, 2009). Salt consumption should be high enough for the mean adult in India to receive the intended amount of iron through DFS (Johnson et al., 2017).[10]

DFS, therefore, is expected to increase hemoglobin among iron-deficient people who eat enough of the salt consistently; who do not suffer illnesses (parasites, malaria, etc.) or have any complementary micronutrient deficiencies (Vitamin C) severe enough to block iron absorption. For a person from this population weighing about 60 kgs, hemoglobin is expected to rise 1 g/dL over a nearly two-year period with a consistent consumption of 10 grams of DFS daily. Iron-deficiency anemia, therefore, would be expected to decrease for DFS consumers who are within 1 g/dL from the anemia cutoff for their age-sex group at baseline.

2.2 Experimental Design

Sampling frame

We conduct our experiments in the Bhojpur district of Bihar. Our sampling frame is the list of villages and households compiled when the District Rural Development Agency updated the number of above and below poverty line households in 2010. Across Bhojpur, there were 999 villages on the list. We consider all villages with fewer than 50 households to be ineligible for inclusion in the experiment. We stratify by Block (an administrative unit smaller than the District) and then in each of eight Blocks randomly select 29 villages and in each of six Blocks randomly select 28 villages. This gives us a total of 400 study villages.

We estimated the required sample size for our experiments using three outcomes: hemoglobin level, days of work lost due to illness (for adults) and cognitive ability (age 50 and above). The calculations suggested that by conducting our main Sales experiment in 200 treatment villages and 200 control villages with 25 measurement households per village, assuming (1) take-up by approximately 30% of households, (2) a .95 confidence level, (3) 80% power, (4) intra-village correlation of 0.024 from pilot data and (5) a Bonferroni correction for multiple outcomes, we would be able to detect an increase in hemoglobin of 0.7 g/dL, or 0.3 standard deviations among the elderly and 0.25 among adults and children. Due to budgetary reasons, we were only able to sample 15 households in each village. We report standard errors in our tables and discuss

[10] The WHO provides recommended iron intakes for populations with very low dietary intake of iron. For adult males the WHO recommends 27.4 mg/day. For adult, non-menopausal females, this number is 58.8 mg/day.

confidence intervals throughout the results section.

Within each village, we randomly select 15 households to be measurement households, for a total of 6,000 households in the baseline sample.[11]

Experiments

Figure 1 summarizes the experimental design. First, we stratify by Block and then randomly assigned half of the villages to treatment and half to control, giving us 200 sales and 200 control villages. Control households are therefore not aware of the existence of Double Fortified Salt when we start the sales experiment. In all sales villages, we provide the option to stock Tata Salt Plus to all *kirana* shops that operate in the village and the Public Distribution (PDS) shop that serves the village. Packets of DFS are clearly marked with the special research Maximum Retail Price of ₹9 per kg.

The period for stocking salt is August 2012 through February 2015. A dedicated team takes orders from shops and oversees delivery to shops by TCL stockists.[12] TCL launches the product in sales villages using a marketing team of five individuals. This team puts on street plays, plays games with children and others in the village, and gives away prizes. These shows and activities highlight the symptoms of anemia, the body's need for iron, and the benefits of consuming Tata Salt Plus. The shows happen in the most central location of each village or section of the village to maximize attendance. These information campaigns take place in all sales villages within a few days of initial stocking, with each village receiving between one and three shows or activities depending on the size of the village.

For our second test, we randomly select 62 sales villages to be the locations of the Free DFS experiment. In those sales villages, we randomly assign 7 out of the 15 measurement households to receive Free DFS delivered to their homes. Another dedicated team delivers 2 kilograms of salt to 438 homes every 2.5 to 3 months for nearly two years. Households not receiving DFS for free may know that a few households in their villages are receiving it for free at home.

In addition, we perform a number of additional marketing experiments in sales villages,

[11] Four additional households were surveyed at baseline due to independent, unnecessary replacements. They are kept in the studies and the individuals in all 6,004 households are included in the baseline.

[12] Over 30 months of stocking, there are a few instances in which PDS store operators who serve both treatment and control villages indicate that they wished to also buy the salt for shops in control villages. They are refused and told that Tata Salt Plus is being sold using a lottery system, and that those shops have not been selected.

which are described in detail in Banerjee et al. (2017).[13]

Data Collection

We conduct surveys in all 400 villages using teams of surveyors who are trained by a medical doctor to take a drop of blood and measure hemoglobin using the HemoCue method.[14]

We use an in-depth survey to assess a range of additional household-level and individual-level outcomes. At the household level, our team collects household composition, income, assets, consumption, nutritional intakes, socio-economic characteristics, household health service usage, and DFS adoption.[15]

Developmental and cognitive outcomes are captured through four standardized measures, depending on the age of the subject. First, our Infant Development Module was based on the Lucknow Development Screen, which captures developmental delays among infants between the ages of 1 to 30 months and has been validated for use among children in India (Bhave et al., 2010).[16]

Child and adult cognition are measured using sections from the National Institute of Mental Health and Neuro Sciences (NIMHANS) neuropsychological battery and the Post Graduate Institute of Medical Education and Research (PGIMER) battery of memory dysfunction, which were previously validated in India.[17] To measure cognition among the elderly, we use the Hindi Mental State Exam, designed to measure basic cognitive awareness and alertness among a lower-income, Indian population (Ganguli et al., 1995).[18]

[13] The additional cross-randomized marketing experiment treatments comprise an edutainment film shown in villages, a flyer delivered to households, and a one-time discount to a set of chosen retailers.

[14] This testing method has been found to be comparable to standard laboratory techniques for measuring hemoglobin in normal and anemic children (Cohen and Seidl-Friedman, 1988) and to adequately estimate population anemia (Neufeld et al., 2002).

[15] Given our focus on nutrition, a household in this project is defined as a group of people living together under the same roof and eating from the same pot for six out of the past twelve months.

[16] For the Lucknow Development Scree, mothers or another close family member are requested to report on the child's ability to perform a certain age-appropriate task. The task is considered age-appropriate if 97% of infants at that age are able to perform the task.

[17] The child cognition module uses the Digit Span test from the PGI battery and the Visuospatial Working Memory Span Test from NIMHANS, which measures memory and retention. The adult cognition module consists of four sub-tests from the PGI battery, including the Digit Span Test, Word Recall Test, Sentence Recall Test, and Word Pairs Test. Scores from the sub-tests are aggregated to measure memory and retention. The NIMHANS battery is composed of tests taken from other standardized test batteries, such as the Luria-Nebraska Neuropsychological battery. The PGIMER battery of memory dysfunction is comprises 10 subtests including forward and backward digit spans, one minute delayed recall of a word list, immediate recall of sentences, retention of similar word pairs, retention of dissimilar pairs, visual retention, visual recognition, recent memory, remote memory and mental balance test.

[18] The Hindi Mental State Exam (Ganguli et al., 1995) was developed by researchers from the University of

The team also directly measures physical health including weight, height, mid-upper arm circumference, and aerobic capacity and balance (Queens Step Test). Additionally, we collect information on self-reported health, including ability to perform various Activities of Daily Living, as well as symptoms and diagnoses of illness (cold, pain, diarrhea; blood loss, malaria, etc). Women aged 15 years and over are also asked about the outcome of pregnancies within the intervention period.

For mental health, we focus on measuring depression, and the set of self reported questions we use to construct our measurement is modeled after the CES-D depression index (Radloff, 1977).

Across all the surveys, eligibility to participate in a particular section is determined by age. Table 1 describes the age requirements for the outcome measures. The age cutoffs are defined as shown in order to match the National Family Health Survey (NFHS), which defines elderly as 50 years or older and adults as between 15 and 49 years of age (International Institute for Population Sciences (IIPS) and Macro International, 2008). We created several indices to bring together multiple measures of cognition, physical health, and mental health.

2.3 Empirical specifications

First, we analyze the sales experiment by comparing individuals in sales villages to those in control villages.

In our basic specification, for any outcome y_{ik} for individual i in village k and block b, we run the following regressions:

$$y_{ik} = \alpha + \beta Sales_k + X_{ik}\gamma_s + \delta BaseHb_{ik} + \mu_b + \epsilon_{ik} \quad (1)$$

where $Sales_k$ is a dummy equal to 1 if DFS was made available in village shops, $BaseHb_{ik}$ is hemoglobin concentration at baseline, μ_b is Block fixed effects, and X_{ik} is a vector of control variables (age, age squared, a dummy which indicates if the individual is anemic at baseline, a dummy for household split, education above 5th grade, BMI at baseline –all set to zero if we

Pittsburgh and the Centre for Ageing Research, India.The research team adapts a battery of tests that were developed to diagnose dementia in an educated English-speaking population to be used in Northern India with Hindi speakers who have little or no formal education and were largely illiterate. The test includes sections on orientation to time, orientation to place, registering and recalling basic objects, attention by subtracting serials threes, naming everyday objects, repetition of a phrase, following a visual command, executing a three-step task, saying something about one's houses, and copying a simplified figure.

don't observe the individual at baseline – a dummy for whether we have baseline measurements, and the household wealth index). We also control for the hemocue machine used at endline when hemoglobin or anemia is the outcome. The inclusion of control variables makes very little difference to the point estimates. The standard errors are clustered by village, the unit of randomization for the sales experiment, or by household, the unit of randomization for the Free DFS experiment.

We also use this specification to look at pre-defined subgroups (male, female, children, adults, and elderly) and those suggested through the review process (reproductive age/adult females and infants) because of different biological requirements for iron by sex and age. When we do this, we provide both conventional standard errors, and a Q-value that adjusts for multiple testing across subsamples because a statistically significant coefficient could still emerge by chance due to the greater number of simultaneous tests. We generate sharpened two-stage Q-values(Benjamini et al., 2006) as described in Anderson (2008), which can be interpreted in the same way one interprets a P-value.

We also run this specification separately for anemic individuals, since, consistent with medical evidence, the literature before us finds effects of iron supplementation on hemoglobin concentration and anemia only for those who were anemic at baseline.

Next, we analyze the Free DFS experiment in two ways. First, we compare individuals in households who received Free DFS with those from households in control villages to focus on the highest expected consumption compared to no introduction of the product at all. We start with the following specification and exclude all households in *sales* villages who did *not* receive the Free DFS.

$$y_{ik} = \alpha + \beta FreeDFS_k + X_{ik}\gamma_f + \delta BaseHb_{ik} + \mu_b + \epsilon_{ik} \tag{2}$$

Standard errors are still clustered at the village level, and we conduct the same subgroup analysis (and related adjustment for multiple hypothesis testing).

Second, we exploit the within-village randomization by restricting to the 930 households in the 62 villages where the free DFS household-level randomization takes place and comparing those who are and are not provided free DFS. This sample restriction still results in a randomized sample, since households that receive the free salt are randomly assigned within a set of villages.

This specification allows us to identify the impact of free versus discounted iron fortification, controlling for village-level information campaigns. The basic specification is analogous to the basic sales experiment specification.

$$y_{ik} = \alpha_F + \beta_F FreeDFS_{ik} + X_{ik}\gamma_f + \delta_F BaseHb_{ik} + \mu_k + \upsilon_{ik} \tag{3}$$

Where μ_k is village fixed effects since the randomization is stratified by village. In this case the standard errors are clustered at the household level.

Equations 2 and 3 are both estimated for the appropriate sample, overall, for males and females separately, and then by age groups, across multiple outcomes. We also estimate these specifications separately for those individuals who were anemic at baseline.

Finally, we take an instrumental variables approach, combining two sources of variation; the dichotomous "free DFS household" and "sales village" treatment variables are used as instruments for recent salt consumption, as indicated by either the household currently consuming DFS or the household reporting that it used DFS right before its current salt (penultimately used DFS). The first stage is simply:

$$UseDFS_{ik} = \pi_0 + \pi_1 FreeDFS_{ik} + \pi_2 Sales_k + X_{ik}\pi_3 + \pi_4 BaseHb_{ik} + \upsilon_{ik} \tag{4}$$

where $UseDFS_{ij}$ is a dummy equal to 1 if household i of village k is currently using DFS or was using it last time.[19] We then estimate the equation:

$$y_{ik} = \alpha_U + \beta_U UseDFS_{ik} + X_{ik}\gamma_U + \delta_U BaseHb_{ik} + \upsilon_{ik} \tag{5}$$

by two stage least squares, using $FreeDFS_{ik}$ and $Sales_{ik}$ as instruments for $UseDFS_{ik}$ and the other variables as instruments for themselves.

2.4 Balance

Table 2 presents summary statistics for all experimental groups in Panel A (individual level variables) and Panel B (household-level variables). It demonstrates that the sales arms were very

[19] This will provide smaller and more precise estimates than if we used "currently using DFS," but since the effect of having used DFS should linger, we find it to be the best intermediate variable choice.

balanced, with none of the characteristics showing significant differences between the treatment and control groups (P values in Column 5).

The free DFS sample is quite large, and randomized at the household level within villages. There are a few significant differences between households who received free DFS and control village households (Column 6) and between households who received free DFS and those who did not receive it in the same villages (Column 7), but the point estimates of the differences are very small. Most notably, hemoglobin concentration was a little lower at baseline in the free DFS group (12.07 g/dL versus 12.178 g/dL in non-free), but the proportion of people who have anemia is balanced at around 0.45 across all groups. In free DFS households, the average person is slightly younger (27.0 years versus 28.2 in control, and 28.3 in non-free households) as there are fewer elderly and slightly more children. Education is a bit lower in free DFS households and the differences in downstream health and cognition measures are very small. More free DFS households split between baseline and endline than in the other groups.[20] We do control for all these variables in our regressions.

2.5 Attrition

Attrition would be a threat to the validity of our estimates if it changes the unobserved characteristics of one treatment group relative to another. For example, if repeat contact with a stranger when receiving free salt made free DFS households take more trouble to help schedule a follow-up visit for our surveyor to interview all members at endline, we may have completed surveys with more of the mobile types belonging to households receiving free DFS compared to others. If mobile persons (migrants) have lower (higher) hemoglobin than others, we could under (over)-estimate the impact of free DFS.

Appendix Table 12 looks at household and individual level attrition for both experiments. We did not lose many entire households over the two-and-a-half years between baseline and endline, with the fewest percentage-wise lost from the free DFS treatment group. Household attrition was roughly 4% in the control villages, 5% in sales villages, 4% among households not receiving free DFS in the free DFS villages, but only 2% for households receiving free DFS. These figures are slightly imbalanced across both experiments. At the individual level, attrition is balanced for the sales experiment (though much higher at roughly 19%). For the free DFS

[20] If a household split, we include both resulting households and their members in the endline survey and analysis.

experiment, attrition is again lower among individuals in households receiving free DFS (15%) compared to the non-free households (18.5%) and control (19%).

At the same time, a number of individuals were also added to the sample between the baseline and the endline (for example, only 81.9% of the sales control sample was present at the baseline, and this is balanced across group). Some of these individuals are genuinely new to the household, for example because of marriage, or birth over the long duration of the project. Furthermore, some individuals were missed at baseline, but found at endline.[21] And finally, we suspect that attrition is actually much lower than what we think, because some of the "added" individuals are actually "attrited" individuals who are mismatched. For this reason, and because we see very little chance of endogenous joining of households due to the treatment, we keep all the individuals in the analysis sample, although not including them does not affect the estimates very much. The Online Appendix reproduces our analysis excluding the "new" individuals, and results are substantively the same.

We compare baseline characteristics of non-attritors by group (Appendix Table 13) to check for balance in the endline sample. Column 5 demonstrates that both individual (Panel A) and household-level characteristics (Panel B) are balanced across control and sales villages when we limit the sample to those who were in both baseline and endline.

