Nutrition and Child Development in Low- and Middle-Income Countries –

Evaluation of Three Micronutrient Interventions

Dissertation in order to acquire the doctoral degree

from the Faculty of Economic Sciences

at the Georg-August-Universität Göttingen

Submitted by

Marion Krämer

born in Leverkusen, Germany

Göttingen, 2017

Supervision and examination committee

First supervisor:	Prof. Dr. Sebastian Vollmer
Second supervisor:	Prof. Dr. Stefan Klonner
Third supervisor:	Prof. Dr. Matin Qaim
Date of defense:	2nd June 2017

"The hidden hunger due to micronutrient deficiency does not produce hunger as we know it. You might not feel it in the belly, but it strikes at the core of your health and vitality."

Kul C. Gautam, former deputy executive director of
United Nations International Children's Emergency Fund
(UNICEF)

Table of Contents

Acknowledgements

I have received great support from so many people and institutions over the years.

I thank Prof. Dr. Sebastian Vollmer for his supervision during the last three and a half years. I am grateful that he gave me the opportunity to be part of the *Development Economics Group* of the University of Göttingen, for the freedom I got to pursue research on topics for which I am truly passionate and the possibility to conduct field research in India. I am thankful for the very valuable comments on this dissertation of my second supervisor Prof. Dr. Stefan Klonner, as well as for his mentorship, guidance and support, which already started during my Master studies in Heidelberg. Stefan has inspired me to work in development economics and on impact evaluation. I am thankful that Prof. Dr. Matin Qaim has been willing to be part of my dissertation committee and for his very helpful comments particularly with respect to essay 2.

I thank my fellow students from the *Development Economics Group* and from the RTG 1666 *GlobalFood*. They have not only been the best possible colleagues I could have wished for, but also have become wonderful friends. I thank Lisa Oberländer, Esther Heesemann, Cara Ebert, Hanna Freudenreich, Karin Jongsma, Rivayani Darmawan and Christian Bommer in particular for scientific discussions and/or for proof reading my manuscript (and so much more). I thank my friends outside University for offering such welcome distractions outside academic everyday life.

There are three persons who have played a special role in supporting me during my dissertation. I am deeply thankful to each one of you. Jana Kuhnt, you have become one of my most valued advisors and motivators and dearest friends. Thank you for your everlasting willingness to discuss even the tiniest details of my thesis and all your helpful scientific advises. I strongly appreciate that you always have had an open ear for me, that you let a kite fly over my head and that you jumped from park benches with me. I cannot wait for all the other exciting and lovely adventures we are going to encounter in the future. Peter Pütz, I am wholeheartly grateful for your endless support, for cooking me dinner from saved vegetables and for simply being there. I very much admired that you took the tough climate and difficult working conditions in Bihar with ease and that you kept synchronizing tablets till midnight. I thank you for your determination and your patience while helping me solve several statistical problems. Thank you for showing me how wonderful data analysis can be and how even more

wonderful finishing work early is. Abhijeet Kumar, I thank you for conducting the data collection in India with me. Your sincere and hard-working personality was very inspiring and impressed me greatly. The collaboration with you was uniquely trustworthy, respectful and productive. *Dhanyavad!*

I owe gratitude to my parents, my grandparents, my sister and her family. Thank you for your trust, understanding, patience and full support. The dissertation is finished and I am very happy to finally celebrate this with you. The assurance that – no matter what – the doors of your homes are always open for me, has giving me much strength over the last years – more than you can possibly imagine.

I thank the *Foundation fiat panis* for providing funding for the data collection in Bihar, as well as the *German Research Foundation* (DFG), which provided funding for the same data collection within the scope of the Research Training Group 1666 *GlobalFood* - Transformation of Global Agri-Food Systems.

List of Abbreviations

ASER	Annual Status of Education Report
ATE	Average treatment effect
BMI	Body mass index
BPL	Below poverty line
CI	Confidence interval
CCT	Calonico, Cattaneo and Titiunik (method for bandwidth selection)
DD	Double-difference/ difference-in-difference
DDP	Diphtheria, pertussis and tetanus
DDS	Dietary diversity score
DHS	Demographic and health surveys
DFS	Double-fortified salt
FAO	Food and Agricultural Organization
GDP	Growth domestic product
GNI	Growth national income
Hb	Hemoglobin
HH	Household
HIV	Human immunodeficiency virus
ICC	Inter-cluster correlation
I$	International Dollar
ID	Identification number
IDD	Iodine deficiency disorders
IGF	Insulin-like growth factor
IGFBP	Insulin-like growth factor binding protein
ITT	Intent-to-treat
MDE	Minimal detectable effect
MDM	Midday Meal
MNREGA	Mahatma Gandhi National Rural Employment Guarantee Act
MRC	South African Medical Research Council
MSE	Mean squared error
N	Number
NFHS	National Family and Health Survey
NIN	National Institute of Nutrition

PDS	Public Distribution System
Rs	Indian Rupees
PSU	Primary sampling unit
PCA	Principal component analysis
RCPM	Raven's Colored Progressive Matrices
RCT	Randomized controlled trial
RDD	Regression discontinuity design
RTK	Rapid test kid
SC/ST	Scheduled castes and tribes
Sd	Standard deviation
SD	Simple difference
SE	Standard error
UK	United Kingdom
UNICEF	United Nation's Child Fund
USD	United States Dollar
WHO	World Health Organization
WISC	Wechsler Intelligence Scale for children

List of Tables

List of Figures

1. General Introduction

This dissertation consists of three essays on the link between nutrition and child development (chapters 2-4). In the general introduction (chapter 1) I discuss the significance of the dissertation topic, define important scientific terms, and highlight the similarities and differences between the essays. In section 1.6 of this introduction I elaborate on a general hypothesis that emerges from the findings of the three essays.

1.1 Linking Nutrition, Poverty and Development (Economics)

With the scientific contributions of Amartya Sen (1999), economists have gone on to define poverty as a multidimensional concept that is not only comprised of insufficient income or consumption (i.e. living on less than 1.9 USD a day) but also to include deprivation in non-monetary dimensions that are considered essential for well-being (Bourguignon and Chakravarty, 2003). According to Sen's influential capability approach, a person is considered poor if he or she is deprived in *basic capability*, i.e. valuable options or freedom in real choice. Only if basic capabilities are ensured, can human beings develop other, higher capabilities. For instance, only a healthy, well-nourished and educated person will be able to pursue the career, she or he has reason to value (Alexander, 2008). A limited capability set, means a limited number of options from which an individual can choose, i.e. that only a limited number of *functionings* (i.e. *doings* and *beings*) can be realized. A functioning is the realization of a particular option from the capability set. For example, one might have the capability to eat a healthy meal because it is available and accessible; however, one might not choose the option to eat it and be well nourished (one functioning) but instead to fast and choose to remain hungry (another functioning) (Sen 1999). According to Drèze & Sen (2013) the expansion of capabilities is *"…what development is ultimately about"* (p. 182) and hence only implicitly the realization of certain functionings. Different researchers have come up with different lists of basic capabilities (e.g. Nussbaum 2011) that, among other things, include the possibility of achieving adequate health, sufficient nutrition and sufficient education.

Nowadays expanding the capabilities of human beings has been established as a means to economic development as well as an end of itself. This is reflected in poverty measures that no longer rely only on income or consumption, but are composed of multiple dimensions such as the human development index (UNDP, 2015) or the multidimensional

poverty index (Alkire and Foster, 2007). Out of the 17 sustainable development goals, only one directly relates to income (Goal 1: By 2030, eradicate extreme poverty measured as living on less than 1.25 USD a day), whereas the other goals focus on health, nutrition, education and environmental conservation among others (United Nations, 2017).

There are different reasons why one might want to ensure basic capabilities. Possibly the most quoted are social justice and humanity. This relates to an understanding of the collective responsibility for protecting disadvantaged groups such as the disabled, weak, elderly and poor. Assuming that being in a state of poverty is based on misfortune and unequal opportunities, a wealthy person with altruistic preferences might feel negatively affected by seeing a person in poverty.[1] Therefore, the utility of the wealthy person increases with the provision of goods and services that reduce poverty, e.g. providing nutritious food, clothing or shelter (Zweifel, Breyer and Kifmann, 2009). Furthermore, poverty alleviation contributes to political stability and peace (Collier and Dollar, 2004). Moreover, some areas of social life, for instance many parts of health care, exhibit externalities. Infectious diseases have negative externalities because they are transmitted from one individual to another and the cost of the person to whom the infection was transmitted is not taken into account when making one's own investment decisions. Similarly, undernutrition weakens the immune system and thus inhibits the resistance against infectious diseases. Being undernourished can therefore exhibit a negative externality. Externalities result in an inefficient allocation of resources and require regulation and governmental intervention (Zweifel, Breyer and Kifmann, 2009). Lastly there is a financial argument why basic capabilities should be ensured. Basic capabilities enable human beings to make use of their potentials, to be productive and to generate income. On a macro-level this might translate to economic growth (Dasgupta & Ray, 1986 and Strauss & Thomas 1998).

In this dissertation, I evaluate three different concrete interventions that aim at expanding the perhaps most basic of all capabilities: *nutrition*.

[1] This perception can be illustrated by using the concept of the 'veil of ignorance', developed by Rawls (1971). According to Rawls, agreeing on the degree of social security benefits should be done in the following way: A society anonymously agrees upon social security benefits without anybody knowing if she will eventually belong to the low or high income group. An agreement would likely include some basic level of social support for all.

1.2 Undernutrition and its Implications for Child Development

Nutrition is defined as

> *"… the process through which living organisms use food to maintain life, growth and normal functioning"* (Kent 2005, p. 7).

Malnutrition emerges either because of an inadequate food intake or because of problems in processing the nutrients from the diet (Kent, 2005). There are three forms of malnutrition: First, overweight and obesity, i.e. when a person consumes too many calories. Second, there are two forms of undernutrition: protein-energy or macronutrient undernutrition, i.e. when a person consumes too few calories, and micronutrient deficiencies, i.e. when a person lacks important vitamins and minerals such as zinc, iron, or iodine (WHO, 2017b). This dissertation focuses on the latter aspect: micronutrient deficiencies. One of the differentiating factor between micronutrient deficiencies and protein-energy malnutrition, where people directly notice a feeling of hunger, see small food portions or look very short and skinny, is that a lack of micronutrients is often not directly visible, unnoticed or involves symptoms that are not attributed to malnutrition. This is why micronutrient deficiencies are commonly called the *hidden hunger* (Kennedy, Nantel and Shetty, 2003).[2] A second key factor of micronutrient deficiency lies in the circumstance that a person needs very few of them, but when they are deficient, the consequences are severe (UNICEF, 2017).

Low-income populations are especially prone to micronutrient deficiencies because they often depend on a single staple food. As such, their diet lacks variety and especially products that are rich in micronutrients, such as animal-sourced products, fruits and vegetable or micronutrient-fortified products (Kennedy et al. 2003).

There are 19 micronutrients known to be essential for proper physical and mental development, for the functioning of a bodies metabolic processes and an intact immune system (Kennedy et al. 2007). It is estimated that more than two billion people, one in three individuals and double the number of individuals that suffer from protein-energy malnutrition, are deficient in at least one micronutrient (FAO et al. 2014, Graham et al. 2007). This dissertation focuses on two micronutrient deficiencies: iodine (essay 1) and iron (essays 2 and 3). Along with zinc and vitamin A, iodine and iron deficiencies are the most prevalent nutritional disorders in the world (Kennedy et al. 2003). Approximately 1.8 billion people suffer from iodine deficiency (von Grebmer et al. 2014, de Benoist et al.

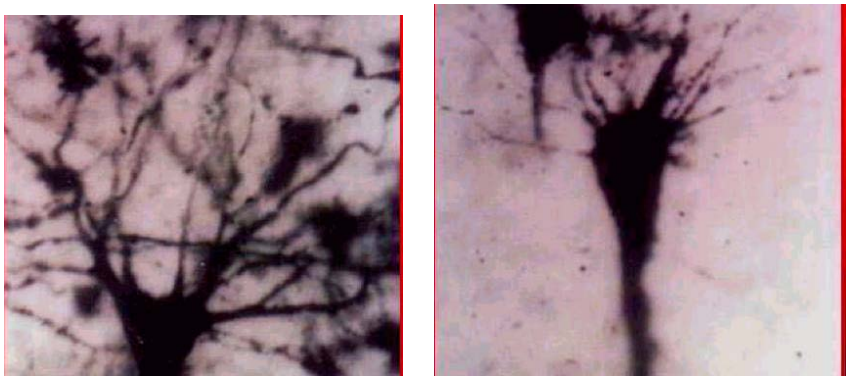
[2] There are few exceptions where micronutrient deficiency becomes directly visible, e.g. blindness from Vitamin A deficiency and goiter from iodine deficiency (von Grebmer *et al.*, 2014).

2008). With 1.6 billion worldwide, the number of people suffering from iron deficiency is similarly high (von Grebmer *et al.*, 2014).

Micronutrient deficiencies can lead to dramatic impairments in *child development*. Child development is defined as the physical, motoric, cognitive, social-emotional and linguistic changes that occur from gestation until adolescence (Bartolotta and Shulman, 2010). Child development is influenced both by genetic-controlled processes, i.e. maturation (also referred to as *nature*) as well as by the impact of the environment (also referred to as *nurture*). Nutrition is counted as an environmental factor (Berk, 2012). Iron and iodine deficiencies both adversely influence cognitive development. Figure 1.1 shows how malnutrition can impair the development of neurons in the human brain.

Figure 1.1: Nutrition and brain development

Normal Malnourished



Source: Kiess, no date

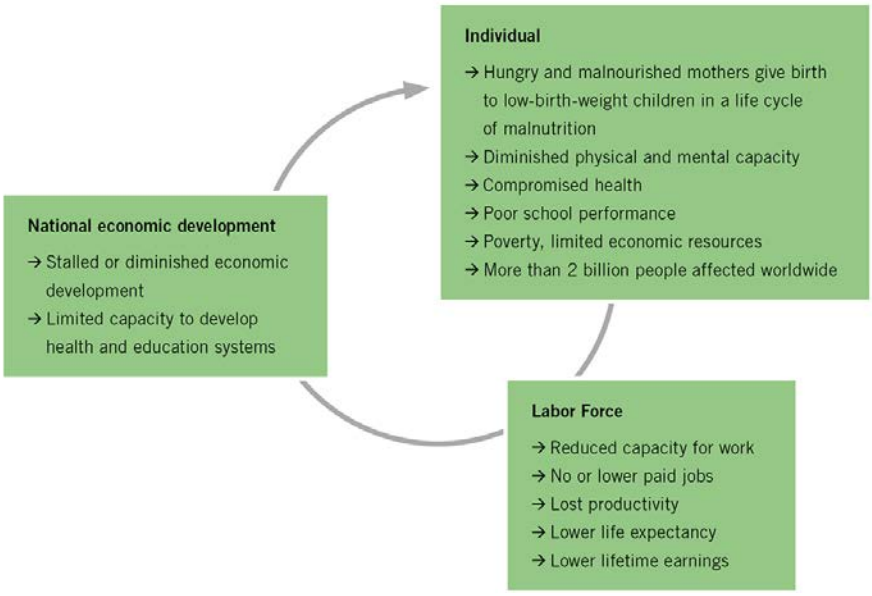
Iodine is also known to affect physical growth.[3] Indirectly, they might also affect other aspects of child development, e.g. motor and language skills, because these other aspects of child development strongly depend on a child's cognitive development (Lozoff, 2007). Iodine is required for the synthesis of different hormones, that among other things, are necessary for skeletal growth and neurological development (Dunn 1992, Samuels et al. 1989). Iron deficiency affects cognitive development through immediate neurobiological processes, i.e. the inhibition of the central nervous system to develop properly (e.g. the brain and the spinal cord) (Beard 2003), and secondly through functional isolation. Functional

[3] The evidence of iron on child growth is inconclusive (Ramakrishnan *et al.*, 2004).

isolation emerges from the symptoms of iron deficiency. Children deficient in iron engage less with their environment, have lower interpersonal interactions, show lower attention and are relatively unresponsive to stimuli in comparison to their non-iron deficient counterparts. Hence, they have difficulties in accumulating new skills (Lozoff et al. 1998). Cognitive development directly influences the education outcomes of a child. For example, if a student is unable to focus their attention and to ignore distraction, they are likely to have trouble concentrating and hence have difficulties in acquiring new skills and knowledge. The same is likely to be true for the symptoms of anemia such as frequent illness or tiredness. Both iron and iodine deficiency therefore hinder human capital formation (Halterman et al. 2001, Bobonis et al. 2006). Sufficient micronutrient intake is therefore essential for a person to reach their physical and cognitive potential, i.e. to expand the set of capabilities.