This table also demonstrates that balance is still a bit of a concern comparing free DFS households with control (Column 6) or non-free households in free DFS villages (Column 7). However, it is the same – or marginally less of a – problem than in the full baseline sample. The proportion anemic is still balanced across all groups, and baseline hemoglobin is balanced as well when we exclude the attritors. The small imbalances in age, other downstream measures, education and splits remain here as well. Therefore, attrition has not worsened balance on observables across treatment and control conditions. Nonetheless, we control for baseline characteristics in our basic specification, including age, hemoglobin, education, and household splits.

[21] No household was added, unless they were a split from a baseline household.

3 Results

3.1 Take Up of Double Fortified Salt

Making Double Fortified Salt available through shops did result in fortified salt consumption. We present descriptive statistics on the take up of Double Fortified Salt in Table 3, by experimental group. Column 2 shows that approximately two years after DFS was launched in local shops, 14.5% of households in sales villages who were not offered DFS for free are consuming the product. Additionally, 21.8% are consuming it at the time of the survey or the previous time they bought salt, and 42.6% had ever tried it. Take up is similar between households who had to buy DFS, though their neighbors received DFS for free (Column 4), as for the overall group of households who had to buy DFS (Column 2). Both the sales and free DFS experiments resulted in higher consumption, which was highly statistically significant (Columns 5 - 7).

In our initial pricing experiment, we saw take up of about 30% at a price of ₹9 per kg (Banerjee et al., 2011), which is higher than consumption at the end of our sales experiments, despite the fact that we had much more intensive, cross-randomized interventions to encourage take up in many of these villages (Banerjee et al., 2017).[22] However, DFS penetration is quite successful in terms of commercial marketing of a new product, according to our partners at TCL. It seems that the DFS launch, and the accompanying interventions, were actually as large a success for a new product launch as what one could hope to see, as just less than half of households at least tried the product at their own expense.

What is a bit worrisome is that take up does appear to fall over time; many people tried it at least once but did not continue with the product. Figure 3 shows that at the store level, purchases also fell over time. This suggests that the product does not have a slow diffusion curve that would eventually culminate in large adoption. Instead it seems that many people tried DFS, but gave it up after awhile. There were a handful of early reports of finding black specks in food as a result of cooking with DFS. We tracked this consistently with free DFS

[22] Our social marketing experiments (reported in detail in Banerjee et al. (2017)) induced greater take up in some villages than in some others. In particular, DFS take up (measured as "currrently using DFS") was about 5 percentage points higher in villages where we showed a high production value "edutainment" movie, and in villages where shopkeepers were given an incentive to market DFS. Using our preferred specification (controlling for baseline hemoglobin level), we examine whether this increased take up was sufficient to generate larger impact on hemoglobin concentration or anemia by endline. We find no difference in any of these villages.

households and this seems to occur very infrequently, if ever, but this is certainly something that was of concern to some number of households at first.

Table 2 shows, however, that the take up of free DFS is high when distributed for free. Not all households try it (only about 75% of the free households ever try DFS for consumption), but many do.[23] Furthermore, those who do are using it until just before the endline, and some of them have stopped because their supply has just run out. We find that 75% of them are using it currently or last time they used a packet of salt (they have recently run out) and 61% are using it currently (which means that they did not buy DFS if the free amount runs out). This suggests a reasonably good acceptability of the product. Thus, if the impacts are large, free distribution of Double Fortified Salt to poor households would in fact be a feasible strategy. Meanwhile, the take up for households in free DFS villages who do not get it for free is 15.6%, and thus this creates a large difference in take up.

3.2 Main results on Anemia and Hemoglobin: Sales Experiment

Table 4a shows the results of estimating equation (1) for the entire sample (Column 1) and females and males separately (Columns 2 and 3). This project is in part funded by the NIA, and our original proposal calls for examining effects separately for children (Column 4), prime age adults (Column 5), and older adults (Column 6) (Banerjee et al., 2010). Through the review process we also added Female Adults (roughly of reproductive age) and Infants (Columns 7 and 8) as these subgroups may benefit the most in terms of health consequences avoided.

We present conventional standard errors as well as Q values that adjust for multiple hypothesis testing for our subgroups for each outcome. A Q value shown in Table 4a, where hemoglobin is the outcome, adjusts for 15 hypotheses tested together.[24]

The bottom line is clear for the full sample and subgroups (Panels A and B). Overall, there is a small increase in average hemoglobin concentration of 0.033 g/dL (95% CI: -0.024 g/dL; +0.09 g/dL) and a small decline in the anemia rate of -0.6 percentage points (95% CI: -2.364 pp; +1.164 pp), but neither is statistically significant. The results are small, and owing to the large sample size, quite precise, so the confidence intervals are tight. We can reject even

[23] Anecdotally, we know that some households give it away or feed it to animals.

[24] We test hemoglobin for the eight subgroups shown in this table plus adolescents in a later table. We also cut each group into a smaller subgroup when we focus on those members who were anemic at baseline (excluding the infant because there is no baseline anemia data). In subsequent tables where the infant subgroup is not of particular interest for a given outcome, Q values will be for 14 simultaneous tests. When the infant subgroup is relevant (as in Table 4a), Q values will be for 15 tests done together.

a small decrease in anemia prevalence.

For both females and males, there is no significant effect of being in a sales village on hemoglobin concentration or having anemia in the full subgroup. For females the 95% CI on hemoglobin concentration is (-0.043 g/dL; +0.079 g/dL) and the 95% CI for anemia is (-2.056 pp; +2.256 pp). For males, the 95% CI on hemoglobin concentration is (-0.02 g/dL; +0.118 g/dL) and the 95% CI for anemia is (-3.36 pp; +0.56 pp). Thus, we cannot reject that being in a sales village had no effect, and we can reject that it had even a small effect.

Interestingly, the point estimates for children overall are larger than in the overall sample, though not statistically signficant. For children, the 95% CI for hemoglobin is (-0.007 g/dL; +0.135 g/dL) and for having anemia it is (-4.548 pp; +0.548 pp). The effects are weaker for prime-age adults, and for the elderly they are wrongly signed, and neither is statistically significant. Even for those groups that are most at risk for anemia, reproductive women and infants, we fail to reject our null hypothesis of no effect.

In Panels C and D of Table 4a, we focus on individuals who were anemic at baseline. The effects of the sales experiment are larger here, but still not significant for any subgroup after adjusting the P values for multiple hypothesis testing. For males, the sales treatment increases hemoglobin by 0.131 g/dL and reduces anemia by 3.6 percentage point. Even here, we can reject a large effect. With a 95% confidence interval for Hb of (+ 0.017 g/dL; + 0.245 g/dL) we can reject an increase of 2.1% or higher in Hb concentration for males who were anemic at baseline. This effect is not robust to adjusting for multiple hypothesis testing.

For children, the effect of sales on hemoglobin among the initially anemic is significant with an increase of 0.127 g/dL, significant at the 10% level. Again, the significance is not robust to the adjustment for multiple hypothesis testing (Q value = 0.300). For children, the effect of sales village on anemia rate is not significant (95% CI: -7.012 pp; +1.612 pp). There are no significant effects for baseline anemic adults, elderly, or female adults. Infants are excluded from Panels (C) and (D) and future "BL Anemic" results because there are only 25 individuals who are still "infants" according to their endline age and who were in the baseline and anemic at that time.

Despite decent take-up results, the overall effect of DFS sales in our experiment is extremely limited. It has no effect on the community overall, even when looking at vulnerable groups (females and elderly) and among the baseline anemic it has only a very small non-robust

effect on baseline anemic males and children.

3.3 Main Results on Anemia and Hemoglobin: Free DFS Experiments

The limited impact of the sales experiment could be the product of two things: modest take up (though probably as high as could be obtained with this kind of product) and the impact of the product.

As mentioned, the level of fortification of NIN's DFS is lower than supplements. The hope was that regular DFS availability by households would be sufficient to complement other sources of iron in an individual's diet.Surprisingly, as we noted, there is no rigorous empirical evidence backing this claim. To test this hypothesis, we distribute free DFS to households, with the aim of increasing take up of fortification. There is, however, neither an obligation for households to use it nor intense monitoring of usage. Our free DFS experiment is, to our knowledge, the first large-scale trial of free DFS fortification looking more closely at efficacy than effectiveness.

Tables 4b and 4c show reduced form impacts of being in the free DFS sample on hemoglobin concentration and having anemia, either compared to the control group, as in equation (2) or compared to other households in the same villages as in equation (3).

In Table 4b, the effect is similar for the whole sample to the results of the sales experiment (point estimates of 0.045 g/dL and -1.5 pp, with larger standard errors). There are no statistically significant effects in Panels A and B of this Table. And in contrast to the sales experiment, the effects are now smaller for males than for females.

We compare the free DFS households to households in free DFS villages who do not receive the product for free (Table 4c) and find even smaller point estimates. For example, for the whole sample, the point estimate on Hb concentration is -0.007 with a 95% CI of (-0.093 g/dL; +0.079 g/dL). There are again no large or statistically significant effects for any of our subgroups: baseline anemic, females, males, children, prime age adults or older adults.

In short, the free DFS experiment yielded no positive impact, comparing the treatment households to either households in control villages or to non-free households in the same villages.

3.4 Downstream Impacts on Cognition, Physical and Mental Health

Tables 5, 6 and 7 present results for cognition, physical health, and mental health from the sales and free DFS experiments. For these tables, we create indices from standardized measures that

take multiple dimensions into account. All indices are constructed using the baseline mean and standard deviation from the sales experiment control group for the given variable of interest.

As done earlier, we repeat each regression limiting the sample to those individuals who were anemic at baseline and present conventional standard errors as well as Q values that adjust for multiple hypothesis testing for multiple subgroups as described above. For the free DFS experiment, we compare only to the control villages as there were no differences in main results between the two ways the free DFS experiment was analyzed.

3.4.1 Cognition

A cognition module was administered to all respondents over the age of one month, with one of four module types being given dependent on age, as shown in Table 1. The cognition scores are standardized such that a positive score indicates a higher level of cognition.

For the sales experiment, cognition results (Table 5, Panels A and B) show some resemblance to the anemia results. Cognition is higher in the sales treatment villages with a point estimate of 0.040 standard deviations (95% CI: + 0.003 SD; + 0.077 SD; Q value = .113). When we limit the sample to those anemic at baseline, the point estimate is a bit larger at +0.054 standard deviations, and the Q-value is still .113. However, for the free DFS experiment, the effects are much smaller and not significant for the full sample.

There is no effect of either sales or free DFS for females, but the cognition results for males mirror the overall results. Males' cognition scores are 0.044 standard deviations higher in the sales villages at endline and 0.077 standard deviations higher for the males who were initially anemic (Column 3, Panels A and B). After adjusting for multiple hypothesis testing, the Q values are .113 for both all males and baseline anemic males. For the free DFS experiment, however, the coefficients on males' cognitions scores are much smaller and not statistically significant, mirroring the insignificant hemoglobin and anemia results for free DFS in Table 4b.

We observe no large or statistically significant effects on cognition for any age subgroup for the Sales (Panel A, Columns 4 - 8) or Free DFS experiments (Panels C and D, Columns 4 - 8), including women of reproductive age and infants. The only thing of note in this table is for the elderly who were anemic at baseline. For the Sales treatment their cognition shows gains at the 95% level (Panel B, Column 6), but the Q value (0.113) indicates that this may be due to chance.

3.4.2 Physical Fitness

For the physical fitness outcomes, only respondents over the age of 10 who were not pregnant and could stand alone unassisted were eligible. Depending on age and whether the respondent had any arthritis or knee conditions, one of two batteries of tests was completed. Each of the individual tests was first standardized and then the entire battery of tests was standardized in order to obtain the standardized physical fitness score. Again, we multiplied the original standardized variable by negative 1 so that a positive physical fitness measure can be interpreted as being more physically fit. While the coefficients in Column 1 indicate worse physical health in all treatment groups and larger gaps for the anemic at baseline, none of these coefficients is statistically significant. The results are also small and not significant for all tests in this table with the exception of children in free DFS households. The point estimate (-0.167) indicates worse health for these kids, but it is not robust to multiple hypothesis testing.

3.4.3 Mental Health

Finally, we create a mental health index to measure depression. Children (ages 10-15) and adults (15+) are asked slightly different questions; therefore, the scores are first summed and then standardized separately.[25] Since a higher score indicates a greater level of depression, we multiply the original standardized variable by negative 1 so that a positive value can be interpreted here as having better mental health.

In the overall sample (Table 7, Column 1), the mental health index is no different between treatment and control for any of the four experiments tested. For example, the coefficient on mental health is 0 in the sales experiment (95% CI: - 0.043 SD; + 0.043 SD), allowing us to reject a very small impact. The result is similar for males, females, and all age subgroups.

Not surprisingly, given that there is no impact on anemia, we also find no consistent impact on downstream measures, both for the sample as a whole, and our sex and age subgroups.

3.5 Instrumental Variable Estimates

The sales and free DFS experiments can be combined to provide an instrumental variables estimate of the actual consumption of salt. There are several candidate variables to measure the exposure to DFS: currently using DFS; used it currently or last time; or ever used it. This choice

[25] Only respondents 10 years or older were eligible for measurement.

of instruments would scale our estimates by a different factor. We choose "using DFS now or having penultimately used DFS" because if households' last packet were DFS, individuals would have been recently exposed, and we should still be able to observe impacts on their health.

Table 8 shows a very robust first stage for actual DFS consumption: households in sales villages are 16 percentage points more likely to consume DFS. In addition, households who got the free DFS treatment are an extra 50 percentage points more likely to consume DFS. Interestingly, individuals who were anemic at baseline (who may be expected to benefit more) are no more likely to consume DFS, most likely because they are not aware of their anemia status.

Table 9 present the results of the 2SLS estimation for all our outcomes. Columns (1) to (7) continue to present our results for the same samples: all, females, males, children, adults, elderly, and adult females. Infants are not included here for the same reason they are excluded from earlier baseline anemic panels: there are only 25 individuals who are still "infants" according to their age at endline and who were in the baseline and anemic. The label Column indicates what outcome is being tested (hemoglobin, anemia, cognition, physical fitness, or mental health) and if we are looking at the full (sub)sample or just those individuals who were anemic at baseline.

Not surprisingly, given the reduced form estimates discussed earlier, none of the estimates are significant. Overall, the point estimate is a 0.043 g/dL increase in Hb (95% CI: - 0.133 g/dL; + 0.219 g/dL) and a 1.8 percentage point reduction in the rate of anemia (95% CI: - 7.484 pp; + 3.884 pp). This means we can reject a 2 percent increase in Hb level for females in general.

For anemic males, the point estimates for all males as well as baseline anemic males are both negative because of the negative impact of free DFS, which more than compensates for the positive effect in the sales villages. For anemic females (95% CI: -0.108 g/dL; +0.432 g/dL), we can reject a 4 percent increase in hemoglobin.

For children, the point estimates are larger, again, than for the most other age groups, but none is statistically significant. The non-significant increase in hemoglobin is 0.127 for the full sample of children but only 0.031 g/dL higher among the anemic at baseline children. The group with the largest – though still not statistically significant – effects is adult females. The increase in hemoglobin for adult females who were anemic at baseline is 0.233 and the 95% confidence interval is (-0.065 g/dL; +0.531 g/dL). This means we can reject a effect size smaller

than what we initially established as meaningful (0.7 g/dL).

We discuss how these point estimates compare with others in the literature in Section 4.

4 Discussion

Given the surprising (disappointing) nature of our results, we find it useful to compare them with two other results. In this section we look at three other well-implemented RCTS: one on fortification in India, one on supplementation efficacy in Indonesia, and one studying supplementation encouragement in Peru. These comparisons suggest our results are accurately depicting what this level of fortification can do alone at scale without intensive and expensive monitoring.