When impairments in child development, due to malnutrition, lead to lower educational attainment this might result in lower adult productivity and income poverty, which in turn might lead to undernutrition of the individual as well as their offspring. Therefore, the *micronutrient poverty trap* enforces the intergenerational transmission of poverty. Through reduced productivity and high health care costs, micronutrient deficiency constrains economic development at the national level as well (Dasgupta and Ray, 1986 and Strauss & Thomas 1998). Stein and Qaim (2007) estimate that the short-term economic costs of micronutrient deficiencies in India amount to between 0.8% and 2.5% of the gross domestic product. Figure 1.2 illustrates the vicious and continuous cycle of hidden hunger and impaired economic development.

Figure 1.2: Cycle of hidden hunger, poverty and stalled development



Source: von Grebmer et al. (2014)

While the biological basis for the positive effects of increased micronutrient intake is clear and well established, the crucial question is in what way and by which means an adequate micronutrient intake of the population in need can be ensured most effectively.

1.3 Interventions to Fight Micronutrient Deficiencies

There are different approaches to cure and prevent micronutrient deficiencies. They can be broadly divided into nutritional supplements and food-based approaches. Supplements refers to the administration of capsules, tablets or injections that contain high concentrations of minerals or vitamins (FAO and WHO, 2002). Food-based approaches are comprised of dietary diversification and fortification (including bio-fortification). Dietary diversification refers to a change in production and consumption patterns towards a broader range of micronutrient-rich foods (Allen *et al.*, 2006). Bio-fortification is the selective breeding or genetic modification of crops with the aim of increasing their nutrition value (Global Panel, 2015). Food fortification is the direct addition of micronutrients to processed foods such as salt, rice or wheat (FAO and WHO, 2002). Food-based approaches do not end with an increase in the availability of iron-rich or iron-enriched products, but can only be successful if

accessibility (i.e. monetary feasibility) is ensured, and a change in feeding practices, i.e. utilization, follows (Allen *et al.*, 2006).

Nutrition interventions – just like interventions in other fields – can be further categorized into supply- and demand-side interventions. While supply-side interventions generally increase availability and often also the accessibility of certain products or technologies, demand-side interventions aim to increase the knowledge and awareness about the benefits of a certain product or technology that might lead to an increase in demand and therefore a behavioral change in feeding practices (World Bank, 2006).

The Copenhagen Consensus Expert Panel continuously ranked nutrition interventions, including vitamin C and zinc supplementation and iron and iodine fortification, among the best interventions for economic development in terms of their cost-benefit ratio (Copenhagen Consensus 2008, Copenhagen Consensus 2012). From the neoclassical point of view, a rational agent,[4] would be expected to adopt and implement these technologies due to their high benefits and low costs.

In this dissertation, I evaluate three different food-based approaches that aim to increase children's consumption of iodine and iron:

(1) Essay 1: The usage of iodized salt at the household-level

(2) Essay 2: The usage of iron-fortified iodized salt in a school-feeding program

(3) Essay 3: Informing parents about the anemia status of their child and the provision of short and simple nutritional advice

Table 1.1 outlines the similarities and differences between the three interventions at a glance.

The possibility to attain a higher consumption level of iron and iodine in children will enlarge their nutritional capability. First, I test whether the enlargement of the nutritional capability indeed results in a state of better nutrition or *functionings* in terms of Sen's capability approach (e.g. anemia, dietary diversity score, child growth). Secondly, since being

[4] Early neoclassical models define rationality in terms of utility maximization. Individuals maximize their utility given their preferences and the constraints they are facing. It is generally assumed that individual utility maximization is exclusively based on self-interest. A rational agent performs an action, e.g. makes an investment or buys a certain good, if the marginal utility of the action is higher than the marginal costs. An individual behaving according to this pattern is called *homo oeconomicus* (Mankiw and Taylor, 2011).

in a state of good nutrition expands other capabilities, I further look at how these interventions affect the educational and cognitive aspects of child development in essays 2 and 3 (higher capabilities).

Table 1.1: Overview and comparison of the three essays

	Title	Research question	Outcomes	Intervention					Data/ Sample	Age group	Method	Quality of evidence	Result
				Nature	Potential Actors	Supply or demand side	Adoption private or public decision	Category					
1st Paper	Association between household unavailability of iodized salt and child growth – Evidence from 89 demographic and health surveys	Is there a general association between household unavailability of iodized salt and child growth across countries?	Stunting, wasting, underweight, low-birth weight	Iodized salt used in households	Public, private, NGOs	Supply	Possibly private, public if mandatory fortification policy	Fortified product	89 Demographic and Health Surveys from 46 low- and middle-income countries	Newborns and children until the age of five	Logistic regression with control variables and primary sampling unit fixed effects	Association	- Universal positive association with low birth weight - Child growth indicators are only statistically significant associated for the sample of India
2nd Paper	School feeding, iron-fortified salt and child cognitive ability – Evidence from a randomized controlled trial in rural India	Does the usage of iron-fortified iodized salt in the Indian school-feeding program reduce anemia and improve cognitive and educational outcomes?	Hemoglobin, anemia, cognitive ability and education outcomes	Iron-fortified iodized salt used in school-feeding	Rather public, maybe NGOs	Supply	Public	Fortified product	Self-collected dataset of about 2000 school-aged children and their households from two blocks in rural Bihar, India	Primary school children	Randomized-controlled trial	Intent-to-treat effect	- Decrease in any form of anemia and mild anemia - No effect on cognition - Weak evidence for a small increase in math and reading skills for high rates of school attendance
3rd Paper	Nutrition information, anemia testing and feeding practices – A regression discontinuity analysis using data from rural India	Does informing the parent about the anemia status of their child and giving them short and simple nutrition advices improve feeding practices, reduce anemia and improve cognitive and educational outcomes?	Feeding practices, hemoglobin, cognitive ability and education outcomes	Information about anemia status and better feeding practices	Public, private, NGOs	Demand	Private	Dietary diversification	Self-collected dataset of about 2000 school-aged children and their households from two blocks in rural Bihar, India	Primary school children	Regression discontinuity design	Marginal treatment effect	- No robust impact on any of the tested outcomes

1.4 The Role of Evaluation

This chapter explains why it is important to identify the true causal impact of these interventions, the challenges in identification of the causal impact and how the identification problem can be solved.

1.4.1 Why Evaluate?

A great deal of money has been and is continuously being spent by local governments, private institutions and in the form of development aid on interventions that aim at reducing poverty in its multiple dimensions, including nutrition. In a world of scarce resources, different interventions must be traded off.[5] Rigorous impact evaluation enables one to *causally* attribute a change in an outcome to a specific intervention. Evaluation answers the question if an intervention is effective at all; and secondly, by which magnitude it has changed the outcome of interest. Costs and benefits or effects of different interventions can be compared to provide policy makers with empirical evidence to decide on one or the other intervention. Evaluation also enables policy makers to be held accountable and provide transparency in resource allocation (Khandker, Koolwal and Samad, 2010).

1.4.2 The Evaluation Problem

To answer the question if and in how far these three interventions causally affected child health, cognitive and educational outcomes, one must answer the counterfactual question: How would these children have done without being exposed to these interventions? How would the growth of one of these children have been different had they not consumed iodized salt at home, or their anemia rate had they not have received iron-fortified salt through their school meal? How would these parents have fed their child had they not been informed that their child suffers from anemia and should be fed differently? However, we are never able to observe the same individual with and without being exposed to the treatment at one point in time. This fact constitutes the evaluation problem (Angrist and Pischke 2008, Duflo, Glennerster, and Kremer 2007).

[5] A famous debate on aid effectiveness was held by Jeffrey Sachs and William Easterly. While Sachs argued that with sufficient money poverty could be eradicated via a big push (Sachs, 2005), Easterly argued that the 2.3 trillion dollars that have been spent on development aid have not been successful in ending poverty, doubting that the next 2.3 trillion dollars could do so (Easterly, 2006). The motivation of impact evaluation can be positioned between these two poles, as impact evaluations assess each potential aid project individually.

One intuitive first approach to create the counterfactual outcomes would be to simply compare how children performed before and after the intervention. In this approach, outcomes before the intervention would then constitute the counterfactual. However, this approach would lead to biased results in most cases because factors other than the treatment itself might have caused a change in the outcome. Hence, in this approach, the treatment effect cannot be disentangled from the effect of other factors. Indeed, it is never possible to estimate the true treatment effect for a given individual. It is however possible to estimate the *average* treatment effect for a *group* of people compared to another *group* of people that did not receive a treatment. The crucial point in evaluation is to create two groups that would have had the same outcome in the absence of the treatment. Unfortunately, when looking at real world policies, it turns out that the treated and not-treated groups created are mostly very different from one another. Real world policies or programs are generally targeted to special groups or attract people with particular characteristics. For instance, households that buy fortified salt possibly have higher literacy and education levels, which enables them to understand the benefits of the fortified product compared to individuals that do not buy fortified products. We never know if the difference in outcomes of interest, e.g. the health and nutrition status, is indeed due to the fortified product, or due to the fact that the outcomes of interest for those individuals that buy the fortified product are anyways better because their better education is also an underlying cause of their more healthy diet. The systematic difference between those who received the treatment or self-selected into it is called the *selection bias*. Using the potential outcomes framework, developed by Rubin (1974), the evaluation problem is formalized as follows.

Let Y_i^T be the average observed outcome of an individual i that was exposed to an intervention T and Y_i^C the average observed outcome of an individual i that was not exposed to an intervention. As described above, we can never observe Y_i^T and Y_i^C at the same time, i.e. we can never estimate the individual treatment effect Y_i^T - Y_i^C. We can however observe these two outcomes as averages within a population.

$$D = E[Y_i^T|T] - E[Y_i^C|C]$$

Subtracting and adding the expected outcome for an individual that was exposed to the intervention had they not been exposed (unobservable), we can rearrange the equation as:

$$D = E[YiT|T] - E[YiC|T] + E[YiC|T] - E[YiC|C]$$

Where

E[YiT |T] - E[YiC|T]

= the causal effect of the treatment, i.e. what we are interested in and

E[YiC|T] − E[YiC|C]

= the selection bias.

The oftentimes called gold standard to overcome selection bias is randomization. If households were randomly allocated into one group that uses the fortified product and another one that does not (the control group), and the number of observations was large enough it would ensure that, on average, all observable and unobservable characteristics between the treatment and control groups are the same, such that the selection bias is:[6]

$$E[YiC |T] − E[YiC |C] = 0$$

and the true causal effect is identified.

Mostly for financial, ethical or practical reasons, randomly allocating individuals to treatment and control groups is not always feasible. That is why economists oftentimes rely on quasi-experimental methods. In quasi-experimental methods, researchers try to identify the causal relationship between two variables without controlling the assignment of an entity to the control or treatment groups themselves, but use groups that have already been created *naturally*. Quasi-experimental methods consist of identifying situations in which the selection bias is either not present, or where methods can be applied such that one can correct for it (Duflo, Glennerster and Kremer, 2007). Quasi-experimental approaches differ in the way the control group (i.e. the counterfactual) is created. Whether the created control group is internally valid depends on the plausibility of a set of *identifying assumptions* (Angrist and Pischke 2008; Duflo, Glennerster and Kremer 2007).

1.4.3 Solving for Selection Bias in the Three Essays

The three essays of this dissertation use the following control groups and are based on the following identifying assumptions:

(1) Essay 1: Controlling for observable (and some unobservable) characteristics in a parametric regression framework. This method assumes that even though the treatment and

[6] The balance in means of the characteristics of the treatment and control group is only true, on average, meaning if a random sample of entities would be drawn many times. It is still likely that due to random chance, and especially if the sample size is small, some means show up to be different between the control and treatment groups (Altman, 1985; Imai, King and Stuart, 2008).

control groups are not randomly allocated, controlling for a set of factors X that influence the treatment status and the outcome of interest at the same time eliminates the selection bias, or put differently, conditional on factors X there would have not been a difference in outcomes in the absence of the treatment.

$$E[YiC \mid X, T] - E[YiC \mid X, C] = 0$$

This however implies that any factor simultaneously influencing the treatment status and outcomes has to be incorporated to overcome the selection bias and any omitted variable will produce biased results (in this context, selection bias is also known as *omitted variable bias*). Furthermore, as we apply a linear regression model, we make the assumption of a linear relationship between the treatment variable, covariates and outcomes. The assumptions of this evaluation method are relatively strong. Since we cannot rule out the possibility of the presence of some degree of selection bias, we prefer using the terms *association* and *correlation* instead of effect in this analysis.

(2) Essay 2: Using randomization (in a randomized controlled trial – RCT), where the randomization of schools into treatment and control groups prevents the presence of a selection bias and hence the true causal effect is identified with the minimal assumption that the randomization process was successful (and a few more minimal conditions that are discussed in chapter 3.6). Hence by construction

$$E[YiC \mid T] - E[YiC \mid C] = 0$$

(3) Essay 3: Randomization at a threshold X' that is the result of the data generating process, where assignment to the treatment is a discontinuous function of a variable X (Regression discontinuity design – RDD).

$$E[YiC \mid T, X < X' + h, X > X' - h\ ] - E[YiC \mid C, X < X' + h, X > X' - h] = 0$$

The assumption implies that given that all observable and unobservable characteristics evolved smoothly with the variable X, conditional on X, the selection bias is 0 around a narrow bandwidth of h from the threshold X'. Individuals just above (below) the threshold are used as the control group for individuals just below (above) the threshold.