4.1 Comparison with Decentralized Iron Fortification of Flour

For two of the authors of this paper, this was not the first (failed) attempt at fortifying food with iron for the poor. From 2006 to 2009, we set up and evaluated a community-level iron fortification program in 134 villages (65 villages were chosen as treatment villages) in Udaipur district, Rajasthan (Banerjee et al., 2011). In this area, households get their grain (maize or wheat) milled once or twice a month by a local miller, or *chakki*. The community-level iron program was designed to increase bio-availability of iron for families who do not buy commercial food, and was designed and implemented by *Seva Mandir*, a well-respected local NGO. On average, each hamlet has four *chakkis* (this is also the median number). In each village, Seva Mandir chose to work with the two *chakkis* serving the majority of households.

A fortification program must meet two objectives: supply a sufficient quantity of iron in the diet, and avoid supplying too much iron. Safety is also a concern, as the process will not be as tightly monitored as it can be in a factory. It is important that the program is robust to accidental over fortification. The technology for fortification begins with a pre-mix, a dry powdered mix with specific concentrations of one or more micronutrients. This pre-mix is diluted into a pre-blend (because pre-mix is too concentrated to be properly hand-mixed into the flour) and then added to flour either during the milling process or after the grain has been milled.

After consultations with micronutrient initiatives and various experts, *Seva Mandir* chose to use ferrous sulfate (FeSO4) and folic acid (which helps with iron absorption). The pre-mix was mixed with flour at *Seva Mandir* (16.66 grams of pre-mix is added to one kg of flour), to produce a pre-blend which had 3300 ppm (or milligrams per kg) of elemental iron (as ferrous

sulfate). This quantity is sufficiently diluted so that if someone were to eat the pre-blend without mixing, there would be no health risk. This pre-blend was then mixed with the ground grain (maize or wheat) in the appropriate quantity, using a measuring scoop, and a simple mixing machine by the *chakkis*. Customers were not charged for the fortification.

The final concentration of iron in the flour ranged from about 20 to 33 mg per kg.[26] A pilot survey on the weight of flour milled showed that the average adult eats 300 grams of flour per day. This implies that the average adult would get an extra 6 to 12 mg of iron from the fortified flour, spanning the 10 mg of iron that would have been provided by the consumption of 10 g of DFS.

Before the program started in a village, a village meeting took place, in which the causes and consequences of iron deficiency anemia were discussed, as well as what steps households could take to prevent it (changes in diet). The program was then explained to the village, and the village collectively agreed to participate (all villages agreed).[27] At the individual level, a household had to initially agree to be a participant. Once a household accepted the program, the *chakkis* were to consider them to be participating households, unless they explicitly declined fortification.

However, in spite of the initial decision to fortify, many households did not regularly fortify, either because the *chakki* did not always fortify the grain, or because households switched to non-participating *chakkis*. Figure 2 plots the take up of the program as a function of the date the program started and separates the households into three groups: those for whom the closest *chakki* fortifies; those who do not fortify, but have a fortifying *chakki* nearby (within 1.5 kilometers); and those who do not have a fortifying *chakki* nearby (within 1.5 kilometers). Take up initially increases in all three groups, but does not reach the same peak for those who do not have a *chakki* nearby. All those who have a *chakki* nearby reach the same peak, but take up falls more quickly for those for whom the participating *chakki* is not the closest one, presumably because households switched back to their normal *chakki* after a while.

By the time of the endline survey there was, as in the DFS experiments, no impact on anemia or hemoglobin (see Table 10, which reproduces Table 4 from Banerjee et al. (2011)).

[26] Except for the top of the first bin.

[27] To avoid creating spurious effects due to the information regarding anemia, *Seva Mandir* held a village meeting in control villages as well, where the discussion was the same (except that the program was not discussed).

In the IV specification, the standard errors are a bit larger than in the DFS study, but the confidence intervals in both studies have large zones of overlap.

These two fortification interventions tried to attack the problem in two ways: by offering households a technology to fortify their own flour or by making Double Fortified Salt available at a price below that of regular iodized salt. Both would have provided nearly the same amount of iron if consistently taken up. Both foundered on the same problems: (1) take up declined over time and (2) the level of fortification was probably too low to lead to significant improvements for most people.

4.2 Comparison with Two Iron *Supplementation* Interventions

This section compares our results to the WISE experiment in Indonesia (Thomas et al., 2006) and the encouragement design for the take up of iron supplements in Peru (Chong et al., 2015).

In Thomas et al. (2006), the treatment individuals received 120 mg of iron per week, and compliance was excellent, due to very active follow up. This was not envisioned as a potential policy, but as a "mechanism experiment" with rich data to investigate the impact of improving iron intake. With perfect compliance, and if each individual consumed 10 g of salt per day, our experiment would only have led to an increase of iron intake of 70 mg per week, or about 60% what was given in Thomas et al. (2006).

Thomas et al. (2006) find effects among those who were anemic at baseline. For anemic males, they find an increase in Hb of 0.399 g/dL. Our corresponding point estimate for the IV specification for anemic males is negative, but with a confidence interval of (-0.397 g/dL; +0.277 g/dL). We can thus reject their point estimate, but since we only fortified at about half the level of what subjects there received (assuming that those who consumed DFS did so regularly) the order of magnitude is reasonable. This would suggest that the level of fortification of the DFS is simply insufficient. For anemic females, they find a 0.203 g/dL increase in Hb. Our 95% confidence interval for the IV estimate is (- 0.108 g/dL; + 0.432 g/dL). In this case, we cannot reject their point estimate.

Chong et al. (2015) encouraged their treatment group to take up iron supplement pills, and members of this group received on average five more pills than the control group. This amounts to an extra 500 mg of iron over three months (the control group also received about 500 mg over three months, so on average the treatment group received 1000 mg over three

months). This is only a third of what individual adults received in Thomas et al. (2006) and about 55% of what someone who ate enough DFS regularly would obtain.

Despite the lower dose in Chong et al. (2015), they report very high estimates for a sample of male and female adolescents, ages 13-17. For males and females who were anemic at baseline, the reported effect on Hb is 0.5 g /dL. They don't report a standard error for this difference, but they also find a decrease of 21 percentage points in anemia among the initially anemic adolescent, with a 95% confidence interval of (-0.42; +0). This is much larger than the 3 to 4 percentage point, (insignificant) reduction in anemia found by Thomas et al. (2006), and of course much larger than what we find for the sample as a whole. The contrast between Chong et al. (2015) and Thomas et al. (2006) suggests that the impact of supplementation may be much larger for adolescents than for adults.

To investigate the effect of DFS on adolescents in our experiments, we run our main OLS and 2SLS estimates for the same age group (13 to 17 year olds) as in Chong et al. (2015), broken down by anemia status and sex. Table 11 presents all comparisons and models: Panel A shows the Sales experiment, Panel B shows the free DFS OLS results comparing with the control group, Panel C shows the free DFS OLS results comparing to non-free households, and Panel D presents the second stage 2SLS results. We look first at hemoglobin for each subgroup (all adolescents, anemic at baseline adolescents, male adolescents, anemic at baseline male adolescents, female adolescents and finally anemic at baseline female adolescents). We then look at anemia for the same subgroups.

The OLS results for the first two experiments (Panels A and B) are correctly signed and the coefficients are larger for the anemic subgroups, but they are also small and not statistically significant. However, the instrumental variables analysis reveals an increase of 0.401 g/dL (95% CI: + 0.033 g/dL; + 0.769 g/dL) for adolescents overall (Panel D, Column 1) with a corresponding decrease in anemia (Column 3) of 12.1 percentage points (95% CI: - 21.9 pp; - 2.3 pp). These coefficients most closely resemble those for non-anemic adolescents (Panel D, Columns 3, 5, 9, 11). These confidence intervals include the Chong et al. (2015) result for anemic children, both for anemia and for Hb concentration. These results are not driven by the adolescents who were anemic at baseline (Panel D, Columns 2 and 8), indicating DFS may have protected the non-anemic from declines in hemoglobin.[28] However, Q values suggest these significant results

[28] We find the results when we have all adolescents together, not just among those who were anemic at baseline.

are from multiple hypothesis testing. We also investigated whether we have impacts on the health, mental health and cognitive outcomes for this sample of adolescents (not reported), but do not find anything, contrary to Chong et al. (2015).

To summarize, the comparison with three very well implemented field RCTs of fortification or supplementation suggests that our estimates are consistent with the literature, given how fortification could happen via DFS. Unfortunately, it seems that these effects are just too weak for the population at large. The adolescent results, however, gives some indication that supplementation targeting specifically adolescents (though pills or through food fortification in school meals) is a promising avenue.

5 Conclusion

Although these results are only obtained in one setting and should be replicated, they are not encouraging for the prospect of DFS as a way to fight anemia in the general rural population. Our own previous work on decentralized flour fortification had equally disappointing impacts. Overall, the evidence in favor of food fortification as a scalable method to prevent anemia among the very poor is less than overwhelming.

This contrasts with positive results of iron supplementation in the two major studies we reviewed. In one study, Thomas et al. (2006), the dose was large and the compliance excellent. This gives us the upper bound of plausible impact of supplementation (which is consistent with what we find), but does not really open a path for a policy. In the other study, Chong et al. (2015), the focus was to improve the take up for a program that was making a supplement available among adolescents, and this seems to be a replicable strategy.

The issue with fortification may be that to make a notable difference for most people (sufficient to perhaps sustain their interest in the product), the iron dose must be large(r). But to be safe (and avoid poisoning due to over-consumption of iron), the concentration of iron supplementation in the food must be limited. With a single source of fortified food, and a diet that continues to be low in iron, the supplementation is perhaps insufficient to make enough of a difference for individuals to be willing to continue with the program. This of course further

Our baseline, however, was taken two years prior, as opposed to three months before, and it is possible that the anemia status for adolescent fluctuates enough to be a bit irrelevant as a baseline measure. In their sample, the control group lost 0.78 g/dL of hemoglobin over the few months of the study period and the treatment group lost 0.28 g/dL.

reduces impact, and ultimately makes the strategy non-viable.

The one exception to this picture may be children and adolescents, who seem to have larger effects both in our IV results and in other studies, and may also be made to consume iron, either through fortification of salt or other ingredients through schools. It is worth noting, however, that adolescents can also be reached through targeted pill distributions in school, which is a strategy now pursued by several governments in India and elsewhere.

References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. Journal of the American statistical Association 103(484), 1481–1495.

Andersson, M., P. Thankachan, S. Muthayya, R. B. Goud, A. V. Kurpad, R. F. Hurrell, and M. B. Zimmermann (2008). Dual fortification of salt with iodine and iron: a randomized, double-blind, controlled trial of micronized ferric pyrophosphate and encapsulated ferrous fumarate in southern India. The American Journal of Clinical Nutrition 88(5), 1378–1387.

Asibey-Berko, E., S. Zlotkin, G. Yeung, W. Nti-Nimako, B. Ahunu, S. Kyei-Faried, J. Johnston, M. Tondeur, and V. Mannar (2008). Dual fortification of salt with iron and iodine in women and children in rural Ghana. East African Medical Journal 84(10), 473–480.

Banerjee, A., S. Barnhardt, and E. Duflo (2010). Intervention to Fight Anemia and Improve Well-Being in A Very Low Income Setting. Project 6 of NIH grant proposal funded under P01AG005842-28 (David Wise, PI).

Banerjee, A., S. Barnhardt, and E. Duflo (2013). Nutrition, iron deficiency anemia, and the demand for iron-fortified salt: Evidence from an experiment in rural Bihar. In Discoveries in the Economics of Aging, pp. 343–384. University of Chicago Press.

Banerjee, A., S. Barnhardt, and E. Duflo (2017). Insights in the Economics of Aging, Chapter Movies, margins, and marketing: Encouraging the adoption of iron-fortified salt. University of Chicago Press.

Banerjee, A., E. Duflo, and R. Glennerster (2011). Is decentralized iron fortification a feasible option to fight anemia among the poorest? In Explorations in the Economics of Aging, pp. 317–344. University of Chicago Press.

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006). Adaptive linear step-up procedures that control the false discovery rate. Biometrika, 491–507.

Bhave, A., R. Bhargava, and R. Kumar (2010). Development and validation of a

new Lucknow development screen for Indian children aged 6 months to 2 years. Journal of Child Neurology 25(1), 57–60.

Bhutta, Z. A., T. Ahmed, R. E. Black, S. Cousens, K. Dewey, E. Giugliani, B. A. Haider, B. Kirkwood, S. S. Morris, H. Sachdev, et al. (2008). What works? interventions for maternal and child undernutrition and survival. The Lancet 371(9610), 417–440.

Brahmam, G., K. Nair, A. Laxmaiah, C. Gal Reddy, S. Ranganathan, M. Vishnuvardhana Rao, A. Naidu, K. Vijayaraghavan, B. Sivakumar, K. Krishanaswamy, et al. (2000). Community trials with iron and iodine fortified salt (double fortified salt). In Proceedings of the 8th World Salt Symposium, Volume 2, pp. 955–960. Elsevier Amsterdam.

Caravotta, G. (2009, September). UNICEF India Stories: Combatting Malaria in Bihar.

Chong, A., I. Cohen, E. Field, E. Nakasone, and M. Torero (2015). Iron deficiency and schooling attainment in Peru.

Cohen, A. and J. Seidl-Friedman (1988). Hemocue system for hemoglobin measurement. evaluation in anemic and nonanemic children. American journal of clinical pathology 90(3), 302–305.

de Benoist, B., E. McLean, I. Egli, and M. Cogswell (2008). Worldwide Prevalence of Anaemia 1993-2005: WHO Global Database of Anaemia. Technical report, World Health Organization Geneva.

Ganguli, M., G. Ratcliff, V. Chandra, S. Sharma, J. Gilby, R. Pandav, S. Belle, C. Ryan, C. Baker, E. Seaberg, et al. (1995). A Hindi version of the MMSE: The development of a cognitive screening instrument for a largely illiterate rural elderly population in India. International Journal of Geriatric Psychiatry 10(5), 367–377.

Gillespie, S. (1998). Major issues in the control of iron deficiency. Micronutrient Initiative; UNICEF.

Haas, J. D. and T. Brownlie (2001). Iron deficiency and reduced work capacity: a critical review of the research to determine a causal relationship. The Journal of Nutrition 131(2), 676S–690S.

Haas, J. D., M. Rahn, S. Venkatramanan, G. S. Marquis, M. J. Wenger, L. E. Murray-Kolb, A. S. Wesley, and G. A. Reinhart (2014). Double-fortified salt is efficacious in improving indicators of iron deficiency in female Indian tea pickers. The Journal of Nutrition 144(6), 957–964.

Indian Council of Medical Research (2009). Nutrient requirements and recommended dietary allowances for Indians. Technical report, National Institute of Nutrition (NIN).

International Institute for Population Sciences (IIPS) (2017). National Family Health Survey 2015-16 (NFHS-4) State Fact Sheet Bihar. Technical report.

International Institute for Population Sciences (IIPS) and Macro International (2008). National Family Health Survey (NFHS-3), India, 2005-06: Bihar. Technical report, IIPS, Mumbai.

Johnson, C., D. Praveen, A. Pope, T. S. Raj, R. N. Pillai, M. A. Land, and B. Neal (2017). Mean population salt consumption in india: a systematic review. Journal of Hypertension 35(1), 3–9.

Lozoff, B. (2007). Iron deficiency and child development. Food & Nutrition Bulletin 28(Supplement 4), 560S–571S.

Lozoff, B., J. Beard, J. Connor, B. Felt, M. Georgieff, and T. Schallert (2006). Long-lasting neural and behavioral effects of iron deficiency in infancy. Nutrition Reviews 64(s2), S34–S43.

Mudur, G. (2013, June 9). Doubts on fortified midday-meal salt - safe but no clear proof it increases haemoglobin: Scientists. The Telegraph.

Nagar, R., S. Sinha, and R. Raman (2015). Haemoglobinopathies in eastern Indian states: a demographic evaluation. Journal of community genetics 6(1), 1–8.