1.5 Summaries of the Three Essays

1.5.1 Essay 1

In essay 1, we investigated if there exists a general association between a household's unavailability of iodized salt and child growth across countries. That means we basically test if the expansion of the nutritional capability of a household through the availability of iodized salt indeed results in child growth, which is an indicator for nutrition and health status (i.e. the realization of certain functionings). We used 89 nationally representative, repeated cross-sectional and mutually comparable demographic and health surveys (DHS), conducted between 1994 and 2012 across 46 low- and middle-income countries. We analyzed the data for the outcome variables stunting (low height-for-age), underweight (low weight-for-age), wasting (low weight-for-height) and low birth weight in children between 0 and 59 months at the time of the interview, using logistic regression models. The unavailability of iodized salt at the households was tested using a rapid chemical test. Our samples consisted of 390,328 children for the stunting analysis, 397,080 for the underweight analysis, 384,163 for the wasting analysis and 187,744 for the low birth weight analysis. Models were adjusted for individual, maternal and household covariates and fixed effects on the level of the primary sampling unit (PSU). In the fully adjusted models (including all covariates and PSU fixed effects), the unavailability of iodized salt was associated with a 3 % higher odds of being stunted, a 5 % higher odds of being underweight, and a 9 % higher odds of low birth weight. When excluding India from the sample, the association was only statistically significant for low birth weight. Though we do not establish causality in our analysis, the findings might indicate that the causal effect of iodized salt on child growth, if it exists, is most profound in utero, and is not universally effective across all countries with respect to longer run child growth outcomes such as stunting and underweight. The findings are very much in line with the previous literature that found mixed results of increased iodine consumption on stunting, wasting and underweight, but mostly positive associations with low birth weight. These findings support the general notion in the health and nutrition literature that emphasizes that pregnancy is a crucial period for child development.

1.5.2 Essay 2

In essay 2, we analyzed if the Indian school-feeding program is an effective channel to deliver iron and reduce the high-risk of iron deficiency that is common among children in rural India. We test the effectiveness of such an intervention by running a randomized controlled trial in which 54 randomly selected government-funded schools, from two blocks in Bihar, India, used iron-fortified iodized salt to prepare the school lunch for a period of one year. Fifty-three randomly selected schools functioned as control schools and continued using conventional iodized salt. In addition to anemia, we investigated the effect on cognitive and educational outcomes. Framed in the capability framework, we expand the capability set through the provision of iron-fortified iodized salt to government-funded schools and test if this results in the realization of a state of better health and nutrition (of which anemia constitutes the indicator). Since a state of good health and nutrition (being non-anemic) is the foundation for a good education and cognitive skills (i.e. health and nutrition enlarge the educational and cognitive capability set), we also look at these additional outcomes.

The treatment reduced the prevalence of any form of anemia by 20% and of mild anemia by 30%. There is no statistically significant interaction effect between treatment and school attendance, i.e. the treatment effect does not change significantly with increasing compliance. However, despite the short treatment period, there is weak evidence that at a school attendance rate of 80% or 90%, the treatment had a positive effect on reading and math scores. We further find that the intervention is cost-effective. These results provide encouraging evidence of the positive effects from using school-feeding programs as a channel to provide school-aged children with iron-fortified salt.

1.5.3 Essay 3

In essay 3, we studied if revealing the anemia status of a child and informing the child's parents about the need to feed the child more iron-rich food items is effective in changing their feeding practices, the child's hemoglobin level and subsequently cognitive and education outcomes. This research is motivated by the circumstance that many low-cost technologies to improve health and nutrition currently exist, but the adoption of these technologies remains low or lower than expected (Dupas 2011, Banerjee & Duflo 2012). This has led researchers to the hypothesis that the supply of or access to health and nutrition technologies, i.e. availability and monetary feasibility, might not be the only constraints poor households in low-income

countries are facing. There is a growing strand of literature that investigates how far the lack of information constitutes an additional constraint, which limits the demand for and the proper use of these technologies (Dupas 2011b, Karlan et al. 2014). The theory being that individuals would make better health and nutritional investments if the information constraint was loosened by providing them with the required information.

Using a dataset from rural India we exploit the discontinuity in the provision of information that resulted from the ethical need to inform parents about the anemia status of their child in a randomized trial, if their child's hemoglobin level was below a clinical threshold. This circumstance allows us to apply a regression discontinuity design, which has the potential to identify the causal effect of the intervention. We do not find any robust treatment effect on any of the tested outcomes. Information alone, even when combined with revealing the anemia status of a child, does not seem to be effective in changing nutritional behavior. This might indicate that other forces constrain people in making rational nutrition investments. We speculate that the unavailability and inaccessibility of iron-rich food items, as well as other social and psychological factors, might be the underlying causes for why we did not find robust effects. A drawback of this study is that the given dataset only provides the statistical power to detect relatively large effect sizes.

1.6 General Conclusion

What are the general implications that can be drawn from these three papers? What can these three evaluations teach us about the life of poor people and the effectiveness of nutrition interventions to improve their well-being? In how far are these three interventions in fact able to expand the capability set of poor people? Starting with essays 2 and 3: Assuming that too little power is not the underlying reason for the non-detection of an effect in essay 3, why is the school-feeding intervention effective and why did the nutrition information intervention not show any statistically significant effects? One important condition that differentiates these two interventions is how much freedom of choice people have in response to these interventions. The usage of iron-fortified salt in the school feeding program was mandatory for the treatment schools. Parents and children were not given any choice to take-up the treatment or not. By going to school and eating the lunch, children received the iron-fortified salt without any further decision or choice to make. This is comparable to a mandatory food

policy.[7] In contrast, in the nutrition information intervention in essay 3, it was fully up to the parents to decide how to act on the information they received. Does this mean that if poor people are given freedom to choose they make bad, i.e. irrational, decisions? These results are from only two evaluations from one particular area in rural India and essay number 3 might be suffering from too little power; therefore, the generalizability of these findings should not be overstretched. Still, the results of this thesis are very much in line with the latest literature on the conditions that influence the human decision making process and the particular influence that a life in poverty has on it (see The World Bank 2015, Karlan et al. 2014 and Dupas 2011 for overviews). This thesis adds pieces of evidence to the potential presence of these conditions.

One of these conditions is missing or imperfect markets. Even standard economic theory recognizes that individuals might be prevented from making rational choices when markets are missing, imperfect or when individuals are not well informed. Essay 3 presents an intervention that aimed at loosening the information constraint. However, missing markets might still constrain individuals from changing nutritional behavior, even if they intend to do so. The financial means (credit market) and the goods and services (commodity market) required for this behavioral change might not be available or accessible. If the diversity of available food items is low, or certain food items are too expensive, people are not capable of changing their diet. Imperfect and missing markets are particularly prevalent in rural areas in low-income countries (e.g. Ray 1999, Mankiw and Taylor 2011).

More recently, insights from sociology and psychology have entered economics and researchers have started to depart from the neoclassical perception of a human being as a self-controlled, fully rational decision making *homo oeconomicus*, who, as long as she or he is well informed and markets are perfect, makes decisions that maximize their well-being. Behavioral economists have promoted the hypothesis that social norms, beliefs and certain cognitive tendencies often prevent human beings from rational decision making even when markets are perfect and they are fully informed (e.g. Sunstein and Thaler 2009, Tomer et al. 2013). The *World Development Report 2015: Mind, Society and Behavior* has emphasized the idea that the insides of sociology and psychology should be used in development as its theme (The World Bank, 2015). In recent years, increasingly more empirical research has combined sociology, psychology and development economics to understand why individuals do not

[7] Only by not sending their child to school, but since they did not know about the intervention, this is not of any concern in this context.

always make choices that improve their well-being. What are the central insights from this interdisciplinary research?

Beliefs, social norms, habits and culture play a crucial role when individuals make decisions about food intake (Fox 2003, Ma 2015, Wood and Neal 2009). In case new information contrasts with deep-seated beliefs and habits, individuals will hardly comply with new information. For instance in Maharashtra, a west Indian state, fever is treated with rice, whereas in West Bengal, an east Indian state, renouncing rice when having fever is a deeply-rooted belief (Childs et al. 1997, Banerjee & Duflo 2012).

Furthermore, individuals have the mental tendency to strongly discount the future. They give stronger weight to rewards and discomforts that are in the present. This so called *present bias* results in the procrastination of tasks that include discomfort or costs on the one hand, and the disability to resist temptations that cause short-term rewards on the other hand (Laibson 1997, O'Donoghue and Rabin 1999, Banerjee and Mullainathan 2010, Dupas 2011, Karlan et al. 2006). Present bias prevents human beings from putting their intentions into actions. With respect to the nutrition intervention evaluated in essay 3, present bias might prevent parents from spending more money on a diverse diet today, even though they have the intention to do so, because the temptation to invest the additional money into tobacco, sweets, festivals or religious rituals, is stronger. This psychological phenomenon applies to the poor as well to the rich.[8]

Moreover, human beings underlie a *limited degree of attention*. Given the wealth of information a human being is exposed to, they can hardly take into account all dimensions and options in the decision-making process. While this circumstance is true for every human being, according to Mullainathan & Sharif (2013), it is particularly true for the poor. Due to the fact that having very few of the basic requirements for living, e.g. money, food, health and shelter, poor people allocate all of their attention and energy to those things that are scarce in their life to make the most of them. As a consequence, almost no or only very limited mental space is left for other important decisions or certain skills such as self-control or long-term planning. If you have little money and you are unsure if you will be able to buy sufficient calories, all your attention will be dedicated to getting more calories and you will not have the

[8] Also in high-income countries people struggle to eat healthy diets and to be physically active regularly, even though they know that they are putting themselves on a much higher risk for diabetes and strokes; however, the present bias seems to prevent them from putting their intention of physical exercise and more healthy diets into practice.

mental capacity to maximize dietary diversity (Mullainathan and Sharif, 2013).[9] The high potential for irrational decision making of the poor is not only based on scarcity alone, but also on the circumstance that poor people need to make many more decisions in their daily life than rich people. Public and private institutions oftentimes do not provide them with the services that make life easier. For example, they often do not have safe and clean piped water at their home, but must make the decision to add chlorine to purify their water each and every time they collect water from the well. They are generally not covered by social security and must decide whether to spend what little money they have either directly or save it for shocks or pension. An additional hypothesis explaining why it often seems that poor people make irrational decisions is that it is not because they are not clever or uninformed, but because the number of decisions to be taken every day would exceed the mental capacity of any human being (Banerjee & Duflo 2012).

It was not within the scope of this thesis to analyze the true underlying causes of human decision making in the unique setting of each essay. However, when combining the results of the nutrition information from essay 3 and the school-feeding program intervention from essay 2, with the existing literature on human decision making, it gives a good indication that in addition to a lack of information and missing markets, social and psychological forces play a role in decision-making. I do not want to neglect the importance of market access and information as these are two necessary conditions. However, they are possibly not sufficient conditions, particularly when it comes to nutrition and health behavior. If this is true, the obvious next question that needs to be asked is what this implies for policy making? How should policies be designed to help people to make rational health and nutrition investments?

Indeed, the insights from psychology and sociology question the conviction of many societies, governments and schools of thought that emphasize the self-responsibility of citizens, the sovereignty of the consumer and the trust that informing people is sufficient to give them a basis to make choices that enhance their well-being. The findings from this thesis combined with the psychological and sociological knowledge on human behavior questions the statement of the Indian government that lifted the ban on the production of non-iodized salt in 2000 with the reasoning "*matters of public health should be left to informed choice and*

[9] The list of phenomena from sociology and psychology that influence the understanding of economic decision making is much longer and for instance also includes loss-aversion and over-optimism (Karlan, Ratan and Zinman, 2014). I limit the outline to social norms, present bias and limited attention since I consider these to be the most relevant, especially in the context of dietary decisions.

not enforced" (FAO 2006, p. 136). It might also object the emphasis on capabilities in the capability approach, i.e. the idea that development is only about expanding the set of choices, while leaving the decision of which of the different options is chosen, i.e. which functioning is realized, to the individual. Instead, it gives much more responsibility to governments and society as a whole, including different players in international development, to take over some of the decisions that people in low-income contexts have to make, such that certain functionings can be realized with certainty.

To overcome social and psychological barriers in rational decision-making, different tools, such as *commitment opportunities*, *reminders*, *nudging* (i.e. small incentives) and *enforced mandatory policies*, have been developed and empirically tested. Dupas and Robinson (2013) test different commitment opportunities to increase savings for preventive health care and find that simply earmarking funds by providing individuals with a safe box and a key already helps them improve their self-control and overcome present bias. Another well-established rather than newly developed commitment opportunity consists of rotating savings and credit organizations (ROSCAs)[10], where people voluntarily have come up with a organized form of savings commitment (Gugerty, 2007).

Regarding the limited degree of attention, reminders have been tested. In a randomized experiment, Karlan et al. (2013) sent messages to bank clients reminding them of their self-set savings goal and found large increases in savings.

Nudging has been tested for instance in the context of immunization. Setting up immunization camps in one Indian district, Banerjee et al. (2010) find that giving parents one kilogram of lentils for each completed immunization and a set of metal meal plates when having completed all five required immunizations, nearly doubles the full immunization rate. Similarly, Thornton (2005) finds that giving individuals financial incentives to learn their HIV status after having been tested at their household, doubles the take-up of HIV test results. The effectiveness of conditional-cash transfers (Schultz, 2004) is possibly also, at least to some extent, based on the nudging component.

If properly implemented and enforced, mandatory policies might also be the most effective, as they not only ensure that the take-up of certain health or nutrition products is

[10] In ROSCAs, a fixed group of people meets regularly. They commit themselves to contribute a certain amount of money at every meeting into a joint pot. The pot is given to another person from the group at every meeting. Commitment is, however, not the only reason for the existence of ROSCAs; others are insurance and intra-household conflict over savings (Gugerty, 2007).

increased (as in the case of nudging, incentives and reminders), but that these products are universally adopted, as no choice is left to the individual. Mandatory policies can target the population of a whole country or specific groups within a country. The further advantage of mandatory policies is that they provide more integrated solutions, which is particularly important when markets are imperfect. Mandatory policies not only ensure that people put their intentions into actions, but also ensures sufficient supply and financial accessibility (Dupas, 2011b). Mandatory policies are widely spread in low as well as in high-income countries. For example, in Germany, health and unemployment insurance are mandatory and are generally directly deducted from a person's salary. In the United States, children are only enrolled in school when their parents can prove their full immunization status (Ciolli, 2008). Essay 2, and to some extent essay 1, of this dissertation add pieces of evidence to the effectiveness of mandatory policies. The intervention evaluated in essay 2 made the usage of iron-fortified iodized salt mandatory in a selected number of government-funded schools and found meaningful reductions in anemia rates due to this intervention. This finding thus supports the effectiveness of mandatory government food policies. The usage of iodized salt at the household level, as it has been studied in essay 1 of this dissertation, can only be partly attributed to mandatory policies. In the cross-country analysis, we cannot disentangle the reason why households possess iodized salt. It might have been an individual choice at the point of purchase or a national mandatory salt fortification policy.[11] In essay number 1, we find a small universal positive association of iodized salt with low birth weight across countries and with stunting and underweight when we look only at India. This is very encouraging evidence with respect to the potential of a mandatory iodized salt policy, or of nudges or reminders combined with the availability of the technology.

Intervening into the lives of people via the outlined tools and particularly through mandatory policies might sound paternalistic at first and initially I was hesitant to make this concluding statement. However, over and above the market failure argument that legitimates governmental intervention based on standard economic theory, it becomes clear that psychological and social tendencies, such as social norms, present bias and limited attention prevent people from putting their intentions into actions. The presented tools would therefore help people to make the choices they indeed strive for. While this argument for limiting free choice applies to all human beings to some extent, there is another argument that is particular

[11] In all countries in the sample, at least a small portion of the population does not possess iodized salt, such that the presence of universal mandatory salt fortification is implausible. However, the reason for the unavailability of iodized salt can also be that iodine dissipates if not properly stored.

to low-income settings and that is borrowed from Banerjee & Duflo's (2012) book *Poor Economics*. Banerjee & Duflo (2012) state that taking over decisions from poor people is indeed not paternalistic or at least not more paternalistic than the behavior of governments and society in high-income countries, with the small difference that citizens in high-income countries do not realize it anymore. As already stated, in high-income countries safe drinking water enters households via piped water. The decision between chlorinating it or not does not emerge in the first place. Social security contributions are directly deducted from a person's salary and there is no decision making process required to choose between spending that money on temptation goods or saving for shocks. Incentives, nudging and enforced mandatory policies surround people from high-income countries all the time and those means free their minds for the rest of their daily decisions. The crucial question to ask is *which* decisions should be taken over by government (or other institutions) and *which* should be left to the individual. This question brings us back to the first section of this introduction. As outlined in section 1.1, most societies see a social responsibility for ensuring that basic capabilities are realized. Being properly nourished forms one of the most basic capabilities, if not the most central. With respect to the capability approach, the reasoning outlined in this section would therefore imply that the requirement to ensure that poor people can reach a stage of good nutrition and health is not the expansion of the capability set, i.e. giving people more options, but rather taking away some of their freedom of choice to directly ensure that certain basic *functionings* of being in a stage of good nutrition and health are reached.

Summarizing the potential implications from these three essays into one sentence: Taking over some of the (nutritional) decisions poor people have to make every day – e.g. taking away some freedoms through mandatory policies, such as mandating the use of fortified products in school-feeding programs – will eventually enable them to widen their sets of several higher capabilities.

2. Association between Household Unavailability of Iodized Salt and Child Growth – Evidence from 89 Demographic and Health Surveys

WITH ROLAND KUPKA, S.V. SUBRAMANIAN AND SEBASTIAN VOLLMER

This article has been published in the American Journal of Clinical Nutrition, 2016, 104(4), pp. 1093-1100.

Abstract

Although a strong biological basis exists for a role of iodine in somatic growth failure in childhood, this relation has not been previously studied on a large scale to our knowledge. We investigated if a general association exists between the household unavailability of iodized salt and child growth across countries. We used 89 nationally representative, repeated, cross-sectional and mutually comparable demographic and health surveys (DHS), conducted between 1994 and 2012 across 46 low- and middle-income countries. We analyzed the data for the outcome variables stunting (low height-for-age), underweight (low weight-for-age), wasting (low weight-for-height), and low birth weight in children aged between 0 and 59 months at the time of the interview, using logistic regression models. Our samples consisted of 390,328 children for the stunting analysis, 397,080 for the underweight analysis, 384,163 for the wasting analysis and 187,744 for the low birth weight analysis. Models were adjusted for individual, maternal, and household covariates and fixed effects on the level of the primary sampling unit. In the fully adjusted models unavailability of iodized salt was associated with 3% higher odds of being stunted (95% CI of ORs: 1.00, 1.06; P-value: 0.04), 5% higher odds of being underweight (95% CI: 1.02, 1.09; P-value < 0.01), and 9% higher odds of low birth weight (95% CI: 1.02, 1.17; P-value: 0.01). When India was excluded from the sample, the association was only statistically significant (95% CI: 1.00, 1.17; P-value: 0.05) for low birth weight. Although we did not establish

causality in our analysis, the findings might indicate that the causal effect of iodized salt on child growth, if it exists, is most profound in utero and is not universally effective across all countries with respect to longer-run child-growth outcomes such as stunting and underweight.

2.1 Introduction

The trace mineral iodine is an essential component of thyroid hormones. Iodine arises naturally in soil, water, plants, and animals. Human beings absorb iodine through regular dietary intake. Iodine concentrations in water and soil vary widely across the globe. Mountain regions, delta, and flood zones are naturally prone to low iodine concentration (Semba and Delange, 2008). Insufficient iodine nutritional intake and related deficiencies in thyroid hormones result in a cluster of adverse effects termed iodine deficiency disorders (IDD). These include adverse fetal outcomes (abortion, stillbirth, congenital anomalies, and perinatal mortality), goiter, impaired mental function and growth among children (Zimmermann, 2009). IDDs are major public health problems across the globe. An estimated 2 billion people worldwide are at risk of IDD, among which 266 million are school-aged children (de Benoist *et al.*, 2008).

Since 1994, the iodization of salt for human and animal consumption is recognized as the best method to effectively and inexpensively prevent IDD in most settings (UNICEF, 2008). Thus far, the advocacy efforts in favor of salt iodization programs have focused on improvements in intelligence quotients and goiter prevention as the main expected programmatic benefits (Qian *et al.*, 2005; de Benoist *et al.*, 2008). However, there is also a strong biologic basis for a role of iodine in child growth. Growth failure is the primary manifestation of undernutrition during gestation and in childhood and is identified by a height or weight that is too low compared to a normal nourished reference population (stunting and underweight) or a weight that is too low for a given height (wasting) (Rogol, Clark and Roemmich, 2000). Growth failures that are transmitted exclusively during pregnancy are measured in terms of low-birth weight (birth weight <2500 g). Iodine deficiency impairs the synthesis of thyroid hormones, including growth hormone expression in vitro (Ceda et al. 1992, Ezzat et al. 1991) and in vivo (Samuels *et al.*, 1989). The effect of growth hormones is mediated by insulin-like growth factor I, which is generally bound to insulin-like growth factor binding protein-3 in circulation (Jones and Clemmons, 1995; Ferry, Cerri and Cohen, 1999). Iodine supplementation of deficient populations has been shown to increase insulin-like growth factor I and insulin-like growth factor binding protein-3 (Zimmermann *et al.*, 2007).

As reviewed by Zimmermann (2011), existing cross-sectional epidemiologic studies on iodine deficiency and child growth have mostly found a positive relation. The existing randomized controlled studies (discussed below), which generally exhibit a higher internal validity, did only partly showed positive effects from iodine supplementation. In the context of this mixed evidence and, in addition, because of the nonexistence of large-scale cross-country

studies, we aimed to shed further light on a potentially general association between the household unavailability of iodized salt and outcomes of early childhood undernutrition in all low- and middle-income countries for which comparable data exists.

2.2 Methods

2.2.1 Data and Procedures

Data were from demographic and health surveys (DHSs), which were administered by *ICF International* and have been conducted at varying intervals since 1984. The DHSs were designed to collect nationally representative health and welfare data for women of reproductive age, their children, and their households. The DHSs have used a multistage stratified sampling design. Each country was divided into regions. Within these regions, populations were stratified by urban and rural area of residence. Within these stratified areas, a random selection of enumeration areas or primary sampling units (PSUs) was drawn. PSUs were selected on the basis of a probability that was proportional to the population size. In the second stage of sampling, all households within a PSU were listed from the most recent population census and about 25 households were randomly selected for an interview with the use of equal probability systematic sampling. Detailed sampling plans are available from survey final reports (ICF Macro, 2011). For each sampled household the household members were listed and women who were eligible for a more detailed interview were identified. Typically, these women were between the ages of 15 and 49 years. A few surveys were restricted to married or previously married women, and in others, the age range was 10-49 years. The DHSs provided weights for calculating nationally representative statistics. DHS data-collection procedures were approved by the ICF Macro International Institutional Review Board and by the relevant human subjects committee in each country. Oral informed consent for the interview was obtained from respondents.

2.2.2 Outcomes

We analyzed the data for stunting (low height-for-age), underweight (low weight-for-age), and wasting (low weight-for-height) in children aged between 0 and 59 months at the time of the interview, and low birth weight in newborns. Heights and weights of children between 0 and 59 months of age were measured and reported at the time of the interview in selected surveys and were used to calculate whether a child was stunted, underweight or wasted according to WHO standards and classifications. The z-score for stunting was calculated with the use of the child's

observed height minus the median height in the WHO reference population divided by the SD in the reference population (WHO Multicenter Growth Reference Study Group, 2006). Similar calculations were done for weight-for-age (for underweight) and weight-for-height (for wasting). We defined a child to be stunted, underweight or wasted if the child exhibited a *z*-score < -2. Similarly, a child was defined to be severely stunted, severely underweight, or severely wasted if it exhibited a *z*-score < -3 (Global Database on Child Growth and Malnutrition, no date). We excluded from our analysis children for whom *z*-scores were implausibly high or low, because of errors in measurement, data coding or reported age. Z-scores were defined to be implausibly high or low according to WHO standards if they were lower than -6 or higher than +6 for height-for-age, lower than -6 or higher than +5 for weight-for-age and lower than -5 or higher than +5 for weight-for-height (Mei and Grummer-Strawn, 2007). Birth weight was reported from a child's health card or, if the health card was not available, from the mother's recall. Also, according to the WHO definition we defined a birth weight < 2500 g as low birth weight. Extreme values for birth weight (> 9000 g) were considered to be measurement errors and were excluded from the analysis.

2.2.3 Exposure and Covariates

Our main variable of interest was the unavailability of iodized salt within a household. This information was first included in the DHS household's questionnaire in the third phase, which started in 1992 (USAID, 2008). The iodine content of the salt that was present in the household at the time of the interview was determined with the use of a rapid test kid (RTK) (ICCIDD Global Network, 2014). A few drops of a chemical liquid were added onto the surface of a salt sample provided by the household and the developed color was visually compared with a calibration chart, which related to certain amounts of iodine content [in parts per million (ppm)], generally 0, 1–15, 15–29, and ≥ 30 ppm (Pandav *et al.*, 2000). DHSs did not follow a uniform coding procedure for the result of the test. In some DHSs the exact number from the chart was reported, in other surveys it was only reported if the color of the salt test changed at all. The salt test is generally not considered to be very precise, and thus it is highly recommended to use it as a qualitative indicator only (Pandav *et al.*, 2000).

Therefore, we constructed an indicator variable, which took on the value of 1 if the salt used by a household did not include any iodine at all and 0 if it contained any amount of iodine (ppm not specified). The sex and age of the child, if the child was a twin, the birth order, if the

child was ever breastfed, and the diphtheria, pertussis, and tetanus immunization status of the child were included as covariates at the child level. The mother's marital status, education level, height, BMI, age at the birth of the child, and whether she has performed any work in the last 12 months were included as maternal covariates. The household's wealth quintile was included as household level covariate. The wealth quintile was a within-country measure of the wealth of the household relative to other households in that survey, based on ownership of household assets like radio, television, refrigerator, phone, motorcycle, bike, car, electricity, quality of floor, wall and roof material, flush toilet and piped drinking water. Only household assets for which information was available for ≥ 90% of the observations were included in the calculation of the asset index for a specific survey. Maternal height values > 2 and < 1 m were excluded from the analysis because they were likely to be the result of a measurement error.

2.2.4 Statistical Analysis

We specified conditional logistic regression models at the child level with stunting, underweight, wasting and low birth weight as outcome variables. All specifications were adjusted for PSU fixed effects. PSU fixed effects were survey specific. With the inclusion of PSU fixed effects, we controlled for observable and unobservable factors that were common at one point in time to all households within a PSU such as environmental factors, urban or rural location, or the local availability of food items. When we included PSU fixed-effects, observations were dropped from the analysis whenever there was no variation in the outcome variable within the PSU. We examined the association between the unavailability of iodized salt with stunting, underweight, wasting and low birth weight with and without adjustment for the covariates mentioned in the section *Exposure and Covariates*. All SEs were clustered at the PSU level. We also specified logistic regression models separately for different age groups. Age groups included children aged between 0 and 5 months, children aged between 6 and 23 months and children aged between 24 and 59 months. We also specified logistic regression models with the use of severe stunting, severe underweight and severe wasting as outcome variables. To investigate the association of the lack of iodine along the entire continuum of the undernutrition indicators, we also specified ordinary least squares regression models with the use of continuous z-scores and the child's birth weight (in kilograms) as outcome variables. These models were also adjusted for the same covariates and PSU level fixed effects.

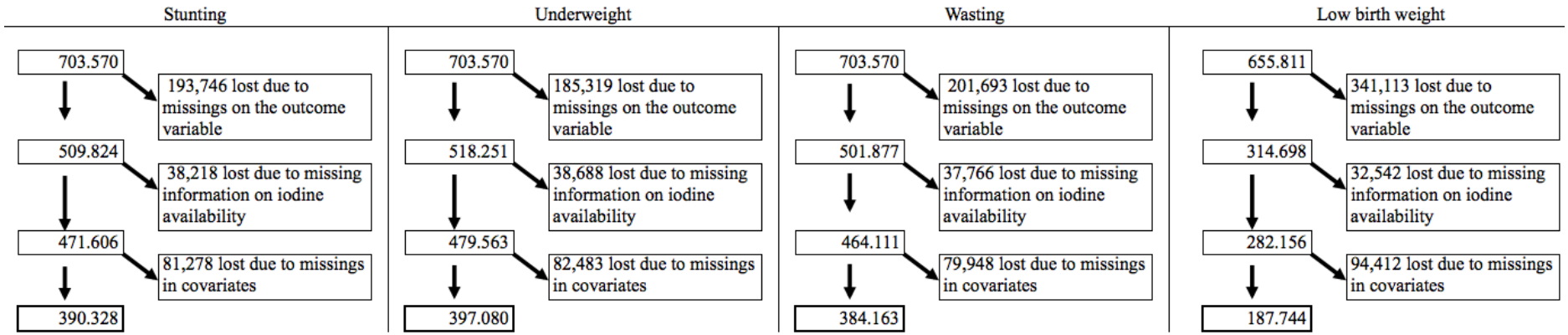
In a sensitivity analysis, we further controlled for the consumption of food items such as fish, eggs, poultry, meat and fruits that potentially included high levels of iodine. In other specifications we also controlled for prenatal healthcare visits and maternal iron consumption. We did not include these variables in our main specification, because there were many missing values. Because India constituted about 18 % of our observations, we also excluded India from our sample in one of our specifications in order to test if this country could drive results. All statistical analyses were done with Stata 14 software (StataCorp LP).

2.3 Results

2.3.1 Sample Description

Our initial sample for stunting, wasting and underweight consisted of the DHSs for which data on the iodine status of the salt used in the household as well as child anthropometric measure were collected. This left us with 89 surveys that were conducted in 46 different countries comprising 703,570 living children aged between 0 and 59 months at the time of the interview that are permanent members of the respective household. Data for low birth weight were only collected for children born after a certain threshold date. In phase 3 this information was collected for children born in 1991 or later, in phase 4 this information was collected for children born in 1995 or later, in phase 5 this information was collected for children born in 2001 or later and in phase 6 this information was collected for children born in 2005 or later. Thus, our initial sample of 655,811 children for low birth weight was built from all DHSs for which data on the iodine status of the salt used in the household as well as birth weight was collected, restricted to those children for whom low birth weight could potentially have been recorded. Data were collected between September 1994 and December 2012. After observations for which variables were missing were dropped, the final analytic sample for stunting, underweight, wasting, and low birth weight consisted of 390,328, 397,080, 384,163, and 187,744 observations, respectively. The different steps of the sample deduction are illustrated in Figure 2.1 (see the appendix 2.A table 2.A.1 for a detailed list of the surveys).

Figure 2.1: Sample deduction



Note: Numbers of observations in regressions were smaller. When including primary sampling unit fixed effects, observations were dropped whenever there was no variation in the outcome variable within the primary sampling unit. There was little variation especially for wasting and low birth weight.

Within the sample of children for whom stunting was measured, 298,743 children (76.5%) lived in households where iodized salt was available at the time of the interview, whereas 91,585 children (23.5%) lived in households, where no iodized salt was available. The corresponding numbers for the samples of children for whom underweight and wasting were measured are reported in the notes of Table 2.1. Prevalence of all child growth indicators were higher among children who lived in households without iodized salt compared to children who lived in households with iodized salt. Only the prevalence of severe wasting was not statistically different in the two groups (Table 2.1). Children from households without iodized salt were also less likely to be immunized, more likely to live in rural areas, more likely to live in households from the poorest wealth quintile and less likely to live in households from the richest wealth quintile. Their mothers were more likely to be married, more likely to have less than primary education and more likely not to work (Please refer to the appendix 2.A table 2.A.2 for means of covariates included).