Nair, K., G. Brahmam, S. Ranganathan, K. Vijayaraghavan, B. Sivakumar, K. Krishnaswamy, et al. (1998). Impact evaluation of iron & iodine fortified salt. Indian Journal of Medical Research 108, 203.

Nair, K. M., B. Sesikeran, S. Ranganathan, and B. Sivakumar (1998). Bioeffect and safety of long-term feeding of common salt fortified with iron and iodine (double fortified salt) in rats. Nutrition Research 18(1), 121–129.

National Institute of Nutrition (India) (2005). Double fortified common salt (DFS) as a tool to control iodine deficiency disorders and iron deficiency anaemia: a report. National Institute of Nutrition.

Neufeld, L., A. García-Guerra, D. Sánchez-Francia, O. Newton-Sánchez, M. D. Ramírez-Villalobos, and J. Rivera-Dommarco (2002). Hemoglobin measured by hemocue and a reference method in venous and capillary blood: a validation study. Salud pública de México 44(3), 219–227.

Penninx, B. W., M. Pahor, M. Cesari, A. M. Corsi, R. C. Woodman, S. Bandinelli, J. M. Guralnik, and L. Ferrucci (2004). Anemia is associated with disability and decreased physical performance and muscle strength in the elderly. Journal of the American Geriatrics Society 52(5), 719–724.

Peters, R., L. Burch, J. Warner, N. Beckett, R. Poulter, and C. Bulpitt (2008). Haemoglobin, anaemia, dementia and cognitive decline in the elderly, a systematic review. BMC Geriatrics 8(1), 18.

Radloff, L. S. (1977). The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measurement 1(3), 385–401.

Rajagopalan, S. and M. Vinodkumar (2000). Effects of salt fortified with iron and iodine on the haemoglobin levels and productivity of tea pickers. Food & Nutrition Bulletin 21(3), 323–329.

Ranganathan, S. and B. Sesikeran (2008). Development of the double-fortified salt from the National Institute of Nutrition. Institute of Food Technologists.

Reddy, K. J. and S. Nair (2014). Double fortified salt and deworming-game changers in the battle against iodine and iron malnutrition in Indian school children. Indian Journal of Community Health 26(6), 175–182.

Sivakumar, B., G. Brahmam, K. M. Nair, S. Ranganathan, M. V. Rao, K. Vijayaraghavan, and K. Krishnaswamy (2001). Prospects of fortification of salt with iron and iodine. British Journal of Nutrition 85(S2), S167–S173.

Stoltzfus, R. J. (2001). Defining iron-deficiency anemia in public health terms: a time for reflection. The Journal of Nutrition 131(2), 565S–567S.

Thomas, D., E. Frankenberg, J. Friedman, J.-P. Habicht, M. Hakimi, N. Ingwersen, N. Jones, C. McKelvey, G. Pelto, B. Sikoki, et al. (2006). Causal effect of health on labor market outcomes: Experimental evidence. California Center for Population Research.

Thomas, D., E. Frankenberg, J. Friedman, J.-P. Habicht, M. Hakimi, N. J. Jaswadi, G. Pelto, B. Sikoki, T. Seeman, J. P. Smith, et al. (2003). Iron deficiency and the well-being of older adults: Early results from a randomized nutrition intervention. In Population Association of America Annual Meetings, Minneapolis.

Wegmüller, R., F. Camara, M. B. Zimmermann, P. Adou, and R. F. Hurrell (2006). Salt dual-fortified with iodine and micronized ground ferric pyrophosphate affects iron status but not hemoglobin in children in Cote d'Ivoire. The Journal of Nutrition 136(7), 1814–1820.

World Health Organization (2001). Iron Deficiency Anaemia: Assessment, Prevention, and Control. Technical report, UNCDF, UNU and WHO.

Zijp, I. M., O. Korver, and L. B. Tijburg (2000). Effect of tea and other dietary factors on iron absorption. Critical Reviews in Food Science and Nutrition 40(5), 371–398.

Zimmermann, M. B., R. Wegmueller, C. Zeder, N. Chaouki, F. Rohner, M. Saïssi, T. Torresani, and R. F. Hurrell (2004). Dual fortification of salt with iodine and micronized ferric pyrophosphate: a randomized, double-blind, controlled trial. The American Journal of Clinical Nutrition 80(4), 952–959.

Zimmermann, M. B., C. Zeder, N. Chaouki, A. Saad, T. Torresani, and R. F. Hurrell (2003). Dual fortification of salt with iodine and microencapsulated iron: a randomized, double-blind, controlled trial in Moroccan schoolchildren. The American Journal of Clinical Nutrition 77(2), 425–432.

6 Figures

Figure 1: Design of Experiments

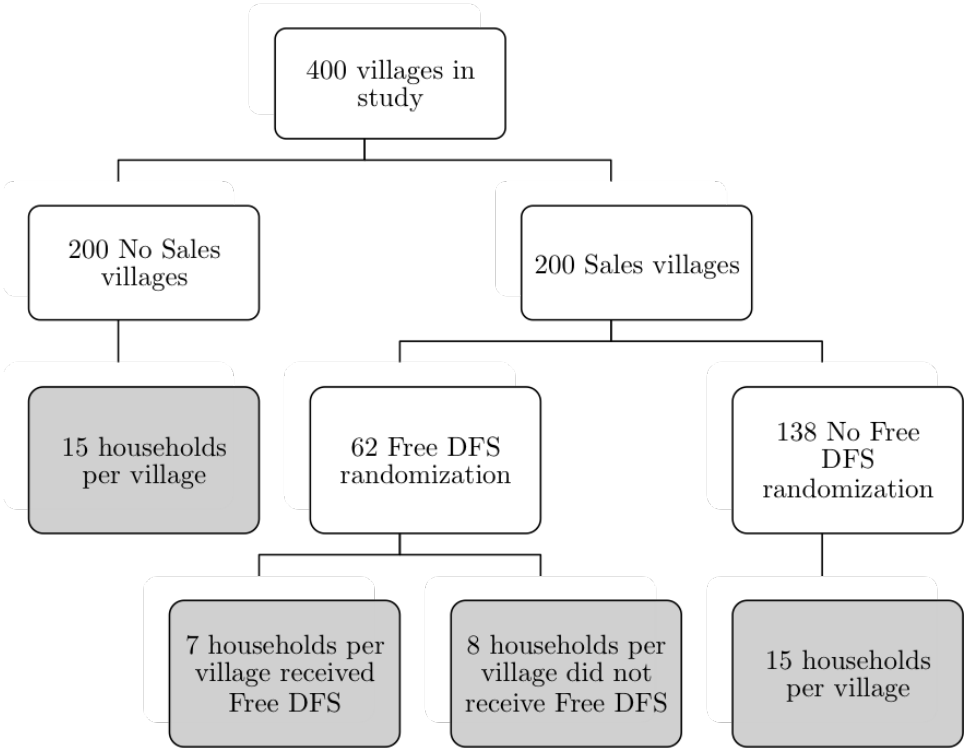


Figure 2: Take-up from Rajasthan

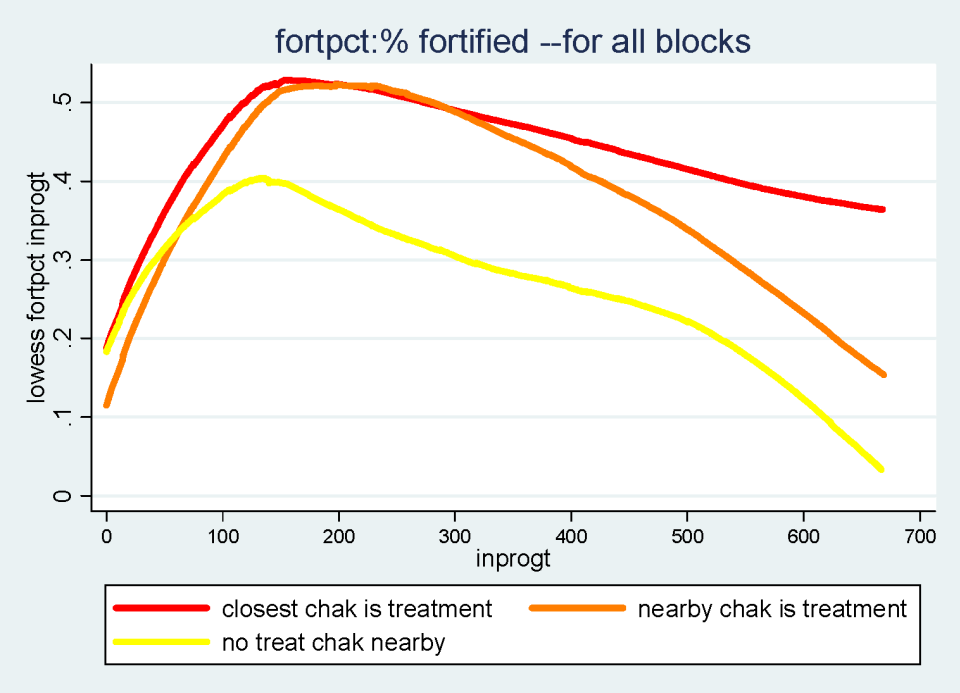
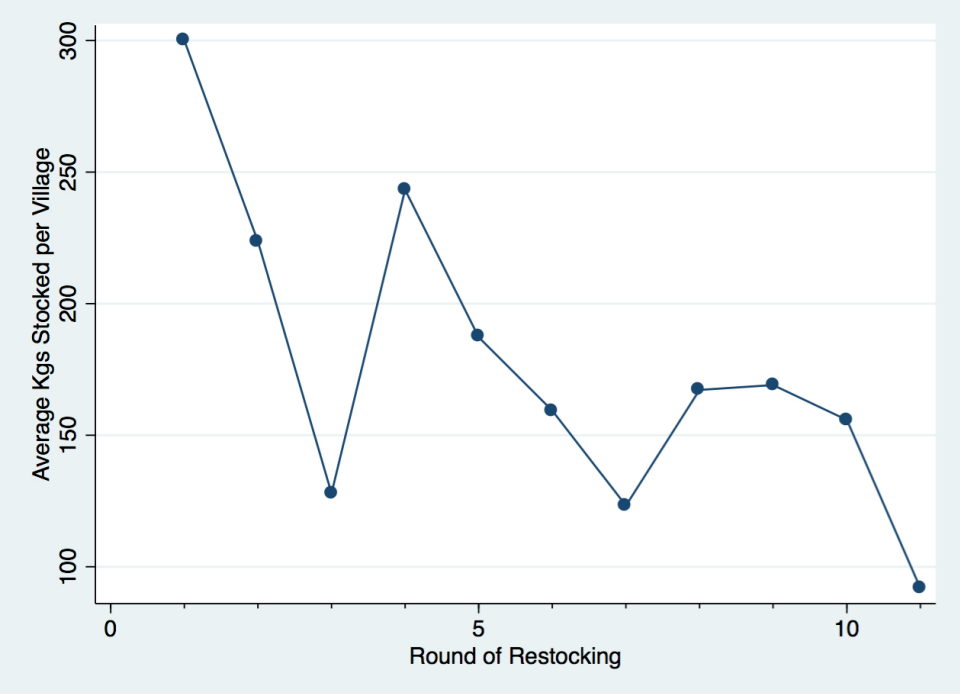


Figure 3: Store Take Up over Time



7 Tables

Table 1: Outcome Eligibility

Topic	**Eligible Age**
Hemoglobin	Ages 6 months and above
Anthropometry	All ages; able to stand
Physical Fitness	Ages 10 years and older
Infant Development	Ages 1-30 months
Child Cognition	Ages 5-14 years
Adult Cognition	Ages 15-49 years
Elderly Cognition	Ages 50 and over

Table 2: Balance Checks for Sales and Free DFS Experiments

	Baseline Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control [2] vs [1]	Free DFS vs Control [3] vs [1]	Free DFS vs Non-Free [3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Individual-Level							
Female	0.515	0.513	0.505	0.516	0.704	0.389	0.150
	(0.500)	(0.500)	(0.500)	(0.500)	(36706)	(22577)	(6301)
Age	28.170	27.882	26.984	28.302	0.262	0.006	0.008
	(20.319)	(20.273)	(20.000)	(20.543)	(36177)	(22254)	(6225)
Child (<15 yrs)	0.323	0.325	0.341	0.322	0.738	0.088	0.203
	(0.468)	(0.468)	(0.474)	(0.467)	(36711)	(22582)	(6302)
Adolescent (13-17 yrs)	0.084	0.082	0.074	0.083	0.395	0.021	0.184
	(0.278)	(0.274)	(0.261)	(0.276)	(36707)	(22578)	(6301)
Adult (15-49 yrs)	0.431	0.430	0.428	0.430	0.842	0.742	0.926
	(0.495)	(0.495)	(0.495)	(0.495)	(36711)	(22582)	(6302)
Elderly (50+ yrs)	0.182	0.178	0.162	0.186	0.334	0.008	0.014
	(0.386)	(0.382)	(0.368)	(0.389)	(36707)	(22578)	(6301)
Infant (6-24 mths)	0.032	0.034	0.031	0.033	0.233	0.833	0.638
	(0.175)	(0.180)	(0.174)	(0.179)	(44839)	(27778)	(7832)
Female Adult (15-49 yrs)	0.193	0.195	0.188	0.191	0.501	0.290	0.539
	(0.395)	(0.396)	(0.391)	(0.393)	(44839)	(27778)	(7832)
Anemia	0.447	0.445	0.460	0.457	0.963	0.673	0.603
	(0.497)	(0.497)	(0.499)	(0.498)	(31710)	(19451)	(5404)
Severe Anemia	0.008	0.008	0.012	0.008	0.724	0.093	0.158
	(0.090)	(0.092)	(0.108)	(0.087)	(31710)	(19451)	(5404)
Hemoglobin	12.197	12.187	12.070	12.178	0.731	0.137	0.051
	(1.856)	(1.837)	(1.914)	(1.862)	(31710)	(19451)	(5404)
Cognition Score	0.000	-0.023	-0.072	0.014	0.326	0.035	0.013
	(1.000)	(0.987)	(0.980)	(0.981)	(29611)	(18151)	(5034)
Mental Health Score	-0.000	0.000	0.018	-0.053	0.976	0.430	0.145
	(1.000)	(1.013)	(0.993)	(1.050)	(24208)	(14869)	(4100)
Physical Fitness Score	-0.000	0.021	-0.070	-0.059	0.369	0.066	0.720
	(1.000)	(0.937)	(0.998)	(0.955)	(14114)	(8733)	(2311)
Body Mass Index (BMI)	18.145	18.106	17.907	18.274	0.605	0.056	0.008
	(4.221)	(4.202)	(4.574)	(4.351)	(32024)	(19651)	(5459)
5+ Education Years	0.530	0.525	0.484	0.528	0.703	0.033	0.044
	(0.499)	(0.499)	(0.500)	(0.499)	(32285)	(19825)	(5548)

Continued on the next page...

Table 2: Balance Checks for Sales and Free DFS Experiments (continued)

	Baseline Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All Non-Free DFS	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control [2] vs [1]	Free DFS vs Control [3] vs [1]	Free DFS vs Non-Free [3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Household-Level							
30-day Consumption per Capita	2249.610	2271.441	2125.663	2276.257	0.827	0.233	0.317
	(2594.825)	(2371.649)	(1977.003)	(2886.207)	(4585)	(2812)	(757)
Head: 5+ Education years	0.562	0.564	0.488	0.581	0.949	0.055	0.017
	(0.496)	(0.496)	(0.501)	(0.494)	(3828)	(2354)	(647)
At least one, Elderly and Anemic	0.479	0.461	0.468	0.480	0.242	0.568	0.704
	(0.500)	(0.499)	(0.500)	(0.500)	(5570)	(3436)	(932)
Number of Members	6.660	6.765	7.018	6.783	0.328	0.219	0.314
	(3.480)	(3.969)	(4.033)	(4.176)	(4621)	(2833)	(762)
Only Immediate Family	0.358	0.374	0.358	0.358	0.342	0.998	0.775
	(0.480)	(0.484)	(0.480)	(0.480)	(4615)	(2830)	(760)
Household Wealth Index	-0.033	-0.045	-0.004	-0.016	0.610	0.554	0.744
	(0.660)	(0.608)	(0.764)	(0.644)	(4621)	(2833)	(762)
Split since Baseline	0.146	0.157	0.226	0.177	0.322	0.000	0.054
	(0.353)	(0.364)	(0.419)	(0.382)	(5570)	(3436)	(932)

[1] The sample for "Individual-Level Variables" includes all baseline respondents from within the respective subsets. "Household-Level Variables" includes all households from baseline from within the respective subsets.
[2] The cognition, physical fitness and mental health scores were standardized using the mean from the sales experiment's control group at baseline. The cognition outcomes were standardized such that a positive score indicates a higher level of cognition. The physical fitness scores were standardized such that a positive score indicates being more physically fit. The score for the mental health outcomes were standardized such that a positive score indicates a lower level of depression.
[3] The Wealth Index is the sum of five standardized components: total household assets, total number of animals owned, house ownership, land ownership, and the amount of land owned.
[4] Standard deviations are in parentheses in columns 1-4 and the observation numbers for the respective samples are in parentheses in columns 5-7. P-values are calculated using block-level fixed effects (columns 5 and 6) or village-level fixed effects (column 7). Standard errors are clustered at the village level (columns 5 and 6) or at the household level in (column 7).

Table 3: Take Up Statistics by Experimental Group

	Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control [2] vs [1]	Free DFS vs Control [3] vs [1]	Free DFS vs Non-Free [3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Currently Using DFS	0.034	0.145	0.611	0.156	0.000	0.000	0.000
	(0.182)	(0.352)	(0.488)	(0.363)	(5770)	(3613)	(1063)
Used DFS (currently or penultimately)	0.046	0.218	0.753	0.230	0.000	0.000	0.000
	(0.209)	(0.413)	(0.432)	(0.421)	(3731)	(2374)	(725)
Tried DFS in the Past Year	0.108	0.426	0.745	0.422	0.000	0.000	0.000
	(0.311)	(0.495)	(0.437)	(0.495)	(4050)	(2585)	(789)
Times Purchased DFS in past year	0.779	3.313	6.656	3.684	0.000	0.000	0.000
	(4.003)	(7.086)	(7.378)	(7.864)	(4050)	(2585)	(789)

[1] Standard deviations are in parentheses in columns 1-4 and the observation numbers for the respective samples are in parentheses in columns 5-7. P-values are calculated using block-level fixed effects (column 6) or village-level fixed effects (column 7). Standard errors are clustered at the village level (columns 5 and 6) or at the household level (column 7)

[2] Observations are at the household level.

Table 4a: Hemoglobin & Anemia - Sales Experiment - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)	(8) Infant (6-24 mths)
Panel A: Hemoglobin								
Sales village	0.033	0.018	0.049	0.064	0.037	-0.041	0.016	-0.032
	[0.029]	[0.031]	[0.035]	[0.036]	[0.038]	[0.047]	[0.038]	[0.080]
Q-value	(0.747)	(1.000)	(0.625)	(0.461)	(0.747)	(0.803)	(1.000)	(1.000)
Control Mean	12.056	11.416	12.769	11.492	12.575	11.982	11.533	10.057
Panel B: Anemia								
Sales village	-0.006	0.001	-0.014	-0.020	-0.007	0.021	-0.002	0.004
	[0.009]	[0.011]	[0.010]	[0.013]	[0.010]	[0.013]	[0.013]	[0.024]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	0.491	0.597	0.375	0.472	0.457	0.605	0.617	0.750
Observations	34732	17941	16726	12775	15576	6295	8772	1242
Panel C: Hemoglobin (BL Anemic)								
Sales village	0.074	0.043	0.131*	0.128*	0.057	0.042	0.028	
	[0.043]	[0.044]	[0.058]	[0.058]	[0.058]	[0.069]	[0.057]	
Q-value	(0.461)	(0.747)	(0.300)	(0.300)	(0.747)	(1.000)	(1.000)	
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090	
Panel D: Anemia (BL Anemic)								
Sales village	-0.017	-0.006	-0.036*	-0.027	-0.019	-0.006	-0.009	
	[0.012]	[0.014]	[0.016]	[0.022]	[0.015]	[0.016]	[0.017]	
Q-value	(1.000)	(1.000)	(0.935)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740	
Observations	11503	7166	4317	3534	4909	3035	3718	

[1] The first line of each panel reports the coefficient on the sales treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.
[2] Regressions include block-level fixed effects and a control for free DFS village. Standard errors are clustered at the village level. The sample includes all respondents present at endline for whom we have the respective outcomes data and excludes free DFS households.
[3] Controls include the following variables from endline: age, age-squared, hemocue machine used to take hemoglobin measurement, a dummy for household split between baseline and endline, and a dummy for not present at baseline. The following baseline variables are also included: hemoglobin concentration, completed 5th standard or higher, body mass index (BMI), and household wealth index. For children under 10, the value for whether the head of the household completed 5th standard or higher is used instead of the value for the child.
[4] Anemia thresholds are determined by age and gender, therefore, a dummy for each relevant age group is included in the regressions for anemia.

Table 4b: Hemoglobin & Anemia - Free DFS Households Compared to Households in Control Villages - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)	(8) Infant (6-24 mths)
Panel A: Hemoglobin								
Free DFS	0.045	0.058	-0.004	0.109	0.032	-0.122	0.059	-0.018
	[0.048]	[0.054]	[0.059]	[0.058]	[0.056]	[0.084]	[0.062]	[0.110]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	12.056	11.416	12.769	11.492	12.575	11.982	11.533	10.057
Panel B: Anemia								
Free DFS	-0.015	-0.015	-0.012	-0.032	-0.008	0.006	-0.009	0.015
	[0.015]	[0.018]	[0.018]	[0.020]	[0.016]	[0.023]	[0.020]	[0.045]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	0.491	0.597	0.375	0.472	0.457	0.605	0.617	0.750
Observations	21623	11163	10412	7960	9670	3925	5455	780
Panel C: Hemoglobin (BL Anemic)								
Free DFS	0.019	0.084	-0.097	0.075	0.044	-0.123	0.092	
	[0.073]	[0.075]	[0.102]	[0.098]	[0.092]	[0.116]	[0.089]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090	
Panel D: Anemia (BL Anemic)								
Free DFS	0.000	-0.014	0.026	0.007	-0.001	-0.013	-0.021	
	[0.023]	[0.024]	[0.030]	[0.034]	[0.028]	[0.028]	[0.029]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740	
Observations	7189	4459	2716	2200	3043	1913	2287	

[1] The first line of each panel reports the coefficient on the the free DFS household treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.
[2] Regression include block-level fixed effects. Standard errors are clustered at the village level. The sample includes all households in the sales experiment control group or Free DFS experiment. All respondents within these households who were present at endline and for whom we have the respective outcomes data are included.
[3] For the list of controls included, please refer to note 3 and 4 under Table 4a.

Table 4c: Hemoglobin & Anemia - Free DFS Experiment within Free DFS Villages - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)	(8) Infant (6-24 mths)
Panel A: Hemoglobin								
Free DFS	-0.007	0.007	-0.075	0.010	-0.022	-0.111	0.038	-0.007
	[0.044]	[0.053]	[0.059]	[0.055]	[0.061]	[0.102]	[0.069]	[0.202]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	12.120	11.467	12.860	11.577	12.656	12.021	11.560	10.069
Panel B: Anemia								
Free DFS	-0.003	-0.016	0.011	-0.011	0.003	-0.014	-0.006	0.029
	[0.014]	[0.018]	[0.017]	[0.021]	[0.017]	[0.028]	[0.025]	[0.064]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	0.476	0.589	0.352	0.451	0.438	0.604	0.608	0.764
Observations	6253	3212	3027	2350	2797	1085	1563	261
Panel C: Hemoglobin (BL Anemic)								
Free DFS	0.010	0.088	-0.181	-0.043	0.074	-0.105	0.158	
	[0.072]	[0.080]	[0.127]	[0.101]	[0.110]	[0.159]	[0.114]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	11.413	11.030	12.087	11.466	11.528	11.165	11.016	
Panel D: Anemia (BL Anemic)								
Free DFS	-0.004	-0.030	0.020	-0.003	-0.027	-0.006	-0.053	
	[0.022]	[0.026]	[0.038]	[0.041]	[0.032]	[0.035]	[0.036]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.663	0.734	0.542	0.523	0.698	0.773	0.771	
Observations	2057	1288	763	657	867	524	651	

[1] The first line of each panel reports the coefficient on the the free DFS household treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.
[2] Regressions include village-level fixed effects. Standard errors are clustered at the household level. The sample includes all respondents located in the villages where the Free DFS experiment took place, who were present at endline and for whom we have the respective outcomes data.
[3] For the list of controls included, please refer to note 3 and 4 under Table 4a.

Table 5: Standardized Cognition Scores - Sales and Free DFS Experiments - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)	(8) Infant (6-24 mths)
Panel A: Sales Experiment								
Sales village	0.040*	0.035	0.044*	0.037	0.037	0.046	0.042	0.088
	[0.019]	[0.023]	[0.019]	[0.023]	[0.023]	[0.029]	[0.026]	[0.050]
Q-value	(0.113)	(0.133)	(0.113)	(0.129)	(0.129)	(0.129)	(0.129)	(0.129)
Control Mean	-0.113	-0.343	0.118	0.258	-0.442	-0.022	-0.612	-0.030
Observations	29783	14948	14459	10711	13418	5279	7455	1506
Panel B: Sales Experiment (BL Anemic Only)								
Sales village	0.054*	0.036	0.077**	0.043	0.033	0.079*	0.037	
	[0.025]	[0.027]	[0.030]	[0.036]	[0.027]	[0.037]	[0.029]	
Q-value	(0.113)	(0.180)	(0.113)	(0.191)	(0.180)	(0.113)	(0.180)	
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639	
Observations	9397	5793	3586	2725	4174	2480	3137	
Panel C: Free DFS vs. Control								
Free DFS	-0.009	-0.012	-0.012	0.026	-0.022	-0.069	-0.007	0.058
	[0.027]	[0.029]	[0.032]	[0.039]	[0.033]	[0.049]	[0.035]	[0.077]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.113	-0.343	0.118	0.258	-0.442	-0.022	-0.612	-0.030
Observations	18515	9242	9031	6621	8379	3266	4643	920
Panel D: Free DFS vs. Control (BL Anemic Only)								
Free DFS	-0.020	-0.037	0.011	0.037	-0.042	-0.105	-0.041	
	[0.033]	[0.035]	[0.053]	[0.056]	[0.040]	[0.063]	[0.041]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639	
Observations	5871	3621	2238	1677	2623	1547	1954	

[1] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized cognition score as the outcome. The cognition scores were standardized using the mean from the sales experiment control group at baseline, with a positive score indicating a higher level of cognition. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.
[2] Regressions include block-level fixed effects. Standard errors are clustered at the village level.
[3] Samples include all respondents present at endline for whom we have the respective outcomes data. Panel A & B: Samples exclude free DFS households. Panels C & D: Samples include all respondents located in the sales experiment control group and free DFS households.
[4] Controls include the following variables from endline: age, age-squared, a dummy for household split between baseline and endline, and a dummy for not present at baseline. The following baseline variables are also included: hemoglobin concentration, completed 5th standard or higher, body mass index (BMI), and household wealth index. For children under 10, the value for whether the head of the household completed 5th standard or higher is used instead of the value for the child.
[5] The regression also includes a dummy for which type of cognition test taken (infant, child, adult, elderly) and baseline standardized cognition score.

Table 6: Standardized Physical Fitness Performance - Sales and Free DFS Experiments - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel A: Sales Experiment							
Sales village	-0.001	0.013	-0.027	-0.015	0.024	-0.074	0.043
	[0.032]	[0.036]	[0.039]	[0.052]	[0.035]	[0.049]	[0.039]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.422	-0.537	-0.259	-0.313	-0.396	-0.566	-0.500
Observations	15607	9093	6500	2503	9574	3527	6136
Panel B: Sales Experiment (BL Anemic Only)							
Sales village	-0.017	0.007	-0.073	-0.082	0.011	-0.067	0.034
	[0.039]	[0.043]	[0.059]	[0.091]	[0.045]	[0.062]	[0.046]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	5555	3856	1693	748	3230	1577	2620
Panel C: Free DFS vs. Control							
Free DFS	-0.061	-0.056	-0.087	-0.167*	-0.062	0.045	-0.053
	[0.057]	[0.065]	[0.064]	[0.085]	[0.062]	[0.090]	[0.072]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.422	-0.537	-0.259	-0.313	-0.396	-0.566	-0.500
Observations	9724	5698	4017	1540	6003	2180	3861
Panel D: Free DFS vs. Control (BL Anemic Only)							
Free DFS	-0.067	-0.076	-0.070	0.039	-0.132	-0.004	-0.115
	[0.071]	[0.081]	[0.109]	[0.142]	[0.076]	[0.126]	[0.081]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	3443	2394	1046	454	2008	981	1616

[1] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized physical fitness score as the outcome. The physical fitness scores were standardized using the mean from the sales experiment control group at baseline, with a positive score indicating a higher level of physical fitness. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Only respondents 10 years of age or older were eligible to perform physical fitness tests.

[2] Regressions include block-level fixed effects. Standard errors are clustered at the village level.

[3] Samples include all respondents present at endline for whom we have the respective outcomes data. We did not measure the physical fitness of the "Infant (6-24 months)" category, so this category is excluded from the table. Panel A & B: Samples exclude free DFS households. Panels C & D: Samples include all respondents located in the sales experiment control group and free DFS households.

[4] For the list of controls included, please refer to note 3 under Table 5.

[5] The regression includes a dummy for which type of physical fitness test performed (Queen Step Test or a series of tests) and baseline standardized physical fitness score.

Table 7: Standardized Mental Health Scores - Sales and Free DFS Experiments - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel A: Sales Experiment							
Sales village	-0.000	0.006	-0.013	-0.092	0.012	0.033	0.038
	[0.022]	[0.033]	[0.023]	[0.049]	[0.023]	[0.033]	[0.035]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.090	-0.315	0.170	-0.174	0.006	-0.264	-0.220
Observations	26407	14142	12234	4573	15557	6269	8888
Panel B: Sales Experiment (BL Anemic Only)							
Sales village	0.022	0.030	-0.001	-0.074	0.045	0.017	0.074
	[0.031]	[0.040]	[0.037]	[0.081]	[0.035]	[0.044]	[0.043]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	9210	6061	3137	1320	4888	3002	3739
Panel C: Free DFS vs. Control							
Free DFS	-0.017	0.000	-0.044	-0.030	-0.006	-0.035	-0.009
	[0.033]	[0.048]	[0.038]	[0.080]	[0.037]	[0.049]	[0.056]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.090	-0.315	0.170	-0.174	0.006	-0.264	-0.220
Observations	16369	8775	7570	2848	9632	3884	5510
Panel D: Free DFS vs. Control (BL Anemic Only)							
Free DFS	0.007	0.007	-0.017	0.110	-0.008	-0.028	-0.010
	[0.055]	[0.065]	[0.069]	[0.116]	[0.065]	[0.072]	[0.073]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	5697	3734	1955	803	3018	1876	2288

[1] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized mental health score as the outcome. The mental health scores were standardized using the mean from the sales experiment control group at baseline, with a positive score indicating a lower level of depression. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Only respondents 10 years of age or older were asked questions regarding their mental health.