Table 2.1: Estimated percentage distribution of outcome variables across exposure categories

	Availability of iodized salt in the household, % (N = 298,743)	Lack of iodized salt in the household, % (N = 91,585)
Outcome variables		
Stunting	36.6	40.4^{+}
Severe Stunting	23.6	24.8^{+}
Underweight	21.0	26.0^{+}
Severe underweight	16.0	17.1^{+}
Wasting	9.6	11.0^{+}
Severe Wasting	11.1	10.7
Low birth weight	10.3	13.1^{+}

N denotes the number of observations in the stunting sample. For other samples, total numbers of observations were slightly different as follows: underweight, N = 397,080; wasting, N = 384,163; and low birth weight, N = 196,238. Values were population weighted. $^{+}$ Proportions in the two groups were significantly different from each other (P-value $\leq$ 0.05).

2.3.2 Association between the Unavailability of Iodized Salt and Child Growth

Tables 2.2 and 2.3 show the adjusted and unadjusted ORs of the association between unavailability of iodized salt and child stunting, underweight and wasting for the full sample, and for severe stunting, underweight and wasting as alternative outcome variables, respectively. In models that were only adjusted for PSU fixed effects (unadjusted model), the unavailability of iodized salt was associated with 7.7% higher odds of being stunted, 11.6% higher odds of being underweight and 7.1% higher odds of being wasted (95% CIs of ORs: 1.05, 1.11; 1.08, 1.15; and 1.03, 1.12, respectively; P-value < 0.01 all three outcomes). In models that were adjusted for a full set of covariates (adjusted model) and PSU fixed effects, unavailability of iodized salt was

associated with 2.9% higher odds of being stunted (95% CI: 1.00, 1.06; P-value: 0.04), 5.4% higher odds of being underweight (95% CI: 1.02, 1.09; P-value < 0.01), and 3.4% higher odds of being wasted (95% CI: 0.99, 1.08; P-value: 0.12), whereas only stunting and underweight were significant at the 5% and 1% levels, respectively. Results for the covariates are reported in the appendix 2.A in table 2.A.3. After adjustment, results for severe undernutrition were not statistically significant (Table 2.3).

Results for different age groups are reported in the appendix 2.A table 2.A.4. In the fully adjusted model, the associations between the unavailability of iodized salt and stunting and underweight were quantitatively stronger in children between 24 and 59 months of age than in the reference category of children between 0 and 5 months of age and were not statistically significant for wasting. In children between 6 and 23 months of age, the association was quantitatively stronger for wasting than for the reference category of children between 0 and 5 months and not statistically significant for stunting and underweight.

Table 2.2: Association between unavailability of iodized salt within a household and growth failure (z-scores <-2) of children under the age of 5 for the full sample

	Unadjusted			Adjusted		
	Stunting (1)	Underweight (2)	Wasting (3)	Stunting (4)	Underweight (5)	Wasting (6)
Full sample						
Absence of iodized salt (OR)	1.08***	1.12***	1.07***	1.03**	1.05***	1.03
95% CI	[1.05,1.11]	[1.08,1.15]	[1.03,1.12]	[1.00,1.06]	[1.02,1.09]	[0.99,1.08]
P-value	(<0.01)	(<0.01)	(<0.01)	(0.04)	(<0.01)	(0.12)
N	349,341	294,835	215,208	349,341	294,835	215,208
PSU fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.3: Association between unavailability of iodized salt within a household and severe growth failure (z-scores <-3) of children under the age of 5

	Unadjusted			Adjusted		
	Severe Stunting (1)	Severe Underweight (2)	Severe Wasting (3)	Severe Stunting (4)	Severe Underweight (5)	Severe Wasting (6)
Absence of iodized salt (OR)	1.07***	1.11***	1.07**	1.01	1.04#	1.04
95% CI	[1.03,1.10]	[1.06,1.16]	[1.01,1.14]	[0.98,1.05]	[0.99,1.09]	[0.98,1.11]
P-value	(<0.01)	(<0.01)	(0.03)	(0.44)	(0.09)	(0.23)
N	280,453	190,427	127,317	280,453	190,427	127,317
PSU fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively. # denotes significance at the 10% level.

Table 2.4 shows results for low birth weight. In the unadjusted model unavailability of iodized salt was associated with 15.7% higher odds of being born with low birth weight (95% CI: 1.09, 1.24; P-value < 0.01). In the adjusted model, the unavailability of iodized salt was associated with 9.1% higher odds of low birth weight (95% CI: 1.02, 1.17; P-value: 0.01). Both coefficients were significant. Both coefficients are statistically significant.

Table 2.4: Association between unavailability of iodized salt within a household and low birth weight

	Unadjusted	Adjusted
	Low birth weight	Low birth weight
Absence of iodized salt	1.16***	1.09**
95 % CI	[1.09,1.24]	[1.02,1.17]
P-Value	(<0.01)	(0.01)
N	93,697	93,697
PSU fixed effects	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

In Table 2.5, we report regression results with the use of z-scores and birth weight (in kg) as continuous dependent variables. The unavailability of iodized salt was associated with a 0.024-lower weight-for-age z-score (95% CI: 20.04, 20.01; P-value < 0.01;) and a 25-g lower birth weight (95% CI: 20.04, 20.01; P-value < 0.01). Results for height-for-age and weight-for-height were not statistically significant.

Table 2.5: Association between unavailability of iodized salt within a household and growth failure of children < 5 years old using z-scores and birth weight (in kg) as continuous dependent variables

	Adjusted			
	Height-for-age	Weight-for-age	Weight-for-height	Birth weight
	(1)	(2)	(3)	(4)
Absence of iodized salt	-0.02	-0.02***	-0.01	-0.03***
95% CI	[-0.04, 0.00]	[-0.04,-0.01]	[-0.03, 0.00]	[-0.04,-0.01]
P-value	(0.12)	(<0.01)	(0.12)	(<0.01)
N	390,328	397,080	384,163	187,744
PSU fixed effects	Yes	Yes	Yes	Yes

Associations were estimated by ordinary least square models and marginal effects are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

In a sensitivity analysis we adjusted for the consumption of food items such as fish, eggs, poultry, meat and fruits, which could potentially have contained some amount of iodine.

This adjustment resulted in the insignificance of all undernutrition indicators. However, this result might have been driven by the reduced sample (in this specification, the stunting sample consisted of only 144,316 children because of missing observations). The same was true when we controlled for maternal iron consumption. In addition, controlling for prenatal health care visits, controlling for multiple children in the household or including survey fixed effects instead of PSU effects did not change the results. The results of the sensitivity analysis are presented in the appendix 2.A tables 2.A.5–2.A.12. The exclusion of India from the sample removed the effect for all outcome variables except of low birth weight, and consideration of only the India sample gave the strongest associations (Tables 2.6 and 2.7); the unavailability of iodized salt was associated with 8.8% higher odds of being stunted (95% CI: 1.02, 1.17; P-value: 0.02), 15.8% higher odds of being underweight (95% CI: 1.08, 1.24; P-value < 0.01), and 10.6% higher odds of being wasted (95% CI: 1.02, 1.20; P-value: 0.02) in the India only sample. The association with low birth weight was not statistically significant in the India sample.

Table 2.6: Association between unavailability of iodized salt within a household and growth failure of children < 5 years old with the exclusion of India from the sample

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.01	1.00	0.98	1.08**
95% CI	[0.98,1.04]	[0.97,1.04]	[0.94,1.03]	[1.00,1.17]
P-value	(0.65)	(0.95)	(0.53)	(0.05)
N	297,774	242,209	169,462	76,790
PSU fixed effects	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.7: Association between unavailability of iodized salt within a household and growth failure of children < 5 years old with only India included

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.09**	1.16***	1.11**	1.10
95% CI	[1.02,1.17]	[1.08,1.24]	[1.02,1.20]	[0.97,1.25]
P-value	(0.02)	(<0.01)	(0.02)	(0.142)
N	51,567	52,571	45,744	16,907
PSU fixed effects	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

2.4 Discussion

2.4.1 Interpretation

We showed that absence of iodized salt at the household level had a statistically significant association with childhood stunting, underweight and low birth weight for the full sample. The finding was robust to the adjustment for various confounding factors. The absence of any association for children until the age of 5 months was reasonable because these children were unlikely to have consumed salt directly, although some benefits via maternal breast milk could have been possible. The influence of adequate maternal iodine consumption during pregnancy was reflected in the statistically significant association with low birth weight, which was robust to the exclusion of India from the sample. The significant, although quantitatively small, association shown in our analysis might have been related to the fact that undernutrition outcomes are a result of a complex interplay of different nutritional factors, in which iodine deficiency constitutes only one factor. In addition, also environmental, psychosocial and socioeconomic factors are known to influence undernutrition indicators (Bhutta et al. 2013, Fenske et al. 2013).

However, our findings also suggest that the global association between the absence of iodized salt at the household level and child growth is largely driven by India, which was the country with the highest number of observations in the sample, and is not universal across countries but, rather, seems to depend on the specific country context. The association for low birth weight is however robust to the exclusion of India from the sample.

This finding was in line with previous studies, which were typically conducted in individual countries and showed rather mixed results for stunting, underweight, and wasting but generally positive results for low birth weight. For instance, children who were living in iodine deficiency areas in Iran (Azizi *et al.*, 1995) and India (Lal, Srivastava and Chandra, 1996) showed delayed growth, and in Kenya, the household availability of iodized salt was related to improved child growth (Neumann and Harrison, 1994). An intervention study from Tibet in which children aged 5-15 years received intramuscular iodized oil injection improved height-for-age z-scores but decreased other anthropometric indicators (Moreno-Reyes *et al.*, 2003). The addition of iodine to irrigation water in an iodine deficient area in China reduced childhood stunting (Ren *et al.*, 2002). In Morocco, a 10-months intervention in school-aged children with iodized salt improved median height-for-age and weight-for-age z-scores (Zimmermann *et al.*, 2007). A randomized, double blind, placebo-controlled trial from South Africa, investigated the effect of multiple-micronutrient fortified biscuits on child growth and did not find any significant effect (van Stuijvenberg *et al.*, 1999). Double blind iodine-supplementation trials from Mexico (Rivera *et al.*, 2001) and Bolivia (Bautista *et al.*, 1982) did not demonstrate growth-promoting effects. In contrast, a randomized study in Albania showed a decrease in stunting and underweight because of the iodization of oil (Zimmermann *et al.*, 2007). However, it might be that, in these intervention studies, existing effects were not detected because of small sample sizes.

The existing evidence on low birth weight indicates that intake of iodized oil or salt during pregnancy significantly increases the birth weight of the offspring. The increase was 101g in the Democratic Republic of Congo, 260 g in Belgium (Mahomed and Gülmezoglu, 2006) and 200 g in Algeria (Chaouki and Benmiloud, 1994). Mixed results were shown in the Philippines, where the association between iodized salt and capsules and higher birth weight was established in all regions but one. However, the study found lower birth weights when both iodized salt and capsules were combined (Mason *et al.*, 2002). In addition, there are ongoing randomized control trials in India and Thailand concerning the relationship between maternal iodine supplementation and various birth outcomes including birth weight (Melse-Boonstra *et al.*, 2012). A systematic review of the effects of iodized salt and iodine supplementation on prenatal and postnatal growth is currently ongoing (Farebrother *et al.*, 2015).

To the best of our knowledge we have conducted the first study on the basis of repeated cross-sectional and mutually comparable surveys from all low- and middle-income countries for which data on the availability of iodine in the household as well as on child anthropometrics

were available. We confirmed that there is substantial heterogeneity across countries for stunting, underweight and wasting, but also documented that, in the global sample the absence of iodized salt at the household level has a statistically significant positive association with childhood stunting and underweight. For low birth weight the association was also robust when India is excluded from the sample.

2.4.2 Limitations of the study

Our study has a few potential limitations. The household availability of iodized salt and child anthropometric measures were both determined at the time of the survey. However, child anthropometric status was the result of determinants that were effective before the survey. Therefore, the analysis is limited by the assumption that availability of iodized salt at household level at the time of the survey represents availability during a child's previous growth period and during gestation.

We were further constrained in the construction of the exposure variable. The exposure variable was generated with the use of the rapid test kit to assess the salt iodine content. These test kits are good qualitative indicators of iodine presence but have very limited ability to quantify iodine contents and detect adequate amounts of iodine (typically ≥15 ppm iodine), that are considered to meet dietary needs (Pandav et al. 2000, Jooste & Strydom 2010, WHO 2007). This constraint biased our estimate downward.

We also made the assumption that missing data in our sample were not systematically correlated with the true unobserved status of child growth and the availability of iodized salt in the household.

Although PSU fixed effects allowed us to control for variations in natural availability of iodine in soil and water and hence the severity of iodine deficiency across regions, we are not able to control for potential variation in the severity of iodine deficiency within a PSU. Because salt iodization is likely to have greatest benefits for growth in settings where baseline iodine status is low, we again assume that we were underestimating the association of a salt iodization policy in an iodine deficient country compared with a non-deficient country. In addition, salt intake of children may be low, and benefits for child growth in the early years might be limited.

The final concern is confounding of the results by unobserved factors. It could have been the case that the unavailability of iodized salt was a proxy for unobserved factors that are important for child growth and that we only observe the association between unavailability of

iodized salt and impaired child growth because this unobserved factor is correlated with both dependent and independent variables. We tried to address this concern to the extent possibly by adjusting for PSU fixed effects. By including PSU fixed effects, we controlled for factors that were similar at one point in time within a PSU. This included, e.g., the availability of different salt brands or other products that were iodine fortified and that were available in that particular PSU, as well as general dietary patterns within one PSU and local availability of food items. Hence confounding factors must have been on a level that is smaller than the PSU (i.e. at the household or individual level). Unfortunately, more detailed information on nutritional intake was only available for a small subset of observations and therefore we prefer to not use these control variables in the main specifications. The PSU fixed effects control for general dietary patterns and the local availability of food items that are common to all households within a PSU. Also, other control variables, such as household wealth and education were proxies, to some extent, for the quantity and quality of dietary intake. However, we could not fully rule out the possibility of an omitted variable bias.

In conclusion, our study documents the global association between the availability of iodized salt in households and stunting and underweight of children < 5 years old as well as of low birth weight. Whereas results for low birth weight were robust to the exclusion of India from the sample, results for stunting and underweight were not. Even though we did not establish causality in our analysis, the findings might indicate, that the causal effect of iodized salt on child growth, if it exists, is most profound in utero, and is not universally effective across all countries with respect to longer run child growth outcomes such as stunting and underweight.