[2] Regressions include block level fixed effects. Standard errors are clustered at the village level.

[3] Samples include all respondents present at endline for whom we have the respective outcomes data. We did not measure the mental health of the "Infant (6-24 months)" category, so this category is excluded from the table. Panel A & B: Samples exclude free DFS households. Panels C & D: Samples include all respondents located in the sales experiment control group and free DFS households.

[4] For the list of controls included, please refer to note 3 under Table 5.

[5] The regression also includes a dummy for which type of depression screen was performed (child or adult) and baseline standardized mental health score.

Table 8: Currently Using or Penultimately Used DFS - First Stage for Salt Consumption 2SLS - By Anemia Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)	Infant (6-24 mths)
Panel A: All								
Sales village	0.164***	0.167***	0.164***	0.168***	0.165***	0.163***	0.167***	0.154***
	[0.019]	[0.019]	[0.019]	[0.020]	[0.019]	[0.021]	[0.020]	[0.031]
Free DFS	0.495***	0.478***	0.512***	0.485***	0.494***	0.531***	0.483***	0.490***
	[0.045]	[0.049]	[0.042]	[0.048]	[0.046]	[0.047]	[0.049]	[0.062]
P-value: Sales Experiment + Free DFS Household = 0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	0.049	0.049	0.049	0.049	0.048	0.051	0.050	0.058
Observations	28126	14208	13564	10124	12579	4525	6922	1089
Panel B: BL Anemic Only								
Sales village	0.165***	0.161***	0.170***	0.164***	0.171***	0.152***	0.171***	
	[0.020]	[0.021]	[0.022]	[0.025]	[0.021]	[0.024]	[0.022]	
Free DFS	0.522***	0.502***	0.554***	0.496***	0.504***	0.581***	0.507***	
	[0.048]	[0.053]	[0.046]	[0.053]	[0.055]	[0.052]	[0.058]	
P-value: Sales Experiment + Free DFS Household = 0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Control Mean	0.047	0.050	0.042	0.051	0.044	0.050	0.047	
Observations	8923	5535	3373	2635	3833	2150	2896	

[1] The first and third lines of each panel report the coefficients on the sales village and the free DFS household treatment variables from a regression with currently using or penultimately used DFS as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[2] Regressions include block-level fixed effects. Standard errors are clustered at the village level. The regressions include all respondents who were present at endline for whom we have the respective outcomes data.

[3] For the list of controls included, please refer to note 3 under Table 5.

Table 9: All Health Outcomes - 2SLS Estimates - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel A: Hemoglobin							
Used DFS (currently or penultimately)	0.043	0.061	-0.004	0.127	0.048	-0.176	0.149
	[0.090]	[0.097]	[0.110]	[0.118]	[0.107]	[0.150]	[0.111]
Control Mean	12.056	11.416	12.769	11.492	12.575	11.982	11.533
Observations	23404	12057	11119	8605	10387	4159	5939
Panel B: Hemoglobin (BL Anemic Only)							
Used DFS (currently or penultimately)	0.057	0.161	-0.060	0.030	0.142	-0.056	0.233
	[0.130]	[0.138]	[0.172]	[0.201]	[0.166]	[0.197]	[0.152]
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090
Observations	7729	4839	2875	2377	3317	2012	2549
Panel C: Anemia							
Used DFS (currently or penultimately)	-0.018	-0.009	-0.022	-0.029	-0.025	0.017	-0.031
	[0.029]	[0.032]	[0.034]	[0.038]	[0.032]	[0.044]	[0.038]
Control Mean	0.491	0.597	0.375	0.472	0.457	0.605	0.617
Observations	23215	12057	11119	8604	10387	4159	5939
Panel D: Anemia (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.013	-0.013	-0.018	0.026	-0.034	-0.051	-0.053
	[0.041]	[0.045]	[0.050]	[0.062]	[0.050]	[0.050]	[0.052]
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740
Observations	7729	4839	2875	2377	3317	2012	2549
Panel E: Cognition Score							
Used DFS (currently or penultimately)	0.032	0.007	0.042	0.080	-0.005	-0.009	-0.006
	[0.051]	[0.060]	[0.057]	[0.076]	[0.062]	[0.072]	[0.067]
Control Mean	-0.113	-0.343	0.118	0.258	-0.442	-0.022	-0.612
Observations	19932	10090	9619	7284	8926	3484	5053

Continued on the next page...

Table 9: All Health Outcomes - 2SLS Estimates - Results by Gender and Age (continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel F: Cognition Score (BL Anemic Only)							
Used DFS (currently or penultimately)	0.011	-0.051	0.114	0.073	-0.056	-0.094	-0.072
	[0.061]	[0.065]	[0.087]	[0.108]	[0.073]	[0.088]	[0.074]
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639
Observations	6302	3926	2364	1819	2824	1640	2153
Panel G: Physical Fitness Score							
Used DFS (currently or penultimately)	-0.054	-0.096	-0.009	-0.037	-0.094	0.081	-0.132
	[0.096]	[0.109]	[0.103]	[0.165]	[0.103]	[0.153]	[0.113]
Control Mean	-0.422	-0.537	-0.259	-0.313	-0.396	-0.566	-0.500
Observations	10278	6082	4189	1623	6342	2312	4157
Panel H: Physical Fitness Score (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.051	-0.080	0.035	0.273	-0.145	0.016	-0.136
	[0.125]	[0.140]	[0.159]	[0.264]	[0.135]	[0.216]	[0.140]
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	3630	2584	1043	444	2170	1016	1800

Continued on the next page...

Table 9: All Health Outcomes - 2SLS Estimates - Results by Gender and Age (continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel I: Mental Health Score							
Used DFS (currently or penultimately)	-0.069	-0.022	-0.129	-0.137	-0.045	-0.077	-0.007
	[0.064]	[0.089]	[0.071]	[0.141]	[0.070]	[0.090]	[0.100]
Control Mean	-0.090	-0.315	0.170	-0.174	0.006	-0.264	-0.220
Observations	17764	9536	8210	3102	10457	4200	6021
Panel J: Mental Health Score (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.041	-0.038	-0.051	0.120	-0.011	-0.172	0.011
	[0.099]	[0.116]	[0.116]	[0.213]	[0.117]	[0.124]	[0.122]
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	6164	4096	2061	836	3314	2014	2557

[1] Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Regressions include block-level fixed effects. Standard errors are clustered at the village level. The regressions include all respondents who were present at endline, and for whom we have the respective outcomes data. We did not measure the physical fitness or mental health of the "Infant (6-24 months)" category, so this category is excluded from the table. Currently using DFS or penultimately used DFS is instrumented by a dummy for sales experiment and a dummy for free DFS households.

[2] For the list of controls used, please refer to note 3 under Table 5. For each outcome, the respective baseline variable was also included (except for Panels C-D, for which baseline hemoglobin level was used instead). For Panels A - D, the hemocue machine used to take the hemoglobin measurement was also included as a control. Anemia thresholds are determined by age and gender, so a dummy for each relevant age group is included in the regressions for anemia. For Panels E-F, the regression also includes a dummy for which type of cognition test taken (infant, child, adult, elderly). For Panels G-H, the regression also includes a dummy for which type of physical fitness test performed (Queen Step Test or a series of test). For Panels I-J, the regression also includes a dummy for which type of depression screen was performed (child or adult).

[3] The cognition, physical fitness and mental health scores were standardized using the mean from the sales experiment control group at baseline. The cognition outcomes were standardized such that a positive score indicates a higher level of cognition. The physical fitness scores were standardized such that a positive score indicates being more physically fit. The score for the mental health outcomes were standardized such that a positive score indicates a lower level of depression.

[4] Q-values were not reported since none of the results are significant

Table 10: Hemoglobin & Anemia - Rajasthan Wheat Fortification Experiment - Results by Gender

	Adult Female				Adult Male			
	Hemoglobin Level		Anemia Status		Hemoglobin Level		Anemia Status	
	(1) Reduced Form (OLS)	(2) Average: Last 3 Months (IV)	(3) Reduced Form (OLS)	(4) Average: Last 3 Months (IV)	(5) Reduced Form (OLS)	(6) Average: Last 3 Months (IV)	(7) Reduced Form (OLS)	(8) Average: Last 3 Months (IV)
Control Group Means	10.888	10.882	0.732	0.734	12.805	12.771	0.507	0.512
Panel A: Basic Control								
Iron Treatment/Take-up	-0.031	-0.100	0.016	0.050	0.129	0.301	-0.039	-0.088
	[0.077]	[0.189]	[0.020]	[0.049]	[0.086]	[0.207]	[0.024]	[0.057]
Observations	3,890	3,362	3,890	3,362	3,527	3,154	3,527	3,154
Panel B: Controlling for Baseline Anemia								
Iron Treatment/Take-up	-0.024	-0.087	0.014	0.046	0.132	0.309	-0.040	-0.091
	[0.074]	[0.183]	[0.020]	[0.049]	[0.082]	[0.196]	[0.023]	[0.055]
Baseline Anemia	-0.836	-0.817	0.173	0.179	-0.861	-0.893	0.199	0.215
	[0.062]	[0.069]	[0.019]	[0.020]	[0.072]	[0.075]	[0.020]	[0.021]
Missing Baseline Anemia	-0.627	-0.656	0.101	0.121	-0.422	-0.396	0.071	0.077
	[0.087]	[0.088]	[0.025]	[0.027]	[0.090]	[0.095]	[0.027]	[0.027]
Observations	3,890	3,362	3,890	3,362	3,527	3,154	3,527	3,154

[1] Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Standard errors are clustered at the village level.
[2] The controls in Panel A include age, age-squared, and block-level fixed effects. Panel B includes missing values. All IV regressions in Panels A-B use original treatment status as the instrument.
[3] Average take-up in the last 3 months and during the treatment period is from the monthly continuous household survey.

Table 11: Hemoglobin & Anemia - OLS and 2SLS - For Adolescents (13-17 years)

	Hemoglobin						Anemia					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample:	All	All Anemic	Male	Male Anemic	Female	Female Anemic	All	All Anemic	Male	Male Anemic	Female	Female Anemic
Panel A: Sales Experiment (OLS)												
Sales village	0.039	0.134	0.001	0.163	0.091	0.136	-0.007	-0.033	0.008	-0.018	-0.027	-0.044
	[0.055]	[0.099]	[0.067]	[0.146]	[0.063]	[0.104]	[0.018]	[0.029]	[0.023]	[0.051]	[0.023]	[0.033]
Q-value	(1.000)	(0.625)	(1.000)	(0.625)	(1.000)	(0.625)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	4015	1276	1986	499	2026	777	4012	1276	1986	499	2026	777
Panel B: Free DFS vs. Control (OLS)												
Free DFS	0.139	0.042	0.025	-0.138	0.169	0.072	-0.054*	0.007	-0.021	0.020	-0.089*	-0.014
	[0.103]	[0.178]	[0.132]	[0.261]	[0.118]	[0.179]	[0.026]	[0.041]	[0.039]	[0.073]	[0.037]	[0.060]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	2485	792	1243	317	1241	475	2484	792	1243	317	1241	475
Panel C: Experiment within Free DFS Villages (OLS)												
Free DFS	0.048	-0.118	0.012	-0.446	0.004	-0.016	-0.051	0.006	-0.013	0.094	-0.068	0.097
	[0.128]	[0.301]	[0.170]	[0.727]	[0.159]	[0.366]	[0.038]	[0.076]	[0.054]	[0.187]	[0.058]	[0.097]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	704	238	361	104	343	134	704	238	361	104	343	134
Panel D: 2SLS Estimates												
Used DFS (currently or penultimately)	0.401*	0.248	0.356	0.007	0.396	0.182	-0.121*	0.012	-0.071	0.069	-0.152*	0.005
	[0.188]	[0.313]	[0.218]	[0.358]	[0.204]	[0.318]	[0.050]	[0.070]	[0.064]	[0.101]	[0.065]	[0.089]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(0.370)	(1.000)	(0.370)	(1.000)	(0.370)	(1.000)
Control Mean	12.386	11.644	12.386	11.644	12.386	11.644	0.467	0.636	0.467	0.636	0.467	0.636
Observations	2682	849	1322	343	1358	506	2680	849	1322	343	1358	506

[1] The first line of each panel reports the coefficient on the treatment variable in a regression with hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.
[2] Panel A, B & D: regressions include block-level fixed effects. Standard errors are clustered at the village level. Panel A: a control for free DFS village is also included. Panel C: regressions include village-level fixed effects. Standard errors are clustered at the household level. Panel D: currently using DFS or penultimately used DFS is instrumented by a dummy for sales experiment and a dummy for free DFS households.
[3] Samples include all respondents present at endline for whom we have the respective outcomes data. Panel A: sample excludes free DFS households. Panel B: sample includes all respondents in the sales experiment control group and free DFS households. Panel C: the sample includes all respondents located in the villages where the Free DFS experiment took place. Panel D: the sample includes all households in the study.
[3] For the list of controls used, please refer to notes 3 and 4 under Table 4a.

8 Appendix

Table 12: Attrition

	Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control [2] vs [1]	Free DFS vs Control [3] vs [1]	Free DFS vs Non-Free [3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Household Lost to Attrition	0.039	0.052	0.023	0.044	0.043	0.050	0.065
	(0.194)	(0.222)	(0.150)	(0.206)	(5570)	(3436)	(932)
Respondent Lost to Attrition	0.192	0.191	0.150	0.185	0.889	0.000	0.018
	(0.394)	(0.393)	(0.357)	(0.388)	(36711)	(22582)	(6302)
In Baseline Sample	0.819	0.819	0.794	0.814	0.905	0.032	0.085
	(0.385)	(0.385)	(0.405)	(0.389)	(44839)	(27778)	(7832)
Missing Endline Hemoglobin	0.162	0.157	0.129	0.152	0.422	0.000	0.075
	(0.369)	(0.364)	(0.335)	(0.359)	(36711)	(22582)	(6302)

[1] Standard deviations are in parentheses in columns 1-4 and the observation numbers for the respective samples are in parentheses in columns 5-7. P-values are calculated using block-level fixed effects (columns 5 and 6) or village-level fixed effects (column 7). Standard errors are clustered at the village level (columns 5 and 6) or at the household level in (column 7).