2.A Appendix

Table 2.A.1: Sample size and estimated percentage of children living in households (HH) without iodized salt across countries (stunting sample)

	Phase (Year of survey)	N (% of children in HH without iodized salt)
Albania	Phase 5 (2008/2009)	1,345 (8.61)
Armenia	Phase 4 (2000), Phase 5 (2005)	2,392 (5.13)
Azerbaijan	Phase 5 (2006)	1,529 (5.66)
Bangladesh	Phase 6 (2011)	5,463 (17.73)
Burkina Faso	Phase 4 (2003), Phase 6 (2010)	12,705 (11.18)
Benin	Phase 3 (1996), Phase 4 (2001), Phase 5 (2006)	15,428 (13.74)
Bolivia	Phase 4 (2003/04), Phase 5 (2008)	14,283 (8.94)
Brazil	Phase 3 (1996)	3,308 (6.67)
Burundi	Phase 6 (2010/11)	2,488 (4.42)
Congo (DRC)	Phase 5 (2007)	2,272 (6.28)
Central Afr. Republic	Phase 3 (1994/95)	1,665 (72.57)
Congo (Brazzaville)	Phase 5 (2005), Phase 6 (2011/12)	5,938 (1.97)
Ivory Coast	Phase 6 (2011/12)	2,472 (9.92)
Cameroon	Phase 3 (1998), Phase 4 (2004), Phase 6 (2011)	7,702 (2.99)
Dom. Republic	Phase 3 (1996)	2,864 (87.32)
Egypt	Phase 4 (2000/03*), Phase 5 (2005/08*)	32,504 (22.7)
Ethiopia	Phase 4 (2000), Phase 5 (2005), Phase 6 (2011)	17,353 (72.54)
Gabon	Phase 4 (2000/01), Phase 6 (2012)	5,013 (42.48)
Ghana	Phase 4 (2003)	2,548 (66.51)
Guinea	Phase 4 (1999), Phase 5 (2005), Phase 6 (2012)	7,932 (30.84)
Guyana	Phase 5 (2009)	1,190 (78.78)
Haiti	Phase 4 (2000), Phase 5 (2005/06), Phase 6 (2012)	9,557 (80.57)
India	Phase 4 (1998/1999/2000), Phase 5 (2005/06)	53,801 (25.42)
Kenya	Phase 5 (2008/09)	4,171 (0.26)
Cambodia	Phase 4 (2000), Phase 5 (2005/06), Phase 6 (2010/11)	9,316 (41.36)
Lesotho	Phase 6 (2009/10)	1,293 (1.74)
Moldova	Phase 5 (2005)	1,177 (35.79)
Madagascar	Phase 3 (1997), Phase 4 (2003/04), Phase 5 (2008/09)	10,213 (28.37)
Mali	Phase 4 (2001), Phase 5 (2006)	15,047 (17.43)
Malawi	Phase 4 (2000), Phase 6 (2010)	10,533 (13.01)
Mozambique	Phase 4 (2003/04), Phase 6 (2011)	13,915 (49.93)
Nigeria	Phase 4 (2003), Phase 5 (2008)	18,825 (3.29)
Niger	Phase 3 (1998), Phase 5 (2006), Phase 6 (2012)	9,313 (27.86)
Nepal	Phase 6 (2011)	1,960 (3.7)
Peru	Phase 5 (2008/09), Phase 6 (2010/11/12)	24,628 (2.69)
Rwanda	Phase 4 (2000), Phase 5 (2005), Phase 6 (2010/11)	10,935 (1.35)
Sierra Leone	Phase 5 (2008)	1,516 (21.02)
Senegal	Phase 4 (2005), Phase 6 (2010/11)	5,195 (43.07)
Sao Tome and Principe	Phase 5 (2008/09)	990 (7.72)
Swaziland	Phase 5 (2006/07)	1,603 (3.2)
Chad	Phase 4 (2004)	3,159 (20.72)
Turkey	Phase 4 (2003/04)	1,489 (54.33)
Tanzania	Phase 4 (2004/05), Phase 6 (2009/10)	11,018 (23.65)
Uganda	Phase 3 (1995), Phase 4 (2000/01), Phase 5 (2006),	10,654 (12.28)

	Phase 6 (2011)	
Zambia	Phase 3 (1996/97), Phase 4 (2001/02)	8,212 (49.01)
Zimbabwe	Phase 6 (2010/11)	3,414 (6.14)
Total		**390,328 (24.37)**

*Two or more surveys within one phase.

Table 2.A.2: Estimated percentage distribution of outcome variables and covariates across exposure categories

	Availability of iodized salt in the household (percentage) N= 298,743	Lack of iodized salt in the household (percentage) N= 91,585
Outcome variables		
Stunting	40.6	43.1+
Severe Stunting	23.6	24.8+
Underweight	28.7	32.0+
Severe underweight	16.0	17.1+
Wasting	17.6	17.4
Severe Wasting	11.0	10.6
Low birth weight	20.7	22.7+
Covariates		
Child		
Sex Female	49.2	49.3
Age in month		
<1	0.9	1.0
1 -11	23.8	24.3+
12 -35	40.7	40.8
36 - 59	34.5	34.0+
Age in month, mean (Se)	26.9 (3.5)	26.6 (6.1)+
Multiple birth	2.4	2.4
Birth order		
First child	23.6	21.0+
Second child	21.9	19.9+
Third child	16.4	16.1+
Fourth child	11.7	12.2+
Fifth or later born child	26.4	30.8+
Ever breastfed	98.1	98.2
Full DPT immunization	70.6	61.0+
Maternal		
Married	78.1	81.2+
Education		
None	58.6	74.4+
Primary	26.4	20.2+
Secondary or higher	15.1	5.4+
Height		
> 159.9 cm	30.7	28.6+
155 - 159.9 cm	27.0	28.8+
150 - 154.9 cm	24.3	25.9+
145 - 149.9 cm	13.4	12.6+
< 145 cm	4.7	4.1+
Age at birth in years		
< 17	3.2	3.6+
17 -19	12.4	13.1+
20 - 24	29.8	29.1+
25 - 29	25.2	24.6+
> 29	29.4	29.6
Age at birth in years, mean (Se)	26.3 (1.8)	26.2 (3.4)
Not working	40.0	42.9+
Household		
Urban household	33.3	22.3+
Wealth quintile within survey		
First (Poorest)	19.2	23.5+
Second	19.3	22.9+
Third	19.8	20.8+
Fourth	20.4	18.4+
Fifth	21.3	14.4+

The numbers of observations (N) refer to the stunting sample. Values are population weighted.

+ Indicates that proportions in two groups are statistically significantly different from each other (P-value ≤ 0.05).

Table 2.A.3: Associations between the full set of covariates and growth failure of children < 5 years of age and low birth weight (including fixed effects)

	Stunting	Underweight	Wasting	Low birth weight
	(1)	(2)	(3)	(4)
Absence of iodized salt	1.03**	1.05***	1.03	1.09**
	[1.00,1.06]	[1.02,1.09]	[0.99,1.08]	[1.02,1.17]
	(0.04)	(<0.01)	(0.12)	(0.01)
Female	0.81***	0.83***	0.85***	1.34***
	[0.80,0.83]	[0.82,0.85]	[0.83,0.87]	[1.29,1.39]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Age category in month (Reference: 0 to 5 month)				
6-23 month	1.95***	4.79***	0.84	
	[1.29,2.95]	[2.17,10.60]	[0.59,1.19]	
	(<0.01)	(<0.01)	(0.32)	
24-35 month	6.71***	7.50***	0.53***	
	[4.44,10.14]	[3.38,16.64]	[0.37,0.75]	
	(<0.01)	(<0.01)	(<0.01)	
36-59 month	5.90***	6.48***	0.31***	
	[3.91,8.92]	[2.92,14.36]	[0.22,0.44]	
	(<0.01)	(<0.01)	(<0.01)	
Multiple birth	2.25***	2.77***	1.65***	12.80***
	[2.12,2.40]	[2.60,2.96]	[1.52,1.79]	[11.55,14.19]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Birth order (Reference: first child)				
Second child	1.12***	1.06***	1.01	0.75***
	[1.09,1.14]	[1.03,1.09]	[0.97,1.05]	[0.72,0.79]
	(<0.01)	(<0.01)	(0.50)	(<0.01)
Third child	1.22***	1.15***	1.04	0.67***
	[1.19,1.26]	[1.11,1.19]	[0.99,1.09]	[0.63,0.72]
	(<0.01)	(<0.01)	(0.12)	(<0.01)
Fourth child	1.28***	1.21***	1.03	0.65***
	[1.23,1.32]	[1.17,1.26]	[0.98,1.09]	[0.61,0.71]
	(<0.01)	(<0.01)	(0.26)	(<0.01)
Fifth or later born child	1.41***	1.34***	1.10***	0.63***
	[1.36,1.46]	[1.28,1.39]	[1.04,1.17]	[0.58,0.68]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Breastfeeding	0.88#	1.01	1.05	
	[0.76,1.02]	[0.82,1.23]	[0.78,1.42]	
	(0.10)	(0.96)	(0.75)	
Immunization	0.86***	0.80***	0.87***	
	[0.84,0.88]	[0.78,0.83]	[0.84,0.90]	
	(<0.01)	(<0.01)	(<0.01)	
Mother married	0.91***	0.90***	0.92**	0.88***
	[0.88,0.94]	[0.86,0.94]	[0.85,0.99]	[0.82,0.95]
	(<0.01)	(<0.01)	(0.02)	(<0.01)
Maternal Education (reference: no education)				
Primary	0.86***	0.80***	0.89***	0.86***
	[0.84,0.88]	[0.78,0.83]	[0.85,0.92]	[0.81,0.91]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Secondary or higher	0.70***	0.64***	0.83***	0.73***
	[0.67,0.73]	[0.61,0.68]	[0.78,0.88]	[0.68,0.79]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Maternal height (reference: taller than 159.9 cm)				
155 – 159.9	1.47***	1.34***	1.04#	1.21***
	[1.43,1.51]	[1.30,1.38]	[0.99,1.08]	[1.14,1.29]
	(<0.01)	(<0.01)	(0.09)	(<0.01)
150 – 154.9	1.95***	1.73***	1.12***	1.44***
	[1.89,2.01]	[1.67,1.79]	[1.06,1.17]	[1.34,1.54]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
145 – 149.9	2.75***	2.32***	1.25***	1.66***
	[2.65,2.85]	[2.22,2.43]	[1.17,1.33]	[1.53,1.81]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Smaller than 145	4.16***	3.31***	1.34***	2.14***
	[3.95,4.38]	[3.12,3.51]	[1.24,1.45]	[1.92,2.39]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Maternal BMI (reference: normal weight >=18.5)				

>=17.1 <=18.4 slight u.	1.20***	1.54***	1.39***	1.31***
	[1.16,1.23]	[1.50,1.59]	[1.34,1.45]	[1.22,1.41]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
>=16.1 <=17 moderate u.	1.25***	1.83***	1.74***	1.51***
	[1.19,1.32]	[1.74,1.92]	[1.64,1.85]	[1.33,1.71]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
>=16 severe u.	1.42***	2.51***	2.41***	1.87***
	[1.33,1.52]	[2.34,2.68]	[2.23,2.60]	[1.62,2.16]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
>=25 Overweight	0.78***	0.64***	0.74***	0.84***
	[0.76,0.80]	[0.62,0.66]	[0.71,0.77]	[0.80,0.89]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Maternal age at birth in years (Reference: younger than 17)				
17 – 19	0.84***	0.84***	1.01	0.94
	[0.78,0.90]	[0.76,0.92]	[0.87,1.16]	[0.81,1.09]
	(<0.01)	(<0.01)	(0.93)	(0.42)
20 – 24	0.73***	0.77***	0.99	0.84**
	[0.68,0.79]	[0.70,0.84]	[0.86,1.15]	[0.72,0.98]
	(<0.01)	(<0.01)	(0.92)	(0.02)
25 – 29	0.66***	0.73***	1.03	0.81**
	[0.60,0.71]	[0.66,0.80]	[0.89,1.19]	[0.69,0.96]
	(<0.01)	(<0.01)	(0.72)	(0.01)
>29	0.60***	0.70***	1.03	0.91
	[0.55,0.65]	[0.64,0.77]	[0.88,1.20]	[0.77,1.08]
	(<0.01)	(<0.01)	(0.73)	(0.28)
Mother not working	0.99	0.99	0.99	1.01
	[0.97,1.02]	[0.96,1.02]	[0.96,1.03]	[0.96,1.06]
	(0.43)	(0.61)	(0.78)	(0.79)
Wealth quintiles, within surveys (Reference: first =poorest)				
Second quintile	0.96***	0.96***	0.94***	0.90***
	[0.94,0.99]	[0.93,0.99]	[0.90,0.98]	[0.84,0.96]
	(0.01)	(<0.01)	(<0.01)	(<0.01)
Third quintile	0.89***	0.87***	0.88***	0.89***
	[0.86,0.91]	[0.84,0.90]	[0.84,0.92]	[0.83,0.96]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Fourth quintile	0.81***	0.78***	0.87***	0.89***
	[0.79,0.84]	[0.75,0.81]	[0.82,0.91]	[0.82,0.97]
	(<0.01)	(<0.01)	(<0.01)	(0.01)
Fifth quintile	0.67***	0.66***	0.82***	0.85***
	[0.65,0.70]	[0.63,0.70]	[0.78,0.88]	[0.77,0.93]
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
PSU fixed effects	Yes	Yes	Yes	Yes
N	349,341	294,780	215,206	93,697

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively. # denotes significance at the 10% level.

Table 2.A.4: Association between unavailability of iodized salt within a household and growth failure (z-scores <-2) of children < 5 years of age and the interaction with age groups (0 - 6 months as reference category)

	Unadjusted			Adjusted		
	Stunting	Underweight	Wasting	Stunting	Underweight	Wasting
	(1)	(2)	(3)	(4)	(5)	(6)
Absence of iodized salt (OR)	0.96	1.01	1.00	0.93#	0.97	0.97
95% CI	[0.89,1.04]	[0.93,1.10]	[0.92,1.09]	[0.86,1.00]	[0.89,1.05]	[0.89,1.06]
P-value	(0.30)	(0.80)	(0.95)	(0.06)	(0.48)	(0.52)
Age 6 – 23 month	2.59***	1.87***	0.90***	2.64***	1.82***	0.85***
95% CI	[2.47,2.71]	[1.77,1.98]	[0.85,0.95]	[2.51,2.77]	[1.72,1.92]	[0.80,0.90]
P-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age 24 – 59 month	3.83***	1.85***	0.42***	4.03***	1.84***	0.41***
95% CI	[3.66,4.01]	[1.75,1.94]	[0.40,0.45]	[3.84,4.23]	[1.74,1.94]	[0.38,0.43]
P-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Absence of iodized salt *Age 6 – 23 month	1.08**	1.08#	1.12**	1.05	1.05	1.10**
95% CI	[1.00,1.17]	[1.00,1.18]	[1.03,1.22]	[0.97,1.13]	[0.96,1.14]	[1.01,1.21]
P-value	(0.04)	(0.06)	(0.01)	(0.20)	(0.27)	(0.03)
Absence of iodized salt *Age 24 – 59 month	1.18***	1.15***	1.05	1.16***	1.13***	1.04
95% CI	[1.09,1.27]	[1.06,1.24]	[0.96,1.15]	[1.08,1.25]	[1.04,1.23]	[0.95,1.14]
P-value	(0.00)	(0.00)	(0.26)	(0.00)	(0.00)	(0.37)
N	349,341	294,835	215,208	349,341	294,835	215,208
PSU fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively. # denotes significance at the 10% level.