Table 13: Balance Table for Those Observed at Endline

	Baseline Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control	Free DFS vs Control	Free DFS vs Non-Free
					[2] vs [1]	[3] vs [1]	[3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Individual-Level							
Female	0.524	0.519	0.516	0.527	0.213	0.415	0.292
	(0.499)	(0.500)	(0.500)	(0.499)	(29679)	(18368)	(5239)
Age	28.117	27.803	26.777	28.166	0.285	0.002	0.005
	(20.116)	(20.085)	(19.718)	(20.392)	(29290)	(18123)	(5183)
Child (<15 yrs)	0.337	0.342	0.358	0.343	0.535	0.075	0.309
	(0.473)	(0.474)	(0.479)	(0.475)	(29680)	(18369)	(5239)
Adolescent (13-17 yrs)	0.074	0.073	0.068	0.073	0.656	0.145	0.622
	(0.262)	(0.260)	(0.251)	(0.259)	(29680)	(18369)	(5239)
Adult (15-49 yrs)	0.417	0.415	0.410	0.409	0.721	0.689	0.655
	(0.493)	(0.493)	(0.492)	(0.492)	(29680)	(18369)	(5239)
Elderly (50+ yrs)	0.183	0.179	0.160	0.189	0.456	0.004	0.006
	(0.387)	(0.384)	(0.367)	(0.391)	(29680)	(18369)	(5239)
Infant (6-24 mths)	0.030	0.030	0.032	0.030	0.602	0.398	0.521
	(0.169)	(0.172)	(0.177)	(0.170)	(37808)	(23565)	(6769)
Female Adult (15-49 yrs)	0.186	0.185	0.183	0.181	0.670	0.611	0.859
	(0.389)	(0.388)	(0.387)	(0.385)	(37808)	(23565)	(6769)
Anemia	0.447	0.442	0.454	0.455	0.713	0.935	0.824
	(0.497)	(0.497)	(0.498)	(0.498)	(26550)	(16400)	(4635)
Severe Anemia	0.007	0.008	0.011	0.007	0.445	0.049	0.145
	(0.082)	(0.086)	(0.106)	(0.081)	(26550)	(16400)	(4635)
Hemoglobin	12.191	12.192	12.084	12.172	0.940	0.260	0.127
	(1.808)	(1.802)	(1.876)	(1.808)	(26550)	(16400)	(4635)
Cognition Score	-0.010	-0.025	-0.092	0.000	0.520	0.020	0.006
	(0.993)	(0.973)	(0.974)	(0.969)	(24710)	(15264)	(4307)
Mental Health Score	-0.006	-0.009	0.011	-0.058	0.913	0.463	0.250
	(0.997)	(1.011)	(0.995)	(1.035)	(19923)	(12328)	(3437)
Physical Fitness Score	0.017	0.028	-0.046	-0.061	0.692	0.103	0.862
	(0.885)	(0.919)	(0.955)	(0.924)	(12013)	(7486)	(2020)
Body Mass Index (BMI)	18.132	18.092	17.917	18.228	0.628	0.109	0.021
	(4.283)	(4.290)	(4.690)	(4.435)	(26868)	(16598)	(4684)
5+ Education Years	0.523	0.521	0.474	0.518	0.838	0.026	0.046
	(0.499)	(0.500)	(0.499)	(0.500)	(27016)	(16678)	(4739)

Continued on the next page...

Table 13: Balance Table for Those Observed at Endline (continued)

	Baseline Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control	Free DFS vs Control	Free DFS vs Non-Free
					[2] vs [1]	[3] vs [1]	[3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Household-Level							
30-day Consumption per Capita	2225.061	2264.749	2128.890	2222.407	0.688	0.361	0.431
	(2596.529)	(2200.786)	(2018.295)	(2023.532)	(4222)	(2621)	(704)
Head: 5+ Education years	0.565	0.566	0.491	0.577	0.988	0.061	0.037
	(0.496)	(0.496)	(0.501)	(0.495)	(3738)	(2305)	(635)
At least one, Elderly and Anemic	0.483	0.466	0.462	0.489	0.305	0.334	0.438
	(0.500)	(0.499)	(0.499)	(0.500)	(5320)	(3309)	(900)
Number of Members	6.734	6.872	7.141	6.871	0.221	0.154	0.254
	(3.453)	(3.964)	(4.066)	(4.193)	(4248)	(2638)	(708)
Only Immediate Family	0.354	0.363	0.344	0.351	0.567	0.745	0.539
	(0.478)	(0.481)	(0.476)	(0.478)	(4245)	(2636)	(707)
Household Wealth Index	-0.028	-0.028	0.002	-0.002	0.988	0.621	0.877
	(0.654)	(0.569)	(0.781)	(0.616)	(4248)	(2638)	(708)
Split since Baseline	0.152	0.166	0.231	0.185	0.236	0.000	0.083
	(0.359)	(0.372)	(0.422)	(0.389)	(5320)	(3309)	(900)

[1] The table presents summary statistics for only the people who we observed at endline (non-attritors). The sample used varies by column as indicated by the column headers.

[2] The sample for "Individual-Level Variables" includes all baseline respondents from within the respective subsets that we observe at endline. "Household-Level Variables" includes all households from baseline that we see at endline within the respective subsets.

[3] The cognition, physical fitness and mental health scores were standardized using the mean from the sales experiment's control group at baseline. The cognition outcomes were standardized such that a positive score indicates a higher level of cognition. The physical fitness scores were standardized such that a positive score indicates being more physically fit. The score for the mental health scores were standardized such that a positive score indicates a lower level of depression.

[4] Standard deviations are in parentheses in columns 1-4 and the observation numbers for the respective samples are in parentheses in columns 5-7. P-values are calculated using block-level fixed effects (columns 5 and 6) or village-level fixed effects (column 7). Standard errors are clustered at the village level (columns 5 and 6) or at the household level in (column 7).

9 Online Appendix

Table A1: BASELINE RESPONDENTS ONLY - Take Up Statistics by Experimental Group

	Baseline Variable Mean and Standard Deviation				Difference in Means (P-values & Observations)		
	Villages 200 Control Households All	Villages 200 Sales Households Non-Free DFS	Villages 62 Free DFS Households Free DFS	Villages 62 Free DFS Households Non-Free DFS	Sales vs Control [2] vs [1]	Free DFS vs Control [3] vs [1]	Free DFS vs Non-Free [3] vs [4]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Currently Using DFS	0.034	0.145	0.614	0.157	0.000	0.000	0.000
	(0.182)	(0.353)	(0.487)	(0.364)	(5736)	(3591)	(1054)
Used DFS (currently or penultimately)	0.046	0.219	0.760	0.232	0.000	0.000	0.000
	(0.210)	(0.414)	(0.428)	(0.423)	(3704)	(2356)	(716)
Tried DFS in the Past Year	0.108	0.426	0.746	0.423	0.000	0.000	0.000
	(0.310)	(0.495)	(0.436)	(0.495)	(4022)	(2566)	(780)
Times Purchased DFS in past year	0.781	3.323	6.706	3.682	0.000	0.000	0.000
	(4.013)	(7.107)	(7.409)	(7.878)	(4022)	(2566)	(780)

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample includes all baseline households for whom we have the respective outcomes data.

[2] Standard deviations are in parentheses in columns 1-4 and the observation numbers for the respective samples are in parentheses in columns 5-7. P-values are calculated using block-level fixed effects (columns 5 and 6) or village-level fixed effects (column 7). Standard errors are clustered at the village level (columns 5 and 6) or at the household level in (column 7).

Table A2a: BASELINE RESPONDENTS ONLY - Hemoglobin & Anemia - Sales Experiment - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)	(8) Infant (6-24 mths)
Panel A: Hemoglobin								
Sales village	0.035	0.002	0.062	0.054	0.044	-0.026	-0.005	0.817
	[0.031]	[0.033]	[0.037]	[0.039]	[0.040]	[0.049]	[0.041]	[1.164]
Q-value	(0.747)	(1.000)	(0.541)	(0.661)	(0.747)	(1.000)	(1.000)	(1.000)
Control Mean	12.163	11.497	12.891	11.725	12.578	11.977	11.561	10.747
Panel B: Anemia								
Sales village	-0.002	0.007	-0.010	-0.012	-0.004	0.022	0.004	0.029
	[0.009]	[0.012]	[0.010]	[0.015]	[0.011]	[0.013]	[0.014]	[0.354]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	0.476	0.584	0.358	0.425	0.456	0.604	0.612	0.526
Observations	28780	14946	13769	9770	13065	5870	7364	39
Panel C: Hemoglobin (BL Anemic)								
Sales village	0.074	0.043	0.131*	0.128*	0.057	0.042	0.028	
	[0.043]	[0.044]	[0.058]	[0.058]	[0.058]	[0.069]	[0.057]	
Q-value	(0.541)	(0.747)	(0.300)	(0.300)	(0.747)	(1.000)	(1.000)	
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090	
Panel D: Anemia (BL Anemic)								
Sales village	-0.017	-0.006	-0.036*	-0.027	-0.019	-0.006	-0.009	
	[0.012]	[0.014]	[0.016]	[0.022]	[0.015]	[0.016]	[0.017]	
Q-value	(1.000)	(1.000)	(0.935)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740	
Observations	11503	7166	4317	3534	4909	3035	3718	

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in this table includes all baseline respondents for whom we have the respective outcomes data and excludes free DFS households.

[2] The first line of each panel reports the coefficient on the sales treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Regressions include block-level fixed effects and a control for free DFS village. Standard errors are clustered at the village level.

[4] Controls include the following variables from endline: age, age-squared, hemocue machine used to take hemoglobin measurement, and a dummy for household split between baseline and endline. The following baseline variables are also included: hemoglobin concentration, completed 5th standard or higher, body mass index (BMI), and household wealth index. For children under 10, the value for whether the head of the household completed 5th standard or higher is used instead of the value for the child.

[5] Anemia thresholds are determined by age and gender, therefore, a dummy for each relevant age group is included in the regressions for anemia.

Table A2b: BASELINE RESPONDENTS ONLY - Hemoglobin & Anemia - Free DFS Households Compared to Households in Control Villages - Results by Gender and Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)	Infant (6-24 mths)
Panel A: Hemoglobin								
Free DFS	0.038	0.025	0.013	0.114	0.036	-0.139	0.031	-3.324
	[0.051]	[0.055]	[0.061]	[0.059]	[0.060]	[0.084]	[0.064]	[6.263]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	12.163	11.497	12.891	11.725	12.578	11.977	11.561	10.747
Panel B: Anemia								
Free DFS	-0.014	-0.008	-0.014	-0.034	-0.007	0.007	-0.004	2.143
	[0.017]	[0.019]	[0.019]	[0.023]	[0.018]	[0.024]	[0.022]	[2.839]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	0.476	0.584	0.358	0.425	0.456	0.604	0.612	0.526
Observations	17861	9285	8528	6089	8076	3633	4579	25
Panel C: Hemoglobin (BL Anemic)								
Free DFS	0.019	0.084	-0.097	0.075	0.044	-0.123	0.092	
	[0.073]	[0.075]	[0.102]	[0.098]	[0.092]	[0.116]	[0.089]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090	
Panel D: Anemia (BL Anemic)								
Free DFS	0.000	-0.014	0.026	0.007	-0.001	-0.013	-0.021	
	[0.023]	[0.024]	[0.030]	[0.034]	[0.028]	[0.028]	[0.029]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740	
Observations	7189	4459	2716	2200	3043	1913	2287	

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in this table includes all baseline respondents for whom we have the respective outcomes data and includes only households in the sales experiment control group or Free DFS experiment.

[2] The first line of each panel reports the coefficient on the free DFS household treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Regression include block-level fixed effects. Standard errors are clustered at the village level.

[4] For the list of controls included, please refer to note 4 and 5 under Table A2a.

Table A2c: BASELINE RESPONDENTS ONLY - Hemoglobin & Anemia - Free DFS Experiment within Free DFS Villages - Results by Gender and Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)	Infant (6-24 mths)
Panel A: Hemoglobin								
Free DFS	0.003	0.013	-0.051	0.036	0.033	-0.135	0.085	0.000
	[0.046]	[0.054]	[0.062]	[0.059]	[0.066]	[0.105]	[0.077]	[.]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	12.213	11.514	12.992	11.797	12.624	12.028	11.521	10.840
Panel B: Anemia								
Free DFS	-0.011	-0.014	-0.008	-0.026	-0.009	-0.007	-0.011	0.000
	[0.015]	[0.020]	[0.018]	[0.025]	[0.019]	[0.029]	[0.027]	[.]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.468	0.580	0.345	0.414	0.448	0.599	0.612	0.400
Observations	5103	2654	2435	1793	2296	996	1287	11
Panel C: Hemoglobin (BL Anemic)								
Free DFS	0.010	0.088	-0.181	-0.043	0.074	-0.105	0.158	
	[0.072]	[0.080]	[0.127]	[0.101]	[0.110]	[0.159]	[0.114]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	11.413	11.030	12.087	11.466	11.528	11.165	11.016	
Panel D: Anemia (BL Anemic)								
Free DFS	-0.004	-0.030	0.020	-0.003	-0.027	-0.006	-0.053	
	[0.022]	[0.026]	[0.038]	[0.041]	[0.032]	[0.035]	[0.036]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	0.663	0.734	0.542	0.523	0.698	0.773	0.771	
Observations	2057	1288	763	657	867	524	651	

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in this table includes only baseline respondents located in the villages where the Free DFS experiment took place and for whom we have the respective outcomes data.

[2] The first line of each panel reports the coefficient on the free DFS household treatment variable from a regression with either hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Regressions include village-level fixed effects. Standard errors are clustered at the household level.

[4] For the list of controls included, please refer to note 4 and 5 under Table A2a.

Table A3: BASELINE RESPONDENTS ONLY - Standardized Cognition Scores - Sales and Free DFS Experiments - Results by Gender and Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)	Infant (6-24 mths)
Panel A: Sales Experiment								
Sales village	0.036	0.030	0.041*	0.032	0.031	0.054	0.034	-0.026
	[0.020]	[0.024]	[0.020]	[0.027]	[0.023]	[0.029]	[0.026]	[0.298]
Q-value	(0.201)	(0.232)	(0.182)	(0.232)	(0.232)	(0.182)	(0.232)	(0.354)
Control Mean	-0.107	-0.355	0.151	0.319	-0.438	-0.019	-0.610	0.160
Observations	24085	12286	11743	7864	11250	4922	6266	37
Panel B: Sales Experiment (BL Anemic Only)								
Sales village	0.054*	0.036	0.077**	0.043	0.033	0.079*	0.037	
	[0.025]	[0.027]	[0.030]	[0.036]	[0.027]	[0.037]	[0.029]	
Q-value	(0.174)	(0.232)	(0.174)	(0.232)	(0.232)	(0.174)	(0.232)	
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639	
Observations	9397	5793	3586	2725	4174	2480	3137	
Panel C: Free DFS vs. Control								
Free DFS	-0.010	-0.004	-0.024	0.038	-0.029	-0.080	-0.018	2.434
	[0.027]	[0.029]	[0.032]	[0.038]	[0.031]	[0.049]	[0.035]	[2.778]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.107	-0.355	0.151	0.319	-0.438	-0.019	-0.610	0.160
Observations	14922	7590	7292	4858	7005	3013	3902	23
Panel D: Free DFS vs. Control (BL Anemic Only)								
Free DFS	-0.020	-0.037	0.011	0.037	-0.042	-0.105	-0.041	
	[0.033]	[0.035]	[0.053]	[0.056]	[0.040]	[0.063]	[0.041]	
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639	
Observations	5871	3621	2238	1677	2623	1547	1954	

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in Panel A & B excludes free DFS households. The sample in Panel C & D includes only households in the sales experiment's control group or free DFS households. Panels A-D include all respondents present at baseline for whom we have the respective outcomes data within the specified samples.

[2] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized cognition score as the outcome. The cognition scores were standardized using the mean from the sales experiment's control group at baseline, with a positive score indicating a higher level of cognition. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Regressions include block-level fixed effects. Standard errors are clustered at the village level.

[4] Controls include the following variables from endline: age, age-squared, and a dummy for household split between baseline and endline. The following baseline variables are also included: hemoglobin concentration, completed 5th standard or higher, body mass index (BMI), and household wealth index. For children under 10, the value for whether the head of the household completed 5th standard or higher is used instead of the value for the child.

[5] The regression also includes a dummy for which type of cognition test taken (infant, child, adult, elderly) and baseline standardized cognition score.

Table A4: BASELINE RESPONDENTS ONLY - Standardized Physical Fitness Performance - Sales and Free DFS Experiments - Results by Gender and Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)
Panel A: Sales Experiment							
Sales village	0.006	0.034	-0.045	-0.007	0.033	-0.070	0.071
	[0.033]	[0.037]	[0.040]	[0.054]	[0.036]	[0.049]	[0.042]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.435	-0.555	-0.264	-0.330	-0.410	-0.564	-0.517
Observations	13702	7987	5701	2248	8146	3308	5244
Panel B: Sales Experiment (BL Anemic Only)							
Sales village	-0.017	0.007	-0.073	-0.082	0.011	-0.067	0.034
	[0.039]	[0.043]	[0.059]	[0.091]	[0.045]	[0.062]	[0.046]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	5555	3856	1693	748	3230	1577	2620

Continued on the next page...