Table 2.A.5: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age controlling for different food items

	Adjusted		
	Stunting (1)	Underweight (2)	Wasting (3)
Absence of iodized salt	1.02	1.04	1.00
95% CI	[0.98,1.07]	[0.99,1.10]	[0.94,1.07]
P-value	(0.37)	(0.14)	(0.93)
Fish, eggs, poultry	0.78***	0.70***	0.84**
95% CI	[0.73,0.84]	[0.63,0.78]	[0.73,0.96]
P-value	(<0.01)	(<0.01)	(0.01)
Meat	0.82***	0.85***	0.96
95% CI	[0.78,0.87]	[0.78,0.92]	[0.86,1.08]
P-value	(<0.01)	(<0.01)	(0.54)
Fruits	0.84***	0.83***	0.90
95% CI	[0.79,0.90]	[0.75,0.92]	[0.79,1.03]
P-value	(<0.01)	(<0.01)	(0.13)
PSU fixed effects	Yes	Yes	Yes
N	144,316	113,445	80,033

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.A.6: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age using the food sample

	Adjusted		
	Stunting (1)	Underweight (2)	Wasting (3)
Absence of iodized salt	1.02	1.04	1.00
95% CI	[0.98,1.07]	[0.99,1.101]	[0.94,1.08]
P-value	(0.31)	(0.12)	(0.92)
PSU fixed effects	Yes	Yes	Yes
N	144,316	113,445	80,033

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit.

Table 2.A.7: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age controlling for maternal consumption of iron

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.02	1.03	1.01	1.12***
95% CI	[0.99,1.06]	[0.98,1.07]	[0.96,1.07]	[1.03,1.22]
P-value	(0.25)	(0.22)	(0.71)	(0.01)
Maternal iron consumption	0.95***	0.97**	0.99	0.82***
95% CI	[0.93,0.98]	[0.94,1.00]	[0.95,1.03]	[0.77,0.88]
P-value	(<0.01)	(0.05)	(0.60)	(<0.01)
PSU fixed effects	Yes	Yes	Yes	Yes
N	197,864	155,551	109,408	49,711

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.A.8: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age controlling for using the maternal consumption of iron-sample

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.02	1.03	1.01	1.12**
95% CI	[0.99,1.06]	[0.99,1.07]	[0.96,1.07]	[1.03,1.22]
P-value	(0.25)	(0.21)	(0.70)	(0.01)
PSU fixed effects	Yes	Yes	Yes	Yes
N	197,864	155,551	109,408	49,711

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. ** denotes significance at the 5% level, respectively.

Table 2.A.9: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age controlling for prenatal health care visits

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.03#	1.06***	1.05#	1.09**
95% CI	[1.00,1.06]	[1.02,1.09]	[1.00,1.10]	[1.01,1.17]
P-value	(0.09)	(<0.01)	(0.05)	(0.03)
Prenatal health care visits (reference: no prenatal visits				
1 visit	0.99	0.96	1.04	1.10
95% CI	[0.94,1.04]	[0.91,1.01]	[0.97,1.11]	[0.86,1.40]
P-value	(0.63)	(0.17)	(0.34)	(0.43)
2 to 3 visits	0.92***	0.94***	0.98	0.82*
95% CI	[0.89,0.95]	[0.91,0.98]	[0.93,1.03]	[0.66,1.02]
P-value	(<0.01)	(<0.01)	(0.46)	(0.07)
4 to 9 visits	0.85***	0.85***	0.93***	0.65***
95% CI	[0.82,0.88]	[0.82,0.89]	[0.88,0.97]	[0.53,0.80]
P-value	(<001)	(<0.01)	(<0.01)	(<0.01)
10 to 15 visits	0.785***	0.739***	0.925	0.533***
95% CI	[0.74,0.83]	[0.68,0.80]	[0.84,1.02]	[0.43,0.67]
P-value	(<0.01)	(<0.01)	(0.11)	(<0.01)
16 to 30 visits	0.89**	0.86#	0.86#	0.75**
95% CI	[0.801,0.997]	[0.743,1.006]	[0.731,1.016]	[0.571,0.983]
P-value	(0.04)	(0.06)	(0.08)	(0.04)
PSU fixed effects	Yes	Yes	Yes	Yes
N	256,260	210,213	154,508	65,194

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively. # denotes significance at the 10% level.

Table 2.A.10: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age using survey fixed effects

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.03**	1.06***	1.04**	1.08***
95% CI	[1.01,1.05]	[1.04,1.09]	[1.01,1.07]	[1.07,1.13]
P-value	(0.02)	(<0.01)	(0.02)	(<0.01)
N	390,328	397,080	384,163	186,252
Survey fixed effects	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.A.11: Association between unavailability of iodized salt within a household and severe growth failure of children < 5 years of age using survey fixed effects

	Adjusted		
	Severe Stunting (1)	Severe Underweight (2)	Severe wasting (3)
Absence of iodized salt	1.03**	1.06***	1.02
95% CI	[1.00,1.06]	[1.02,1.10]	[0.97,1.08]
P-value	(0.03)	(<0.01)	(0.36)
N	390,328	397,080	384,163
Survey fixed effects	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

Table 2.A.12: Association between unavailability of iodized salt within a household and growth failure of children < 5 years of age controlling for multiple children surveyed in one household

	Adjusted			
	Stunting (1)	Underweight (2)	Wasting (3)	Low birth weight (4)
Absence of iodized salt	1.03**	1.05***	1.03	1.09**
95% CI	[1.00,1.06]	[1.02,1.09]	[0.99,1.08]	[1.01,1.16]
P-value	(0.046)	(<0.01)	(0.12)	(0.01)
Multiple children	1.05***	1.000	0.94***	0.97
95% CI	[1.03,1.07]	[0.98,1.02]	[0.91,0.96]	[0.93,1.01]
P-value	(<0.01)	(0.95)	(<0.01)	(0.19)
N	349,341	294,780	215,206	94,739
PSU fixed effects	Yes	Yes	Yes	Yes

Associations were estimated by logistic regression models and odds ratios are shown in this table. 95% Confidence intervals (CI) are reported in brackets and p-values are reported in parentheses. Standard errors are clustered by the primary sampling unit. *** and ** denote significance at the 1% and 5% level, respectively.

3. School Feeding, Iron-Fortified Salt and Child Cognitive Ability – Evidence from a Randomized Controlled Trial in Rural India[†]

BY MARION KRÄMER, SANTOSH KUMAR AND SEBASTIAN VOLLMER

Abstract

The Indian public school-feeding program might be an effective channel to deliver iron and reduce the high-risk of iron deficiency that is common among children in rural India. We test the effectiveness of such an intervention by running a randomized controlled trial in which 54 randomly selected government-funded schools, from two blocks in Bihar, India, used iron-fortified iodized salt to prepare the school lunch for a period of one year. The same number of randomly selected schools functioned as control schools and continued using conventional iodized salt. In addition to anemia, we investigated the effect on cognitive and education outcomes. The treatment reduced the prevalence of any form of anemia by 20% and of mild anemia by 30%. There is no statistically significant interaction effect between treatment and school attendance, i.e. the treatment effect does not change significantly with increasing compliance. Despite the short treatment period, there is however weak evidence that with high school attendance (80 or 90% school attendance) the treatment had a positive effect on reading and math scores. We further find that the intervention is cost-effective. These results provide encouraging evidence of the positive effects from using school-feeding programs as a channel to providing school-aged children with iron-fortified salt.

[†] Ethical clearance: The Medical ethics commission of the University of Göttingen approved the study. Informed written consent for the survey, including the medical tests, was taken from the parents. Informed consent was also taken from the headmasters. Acknowledgments: We thank the Foundation *fiat panis* for providing funding for this study, as well as the German Research Foundation (DFG), which provided funding within the scope of the RTG 1666.

3.1 Introduction

Adequate health and nutrition, both in terms of micro- and macronutrients, constitute the cornerstones of human capital formation. Only a well-nourished and healthy child will attend school regularly, react properly to the stimuli from its environment and learn sufficiently to achieve good education outcomes. In contrast, a low level of education might result in low adult productivity and income poverty, which in turn might lead to undernutrition for the individual as well as his or her offspring. This *nutritional poverty trap* might hence enforce the intergenerational transmission of poverty (e.g. Dasgupta and Ray, 1986 and Strauss & Thomas 1998) .

In this paper, we study the impact of an intervention that aims at tackling one of the most widespread nutritional impairments in the world: Iron deficiency. Iron deficiency mostly emerges from inadequate dietary intake of iron or the malabsorption of iron stemming from an individual's diet.[13] Iron deficiency is determined through the concentration of hemoglobin, a protein in the blood, which is essential to carry oxygen through the body (McLean *et al.*, 2009). It is estimated that globally about 1.62 billion people, more than 20% of the world's population, are anemic (WHO, 2008), a situation where the concentration of hemoglobin falls below a clinical threshold.[14] [15] Despite the well-known adverse effects of iron deficiency on physical activity (fatigue and loss of energy), iron deficiency also affects the proper development of the central nerve system and can hence directly hinder human capital formation in the form of impaired cognitive development (Halterman et al. 2001, Bobonis et al. 2006). It has been estimated that reduced cognitive development due to iron deficiency is associated with a 5-10 points lower intelligence quotient (WHO, 2001). The impairment of physical activity and cognitive development might result in a substantial productivity loss. Therefore, iron deficiency is argued to be the costliest micronutrient deficiency from an economic point of view. Using data from 10 low-income countries, Horton & Ross (2003) estimate that physical and cognitive impairment, due to iron deficiency, cause a loss of 4.05% of a country's GDP at the median. Nutrient deficiency in childhood is especially harmful

[13] Anemia can have other causes too, e.g. excessive bleeding, hookworm infections or malaria (WHO, 2001). Since all children in the data set used for the analysis are dewormed at school once a year, we are quite sure that most of the anemia observed in our study comes from iron deficiency.

[14] Though anemia can have different causes (e.g. excessive bleeding, hookworm infections or malaria), iron deficiency is known to be the most common (WHO, 2001).

[15] For the identification of an anemic individual, the authors used the age and gender specific WHO hemoglobin cutoffs for mild anemia (equivalent to *any anemia* in our study). The anemia thresholds are defined differently depending on age and gender: For children under 5 years: < 11.0 g/dl; children 6–11 years: < 11.5 g/dl; children 12–14 years: < 12.0 g/dl; adult males: < 13.0 g/dl; adult females (non-pregnant): < 12.0 g/dl; adult females (pregnant): < 11.0 g/dl (WHO, 2008).

since its adverse effects are supposed to accumulate over time (e.g. Maluccio et al. 2009 for a protein-energy intervention and Lozoff et al. 2006 for iron supplementation).

Anemia in India is highly prevalent. Depending on age and gender, anemia rates range between 23% and 58% at the national level (NFHS 2015). Our evaluation took place in the Jehanabad district in the state of Bihar, where the prevalence of anemia is even higher and ranges from 26% for adult males to 63% among children below the age of five (NFHS, 2015a). The Indian low-income population is especially prone to a high-risk of iron deficiency due to a lack of dietary diversity. Moreover, their diet generally includes a large amount of rice and wheat, which inhibits the absorption of iron due to the high concentrations of phytate in these products, and the low consumption of meat from which iron can more easily be absorbed (FAO and WHO, 2002).

There is a strong biological basis for the positive impacts of increased iron consumption on health and cognitive development (e.g. Beard 2008 or Kretchmer et al. 1996). Different strategies and technologies have been developed and discussed to increase iron dietary intake (e.g. supplementary pills, different fortified product, including biofortified products and dietary diversification). Different channels to deliver these products have also been tested (e.g. retail market, agricultural extensions, public institutions such as hospitals or schools etc.). In a world of scarce resources, the challenge for policy makers lies in the identification of the most cost-effective combination of technology and the delivery channel.

In this paper, we test if the Indian school-feeding program (also commonly called *Midday Meal* or MDM) constitutes an effective delivery channel to provide school-aged children with *iron-fortified iodized salt* (also called *double fortified salt* or *DFS*). We ran a randomized controlled trial (RCT) in two blocks in rural Bihar, India. Over a period of one year we delivered DFS to 54 randomly selected government-funded schools that used the DFS to prepare the MDM that is cooked directly at the schools and provided to all school children up to grade eight free of charge. We tested the effect of the supply of DFS to those schools against 53 randomly selected control schools that continued using conventional iodized salt. As such, any effect found through this intervention, can be attributed to the iron component of the salt because iodized salt is already widely used in rural Bihar.[16] We analyzed how this intervention affects hemoglobin and anemia, cognitive ability and education outcomes of children that were in the second grade at the baseline.

[16] We tested all major salt brands we came across during the baseline data collection for its exact iodine content and found that all were sufficiently iodized (at least 15 ppm of iodine).

Salt is a very promising vector to deliver iron to any population, since nearly every household in the world consumes it daily, irrespective of location, socioeconomic status or food preferences, i.e. there is a high likelihood of achieving a high level of coverage. In contrast to supplementation, fortified products are supposed to have high compliance as they only substitute the conventional product. Yet, empirical evidence shows that if people are given the choice between the conventional and the fortified product, they do not necessarily choose the fortified option. In another district of Bihar, Banerjee et al. (2014) evaluated the demand for the same DFS we are using in our study, but offered it in local private shops and via the Public Distribution System (PDS) (local public shops in which people identified as poor by means-testing are entitled to buy certain food items at subsidized prices), where people had the choice between conventional salt and DFS. They found that only 33% of the households that were offered the DFS at a subsidized price of nine rupees per kg (0.13 USD), a price that is already one rupee below the price of conventional salt on the market, purchased it at the local shops and only 13% purchased it at the PDS shops. Their results did not change when standard advertisement campaigns were run.[17] After two to three years of the intervention, less than 10% were consuming DFS at the subsidized rate of nine rupees. In another study arm, Banerjee et al. (2016) distributed DFS directly and free of costs to a randomly selected number of households. Nevertheless, compliance was only about 61-75% (61% were using it at the endline survey and 14% reported that they had just recently ran out of DFS). Adding to this, Banerjee et al. (2011) conducted a RCT where they offered decentralized flour iron-fortification at local mills. Again, they did not find any positive impact at the endline due to a rapid drop in take-up six months after the start of the intervention. These studies indicate that a behavioral change in purchasing and consumption habits, even if it is only a small one, might be difficult to induce at the household level.

Our main point for considering the MDM to be an effective delivery channel of iron (and potentially other micronutrients) is that in contrast to the approach chosen by Banerjee et al. (2014) and Banerjee et al. (2011), the usage of fortified products in public programs is not a private individual decision, but a governmental one and hence, if properly implemented, has the potential of high compliance. Using DFS in the MDM is comparable to a (partly) mandatory fortification policy. The usage of DFS in the MDM further enables a regular (nearly daily) provision of iron dosages to children, which is the recommended proper intake. Third, since government-funded school infrastructure is well established even in rural areas (in India

[17] A stronger communication campaign and an incentive intervention were somewhat effective and raised take-up by about 10 percentage points (Banerjee et al. 2015).

nearly every village has a school), and most children from low-income households visit government-funded schools, school-feeding programs are capable to reach a high-risk population: children of rural low-income families. Children of rural low-income families are at high-risk for iron deficiency because of two reasons. First, as stated above, they often have a lack of dietary diversity; and second, primary school children are at the age of critical brain development where the iron requirement is especially high.[18] Fourth, school-feeding programs are supposed to be very cost-effective since an already existing distributional infrastructure is used (Bobonis, Miguel and Puri-Sharma, 2006).