Table A4: BASELINE RESPONDENTS ONLY - Standardized Physical Fitness Performance - Sales and Free DFS Experiments - Results by Gender and Age (continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel C: Free DFS vs. Control							
Free DFS	-0.071	-0.085	-0.070	-0.180*	-0.056	-0.014	-0.069
	[0.061]	[0.070]	[0.066]	[0.090]	[0.068]	[0.091]	[0.080]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.435	-0.555	-0.264	-0.330	-0.410	-0.564	-0.517
Observations	8470	4986	3475	1383	5064	2023	3287
Panel D: Free DFS vs. Control (BL Anemic Only)							
Free DFS	-0.067	-0.076	-0.070	0.039	-0.132	-0.004	-0.115
	[0.071]	[0.081]	[0.109]	[0.142]	[0.076]	[0.126]	[0.081]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	3443	2394	1046	454	2008	981	1616

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in Panel A & B excludes free DFS households. The sample in Panel C & D includes only households located in the sales experiment control group or free DFS households. Panels A-D include all respondents present at baseline for whom we have the respective outcomes data within the specified samples.

[2] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized physical fitness score as the outcome. The physical fitness scores were standardized using the mean from the sales experiment control group at baseline, with a positive score indicating a higher level of physical fitness. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Only respondents 10 years of age or older were eligible to perform physical fitness tests. We did not measure the physical fitness of the "Infant (6-24 months)" category, so this category is excluded from the table.

[3] Regressions include block level fixed effects. Standard errors are clustered at the village level.

[4] For the list of controls included, please refer to note 4 under Table A3.

[5] The regression also includes a dummy for which type of physical fitness test performed (Queen Step Test or a series of tests) and baseline standardized physical fitness score.

Table A5: BASELINE RESPONDENTS ONLY - Standardized Mental Health Scores - Sales and Free DFS Experiments - Results by Gender and Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)
Panel A: Sales Experiment							
Sales village	0.003	0.014	-0.017	-0.089	0.022	0.020	0.050
	[0.023]	[0.034]	[0.024]	[0.051]	[0.024]	[0.033]	[0.036]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.099	-0.330	0.170	-0.165	-0.001	-0.268	-0.233
Observations	22980	12317	10636	4093	13039	5848	7459
Panel B: Sales Experiment (BL Anemic Only)							
Sales village	0.022	0.030	-0.001	-0.074	0.045	0.017	0.074
	[0.031]	[0.040]	[0.037]	[0.081]	[0.035]	[0.044]	[0.043]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	9210	6061	3137	1320	4888	3002	3739

Continued on the next page...

Table A5: BASELINE RESPONDENTS ONLY - Standardized Mental Health Scores - Sales and Free DFS Experiments - Results by Gender and Age (continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel C: Free DFS vs. Control							
Free DFS	-0.021	0.007	-0.058	-0.005	-0.015	-0.039	-0.011
	[0.037]	[0.050]	[0.042]	[0.083]	[0.042]	[0.052]	[0.062]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.099	-0.330	0.170	-0.165	-0.001	-0.268	-0.233
Observations	14186	7624	6540	2555	8035	3595	4616
Panel D: Free DFS vs. Control (BL Anemic Only)							
Free DFS	0.007	0.007	-0.017	0.110	-0.008	-0.028	-0.010
	[0.055]	[0.065]	[0.069]	[0.116]	[0.065]	[0.072]	[0.073]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	5697	3734	1955	803	3018	1876	2288

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. The sample in Panel A & B excludes free DFS households. The sample in Panel C & D includes only households located in the sales experiment control group or free DFS households. Panels A-D include all respondents present at baseline for whom we have the respective outcomes data within the specified samples.

[2] The first line of each panel reports the coefficient on the treatment variable from a regression with standardized mental health score as the outcome. The mental health scores were standardized using the mean from the sales experiment control group at baseline, with a positive score indicating a lower level of depression. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Only respondents 10 years of age or older were asked questions regarding their mental health. We did not measure the mental health of the "Infant (6-24 months)" category, so this category is excluded from the table.

[3] Regressions include block level fixed effects. Standard errors are clustered at the village level.

[4] For the list of controls included, please refer to note 4 under Table A3.

[5] The regression also includes a dummy for which type of depression screen was performed (child or adult) and baseline standardized mental health score.

Table A6: BASELINE RESPONDENTS ONLY - Currently Using or Penultimately Used DFS - First Stage for Salt Consumption 2SLS - By Anemia Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	Female	Male	Child (10-14 yrs)	Adult (15-49 yrs)	Elderly (50+ yrs)	Female Adult (15-49 yrs)	Infant (6-24 mths)
Panel A: All								
Sales village	0.170***	0.174***	0.166***	0.168***	0.171***	0.173***	0.172***	0.051
	[0.019]	[0.020]	[0.019]	[0.021]	[0.019]	[0.022]	[0.020]	[0.434]
Free DFS	0.512***	0.484***	0.540***	0.503***	0.507***	0.552***	0.495***	0.570
	[0.045]	[0.049]	[0.042]	[0.047]	[0.047]	[0.047]	[0.051]	[1.212]
P-value: Sales Experiment + Free DFS Household = 0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.549
Control Mean	0.049	0.049	0.049	0.051	0.048	0.050	0.050	0.000
Observations	22429	11506	10878	7286	10415	4160	5735	25
Panel B: BL Anemic Only								
Sales village	0.165***	0.161***	0.170***	0.164***	0.171***	0.152***	0.171***	
	[0.020]	[0.021]	[0.022]	[0.025]	[0.021]	[0.024]	[0.022]	
Free DFS	0.522***	0.502***	0.554***	0.496***	0.504***	0.581***	0.507***	
	[0.048]	[0.053]	[0.046]	[0.053]	[0.055]	[0.052]	[0.058]	
P-value: Sales Experiment + Free DFS Household = 0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Control Mean	0.047	0.050	0.042	0.051	0.044	0.050	0.047	
Observations	8923	5535	3373	2635	3833	2150	2896	

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. It includes all baseline respondents for whom we have the respective outcomes data.

[2] The first and third lines of each panel report the coefficients on the Sales village and Free DFS Household treatment variables from a regression with Currently Using or Penultimately Used DFS as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Regressions include block-level fixed effects. Standard errors are clustered at the village level.

[4] For the list of controls included, please refer to note 4 under Table A3.

Table A7: BASELINE RESPONDENTS ONLY - All Health Outcomes - 2SLS Estimates - Results by Gender and Age

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel A: Hemoglobin							
Used DFS (currently or penultimately)	0.044	0.004	0.025	0.127	0.073	-0.175	0.090
	[0.096]	[0.099]	[0.115]	[0.121]	[0.118]	[0.142]	[0.116]
Control Mean	12.163	11.497	12.891	11.725	12.578	11.977	11.561
Observations	18889	9881	8969	6452	8550	3829	4923
Panel B: Hemoglobin (BL Anemic Only)							
Used DFS (currently or penultimately)	0.057	0.161	-0.060	0.030	0.142	-0.056	0.233
	[0.130]	[0.138]	[0.172]	[0.201]	[0.166]	[0.197]	[0.152]
Control Mean	11.364	11.027	11.917	11.248	11.544	11.205	11.090
Observations	7729	4839	2875	2377	3317	2012	2549
Panel C: Anemia							
Used DFS (currently or penultimately)	-0.012	0.010	-0.024	-0.030	-0.017	0.022	-0.015
	[0.031]	[0.036]	[0.035]	[0.042]	[0.034]	[0.043]	[0.042]
Control Mean	0.476	0.584	0.358	0.425	0.456	0.604	0.612
Observations	18889	9881	8969	6452	8550	3829	4923
Panel D: Anemia (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.013	-0.013	-0.018	0.026	-0.034	-0.051	-0.053
	[0.041]	[0.045]	[0.050]	[0.062]	[0.050]	[0.050]	[0.052]
Control Mean	0.669	0.722	0.582	0.561	0.674	0.785	0.740
Observations	7729	4839	2875	2377	3317	2012	2549
Panel E: Cognition Score							
Used DFS (currently or penultimately)	0.025	0.013	0.025	0.098	-0.032	-0.001	-0.036
	[0.051]	[0.057]	[0.058]	[0.078]	[0.060]	[0.071]	[0.065]
Control Mean	-0.107	-0.355	0.151	0.319	-0.438	-0.019	-0.610
Observations	15865	8162	7669	5253	7362	3206	4192

Continued on the next page...

Table A7: BASELINE RESPONDENTS ONLY - All Health Outcomes - 2SLS Estimates - Results by Gender and Age (continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel F: Cognition Score (BL Anemic Only)							
Used DFS (currently or penultimately)	0.011	-0.051	0.114	0.073	-0.056	-0.094	-0.072
	[0.061]	[0.065]	[0.087]	[0.108]	[0.073]	[0.088]	[0.074]
Control Mean	-0.253	-0.450	0.063	0.105	-0.550	-0.137	-0.639
Observations	6302	3926	2364	1819	2824	1640	2153
Panel G: Physical Fitness Score							
Used DFS (currently or penultimately)	-0.077	-0.137	-0.008	-0.013	-0.101	-0.031	-0.157
	[0.100]	[0.117]	[0.102]	[0.171]	[0.108]	[0.154]	[0.124]
Control Mean	-0.435	-0.555	-0.264	-0.330	-0.410	-0.564	-0.517
Observations	8871	5260	3604	1439	5295	2137	3498
Panel H: Physical Fitness Score (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.051	-0.080	0.035	0.273	-0.145	0.016	-0.136
	[0.125]	[0.140]	[0.159]	[0.264]	[0.135]	[0.216]	[0.140]
Control Mean	-0.487	-0.558	-0.330	-0.337	-0.433	-0.661	-0.479
Observations	3630	2584	1043	444	2170	1016	1800

Continued on the next page...

Table A7: BASELINE RESPONDENTS ONLY - All Health Outcomes - 2SLS Estimates - Results by Gender and Age(continued)

Sample:	(1) All	(2) Female	(3) Male	(4) Child (10-14 yrs)	(5) Adult (15-49 yrs)	(6) Elderly (50+ yrs)	(7) Female Adult (15-49 yrs)
Panel I: Mental Health Score							
Used DFS (currently or penultimately)	-0.074	-0.020	-0.141	-0.104	-0.054	-0.071	-0.024
	[0.069]	[0.092]	[0.074]	[0.149]	[0.077]	[0.089]	[0.106]
Control Mean	-0.099	-0.330	0.170	-0.165	-0.001	-0.268	-0.233
Observations	15224	8198	7009	2742	8611	3870	4989
Panel J: Mental Health Score (BL Anemic Only)							
Used DFS (currently or penultimately)	-0.041	-0.038	-0.051	0.120	-0.011	-0.172	0.011
	[0.099]	[0.116]	[0.116]	[0.213]	[0.117]	[0.124]	[0.122]
Control Mean	-0.194	-0.361	0.123	-0.176	-0.108	-0.339	-0.247
Observations	6164	4096	2061	836	3314	2014	2557

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. It includes all baseline respondents for whom we have the respective outcomes data.

[2] Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001. Regressions include block-level fixed effects. Standard errors are clustered at the village level. Currently using DFS or Penultimately Used DFS is instrumented by a dummy for sales experiment and a dummy for free DFS households.

[3] For the list of controls used, please refer to note 4 under Table A3. For each outcome, the respective baseline variable was also included (except for Panels C-D, for which baseline hemoglobin level was used instead). For Panels A - D, the hemocue machine used to take the hemoglobin measurement was also included as a control. Anemia thresholds are determined by age and gender, therefore, a dummy for each relevant age group is included in the regressions for anemia. For Panels E-F, the regression also includes a dummy for which type of cognition test taken (infant, child, adult, elderly). For Panels G-H, the regression also includes a dummy for which type of physical fitness test performed (Queen Step Test or a series of test). For Panels I-J, the regression also includes a dummy for which type of depression screen was performed (child or adult).

[4] The cognition, physical fitness and mental health scores were standardized using the mean from the sales experiment control group at baseline. The cognition outcomes were standardized such that a positive score indicates a higher level of cognition. The physical fitness scores were standardized such that a positive score indicates being more physically fit. The score for the mental health outcomes were standardized such that a positive score indicates a lower level of depression.

[5] Q-values were not reported since none of the results are significant

Table A8: BASELINE RESPONDENTS ONLY - Hemoglobin & Anemia - OLS and 2SLS - For Adolescents (13-17 years)

	Hemoglobin						Anemia					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample:	All	All Anemic	Male	Male Anemic	Female	Female Anemic	All	All Anemic	Male	Male Anemic	Female	Female Anemic
Panel A: Sales Experiment (OLS)												
Sales village	0.042	0.134	0.014	0.163	0.048	0.136	0.002	-0.033	0.011	-0.018	-0.009	-0.044
	[0.059]	[0.099]	[0.070]	[0.146]	[0.066]	[0.104]	[0.019]	[0.029]	[0.024]	[0.051]	[0.024]	[0.033]
Q-value	(1.000)	(0.661)	(1.000)	(0.661)	(1.000)	(0.661)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	3551	1276	1769	499	1782	777	3551	1276	1769	499	1782	777
Panel B: Free DFS vs. Control (OLS)												
Free DFS	0.146	0.042	-0.023	-0.138	0.197	0.072	-0.051	0.007	-0.005	0.020	-0.091*	-0.014
	[0.106]	[0.178]	[0.139]	[0.261]	[0.130]	[0.179]	[0.028]	[0.041]	[0.041]	[0.073]	[0.040]	[0.060]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	2185	792	1093	317	1092	475	2185	792	1093	317	1092	475
Panel C: Experiment within Free DFS Villages (OLS)												
Free DFS	0.075	-0.118	0.041	-0.446	0.050	-0.016	-0.052	0.006	-0.014	0.094	-0.059	0.097
	[0.135]	[0.301]	[0.177]	[0.727]	[0.175]	[0.366]	[0.041]	[0.076]	[0.057]	[0.187]	[0.063]	[0.097]
Q-value	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	636	238	330	104	306	134	636	238	330	104	306	134
Panel D: 2SLS Estimates												
Used DFS (currently or penultimately)	0.411*	0.248	0.254	0.007	0.419*	0.182	-0.105*	0.012	-0.040	0.069	-0.145*	0.005
	[0.186]	[0.313]	[0.219]	[0.358]	[0.204]	[0.318]	[0.050]	[0.070]	[0.064]	[0.101]	[0.065]	[0.089]
Q-value	(0.759)	(1.000)	(0.759)	(1.000)	(0.759)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Control Mean	12.398	11.644	12.398	11.644	12.398	11.644	0.461	0.636	0.461	0.636	0.461	0.636
Observations	2341	849	1166	343	1175	506	2341	849	1166	343	1175	506

[1] Online appendix tables exclude 8977 individuals who entered study households after the baseline survey. Panel A: the sample excludes free DFS households. Panel B: the sample includes all households in the sales experiment control group or Free DFS experiment. Panel C: the sample includes all households located in the villages where the Free DFS experiment took place. Panel D: the sample includes all households in the study. Panels A-D include all baseline respondents for whom we have the respective outcomes data within the specified samples. Currently using DFS or penultimately used DFS is instrumented by a dummy for sales experiment and a dummy for free DFS households

[2] The first line of each panel reports the coefficient on the treatment variable in a regression with hemoglobin concentration or has anemia as the outcome variable. Standard errors in brackets. * indicates p-value less than 0.05, ** indicates a p-value less than 0.01, and *** indicates a p-value less than 0.001.

[3] Panel A, B & D: regressions include block-level fixed effects. Standard errors are clustered at the village level. Panel A: a control for Free DFS village is also included. Panel C: regressions include village level fixed effects. Standard errors are clustered at the household level. Panel D: currently using DFS or penultimately used DFS is instrumented by a dummy for sales experiment and a dummy for free DFS households.

[4] For the list of controls used, please refer to notes 4 and 5 under Table A2a.