Different empirical studies have shown that school-feeding programs can effectively be used to deliver micronutrients to school-aged children. Van Stuijvenberg (2005) reports several RCTs conducted by the South African Medical Research Council (MRC) that tested the effect of using different micronutrient carriers (biscuits, a bread spread and a soup powder fortified with multiple-micronutrients including iron) in the South African school-feeding program. In all interventions, they found positive effects on iron status. Improvements in iron status were also found for iron-fortified whole maize flour used at four schools in Kenya (Andang'o *et al.*, 2007) and for iron-fortified rice used to prepare school lunches in Bangalore, India (Moretti *et al.*, 2006) and in Vishakhapatnam, India (Radhika *et al.*, 2011). Those studies rarely looked at productive outcomes such as cognitive ability or education outcomes. Only in the biscuit intervention in the South African school-feeding program, where some cognitive tests were performed, an increase in some of these cognitive tests (the forward digit-span test, verbal fluency task and reading numbers task) was found. Those effects were especially strong for those children that had low iron levels before the intervention (van Stuijvenberg 2005).

Adding to this literature, our study finds that the usage of DFS in the Indian MDM program reduced the prevalence of any form of anemia by more than 20% and the prevalence of mild anemia by about 30%. Effects do not differ when school attendance increases, i.e. with increasing compliance with the treatment. However, despite the short treatment period, there is weak evidence that reading and math scores were positively affected with high school attendance (80 or 90% school attendance). There is some weak evidence that children from historically disadvantaged groups (scheduled castes and tribes, SC/ST) benefitted slightly more from the treatment in the form of reduced anemia. However, children from non-SC/ST households might have benefited more in terms of higher education outcomes. We further calculate that the intervention is cost-effective. This is despite the fact that mild but not

[18] In low-income countries, 43.6% of school-aged children are suggested to be anemic (McLean et al, 2009).

moderate and severe anemia, which account for the highest share of the disease burden from anemia, were reduced. Per disability-adjusted life year (DALY) averted, we calculate total costs of the intervention of 53.52 USD or 231.94 I$ for the statistically significant effect found on mild anemia. These results are encouraging with respect to delivering iron-fortified salt through the Indian school-feeding program.

To the best of our knowledge, this is the first large-scale study that rigorously evaluates the usage of salt as an iron carrier in school-feeding programs. Furthermore, in contrast to existing studies, we assess not only health, but also productive outcomes such as cognitive abilities and education outcomes.

The reminder of this paper is structured as follows: Chapter 3.2 describes the intervention and the context. The study design and data are explained in chapter 3.3 and chapter 3.4 outlines the estimation strategy. Our estimation results are presented in chapter 3.5. In chapter 3.6, we discuss the challenges to the internal validity of our results and in chapter 3.7 we perform a cost-effectiveness analysis of the intervention. We conclude in chapter 3.8.

3.2 Context and Intervention

3.2.1 India's School-feeding Program

India's school-feeding program was initiated in some Indian states, including Tamil Nadu and Gujarat, by the mid-1980s as an initiative of the respective state governments. The implementation of the MDM aimed at overcoming classroom hunger, improving the nutritional level of school children and the enhancement of enrolment and attendance, which should consequently lead to improvements in education outcomes. It also aimed at reducing caste discrimination and enhancing gender equity. Covering an estimated 104.5 million school children in 1.16 million schools during the school year 2013/14, the program is the largest school-feeding program in the world (The Ministry of Human Resource Development, 2016a). In Bihar, it now covers all children from grade one to eight (The Ministry of Human Resource Development, 2016b).

Each primary school child is supposed to receive a daily lunch meal that is predefined in calories and composition of food items. The menu is fixed by the state government for all schools, varies on a daily basis but is repeated every week. Hence, the same food is supposed to be offered at each and every school. In most regions, including the state of Bihar, the Midday Meal is prepared directly at the schools in school kitchens explicitly built for this

purpose.[19] The Food Corporation of India provides grains directly to the schools and a representative of the school, usually the headmaster, individually buys the remaining ingredients at the local market (vegetables, pulses, oil and spices). For every primary school child (class one to five), the schools receive 3.59 Rs. (0.05 USD)[20] from the government, and for every upper-primary school child (class six to eight), the schools receive 5.38 Rs. (0.07 USD) per day to cover the cost of the additional ingredients. The MDM is served every day except Sundays and holidays (The Ministry of Human Resource Development, 2016b).

3.2.2 Distribution of Iron-fortified Iodized Salt

In this study, we delivered DFS to 54 randomly selected government-funded school kitchens in Bihar, India. The DFS was subsequently used to prepare the MDM at the school level. The DFS has been developed by the National Institute of Nutrition (NIN), Hyderabad, and is produced by different manufacturers. TATA Chemicals supplied the DFS for our intervention. TATA Chemicals produces DFS under the name *TATA Salt Plus – Iron fortified iodized salt.* The TATA Salt Plus includes 0.86 mg/g of iron. The daily requirement of a child between 7 and 9 years of age is 10 mg while for children between 4 and 6 years it is 8 mg (WHO, 1959). By matching the daily amount of salt used as reported by the cook and the number of children that go to the respective school we calculate an average amount of DFS per meal served of 4g in our study. This makes up for 34 – 43% of the daily iron requirement of a child.

In laboratory studies, good stability of the iron and iodine content of the NIN formula of DFS has been shown. In a few mostly small-scale experimental studies (where regular consumption was closely monitored) also acceptability and effectiveness in increasing hemoglobin levels were tested (Sivakumar et al. 2001, Nairl et al. 2013, Haas et al. 2014). Most of these studies found good acceptability and increases in hemoglobin levels at least for subgroups.

In our study area DFS was unavailable and mostly unknown throughout the study. Originally DFS was sold in larger Indian cities starting in April 2012. Due to little demand TATA however removed it from the market. As instructed by the Department of Women and Child Welfare, DFS is supposed to be made available in governmental food security programs

[19] In other regions, mostly in urban areas, the MDM is prepared centralized in large kitchens or provided by organizations of international assistance.
[20] Average exchange rate 2016. 1 Rs. equals 0.015 USD.

including the Integrated Child Development Program and the Midday Meal Scheme (Mudur, 2013). However, this instruction is not enforced and availability and usage of DFS is still very limited. TATA was planning to launch DFS at the beginning of 2017 in the study region. Yet, in the last four month of the treatment period DFS got available in Patna, the capital city of the state Bihar, around two to three hours drive by car from the study region and thus difficult to reach for the study population.

TATA directly supplied the DFS to our research team. Depending on need, DFS was delivered to the treatment schools every, or every second, month. Headmasters were encouraged to contact our team in case they ran out of DFS stock before the next regular delivery date. The regular school visits also functioned as a monitoring system (see section 3.6.1 for a detailed discussion of compliance with the treatment). In order to mimic a real public policy program and to account for the potential of a budgetary effect, DFS was sold to headmasters to a subsidized price of Rs. 12 (0.18 USD)[21] [22], which is close to the price of local salt.[23] At endline headmasters and cooks stated that they did not have any major difficulties with the delivery and usage of the DFS. Twenty-two headmasters and 23 cooks (about 40% of treatment schools) mentioned a change in color due to the DFS (iron can make the food slightly bluish), which however did not seem to constitute a problem, despite that sometimes more turmeric was added. There was one isolated case where a school had problems in storage of DFS. More than half of the headmasters in treatment schools and also about 40% of the cooks attributed a positive change in the taste of the food to the DFS. They anecdotally reported that the taste of the DFS is more intense than of the conventional salt. Headmasters, but not parents and children, knew about the interventions.[24]

3.2.3 Theory of Change

Given that DFS is used regularly and sufficiently for cooking the MDM (an assumption that is discussed in section 3.6.1) and that children regularly attend school (an assumption that will be discussed in section 3.5.2), we expected that children's average hemoglobin value will

[21] Average exchange rate 2016. 1 Rs. equals 0.015 USD.

[22] The initial market price of TATA Salt Plus was 20 Rs. (0.3 USD) for one kg, but it was increased to 25 Rs. (0.38 USD) recently.

[23] Initially we sold the DFS for Rs. 15, which was equivalent to the cost of TATA iodized salt, but a little higher than the cost of other salt brands. After a few months of DFS supply, headmasters complained that the costs are too high and we reduced the price to 12 Rs.

[24] The usage of DFS in the MDM had already been instructed by an official note of the Central Government (Mudur, 2013). By sending their children to government-funded schools, parents allow them to participate in the MDM and hence agree to the consumption of DFS as well.

increase and that the anemia prevalence will be reduced. Related to this we expected a reduction in the number of anemia symptoms as reported by the child's mother. If positive effect on hemoglobin and anemia could indeed be found it would give us some reason to believe that cognitive skills and education outcomes might be positively affected too. Samuels, Wierman, Wang, & Ridgway 1989). Iron deficiency affects cognitive development through immediate neurobiological processes, i.e. the inhibition of the central nervous system to develop properly (e.g. the brain and the spinal cord) (Beard 2003), and secondly through functional isolation. Functional isolation emerges from the symptoms of iron deficiency. Children deficient in iron engage less with their environment, have lower interpersonal interactions, show lower attention and are relatively unresponsive to stimuli in comparison to their non-iron deficient counterparts. Hence, they have difficulties in accumulating new skills (Lozoff et al. 1998). Cognitive development directly influences the education outcomes of a child. For example, if a student is unable to focus their attention and to ignore distraction, they are likely to have trouble concentrating and hence have difficulties in acquiring new skills and knowledge. The same is likely to be true for the symptoms of anemia such as frequent illness or tiredness (Halterman et al. 2001, Bobonis et al. 2006). However, it should be noted that a treatment period of one year might be too short to be reflected in these more indirect outcomes like cognition and education.

3.3 Study Design and Data

3.3.1 Sample and Randomization

The study took place in the two blocks Kako and Modanganj in the district Jehanabad, located in the state Bihar, India. From a list of 228 government-funded schools that exists in the two blocks, a simple random sample of 108 schools was drawn prior to the DFS intervention. From the 108 schools, 54 were randomly chosen and allocated to the treatment group. The other 53 schools did not receive any treatment but continued using the conventional iodized salt.[25] For the choice of schools, as well as the allocation of the schools to the treatment and control group, a computer- generated list of random numbers was used.

On average, 20 children from the second grade were randomly selected from each of the 108 schools for the survey, which results in an initial overall sample size of about 2000

[25] Initially we had 54 control schools but one control school could not be reached due to monsoon flooding during the endline.

children.[26] Second graders were chosen because of the strong biological basis for post-infancy effects of iron deficiency on the neurobiological development of the brain. More specifically, the frontal lobes continue to develop until adolescence and experience spurts of development between the age of 7 and 9 years of life and in mid-teenage years (Anderson 2002, Hudspeth & Pribram 1990, Thatcher 1991). Among other functions, the frontal lobes are known to mediate advanced interrelated cognitive skills. These involve abilities, such as response inhibition, task switching, planning and organizing, working memory, abstraction, initiation, self-monitoring and volition, which can be summarized as executive functions (Anderson 2001, Lezak 1995, Salimpoor and Desrocher 2006).[27]

3.3.2 Data

Between November 2014 and January 2015, a baseline survey was carried out. Treatment started in August 2015.[28] An endline survey, where the same children were followed up, was conducted from August until October 2016. Both at endline and baseline three different teams collected data, mostly on different survey days. First the household team went to the home of the children to get parental consent for participation in the study and for a parental interview. Another team went to the schools and performed cognitive and education tests with the children. Headmasters and cooks were interviewed, too. The cooking and distribution of the MDM was also observed. The third team consisted of local medical staff that performed medical tests, including a blood test for hemoglobin, with the children at their home or in the villages in order to give the parents the possibility to attend the procedure.

Outcome variables

Anemia – As an immediate outcome of the intervention, the hemoglobin level of each child was assessed by a minor blood test. The portable HemoCue® Hb 301 device (AB Leo

[26] To determine if the chosen sample size was adequate, we calculated the minimal detectable effects (MDE) for the different outcomes. We assumed a significance level of 5%, a final sample of 1600 children and used the standard deviations and inter-cluster correlations from the baseline dataset. The minimal detectable effects for different outcomes are presented in the appendix table 3.A.1. At a treatment take-up of 80%, we would be able to detect a increase in hemoglobin of 0.2 g/dl or higher and a reduction in any form of anemia of 8.7 percentage points. For the cognitive outcomes, we would be able to detect effects with sizes of about 0.2 to 0.3 standard deviations.

[27] There is little consent on the exact definition of executive functions and there is an ongoing debate within neuropsychology about the dimensions of executive functions. For instance, other authors also classify focused and sustained attention, generation and implementation of strategies, monitoring, and utilization of feedback as executive functions (e.g. Stuss & Benson 1987).

[28] Implementation of the treatment was delayed because of the earthquake in Nepal 2015 that also affected Bihar and led to a postponement of school holidays. It was further delayed because of a contract teacher strike, which led to the circumstance that many schools were not functioning for several months.

Diagnostics, Helsinborg, Sweden) allows measurement directly in the field. In both the baseline and endline surveys, less than 1% of the children refused the hemoglobin test. Instead, the demand for the medical check was very high since parents perceived the health survey as free health service. The fact that medical staff came from the region might have enforced this perception. Following the official cutoff levels for different levels of severity of anemia for the age group 5 to 11 years (WHO, 2001), we define any anemia as a hemoglobin value < 11.5 g/dl, mild anemia as a hemoglobin value ≥ 11 & < 11.5 g/dl, moderate anemia as a hemoglobin value ≥ 8 & < 11 g/dl and severe anemia as a hemoglobin value < 8 g/dl. Since we only have 14 observations of severe anemia at the baseline and two at the endline, we collapse moderate and severe anemia to one category of moderate or severe anemia. We further collected data on the presence of common anemia symptoms in the last 30 days (easily fatigued and loss of energy, shortness of breath particularly with exercise, bluish color of lips and nails, headache, and shaking chills) as reported by the child's mother and we sum them up to a score from 0 to 5.

Cognitive ability – Cognitive ability was measured by five different cognitive tests: forward digit-span, backward digit-span, block design, *Stroop*-like day-and-night test and Raven's Colored Progressive Matrices. They were chosen according to the following criteria. First, apart from minor cognitive functions, these tests also require a range of *executive functions*, which are supposed to especially develop during our intervention period for the selected sample (see section 3.3.1). Second, they are frequently used measures of cognitive ability for child populations, e.g. the forward and backward digit-span tests and the block design tests are tests from Malin's Intelligence Scale for Indian Children (Malins, 1969), the Indian adaption of the Wechsler Intelligence Scale for children (WISC). Third, they are suitable for our study population, children between 6 and 9 in rural areas of Bihar. Fifth, it had to be practical to administer them in the field. The latter two criteria were verified by consulting Indian psychologists and in a pretest. To ensure similar text conditions, female prospective teachers administered all tests one by one in the schools. To minimize disturbance, the cognitive tests were performed in a separated room whenever possible. On average, the sequence of cognitive and education tests lasted for about 15 minutes. See table 3.1 for the characteristics of those tests at a glance. In the appendix figure 3.A.1 the material that was used for the cognitive tests can be found.

Table 3.1: Cognitive tests

Test	Original Source	Cognitive ability	Executive function
Digit span forwards	Malin's Intelligence Scale for Indian Children	- Short-term auditory memory - Simple verbal expression	
Digit span backwards	Malin's Intelligence Scale for Indian Children	- Store, use and manipulate new information - Attention - Impulse control - Shifting	x
Block design	Malin's Intelligence Scale for Indian Children	- Planning and organizing	x
Stroop-like day-and-night test	Gerstadt et al. (1994)	- Inhibition - Memorizing two rules simultaneously	x
Raven's Colored Progressive Matrices	Raven et al. (1998)	- Abstract reasoning - Capacity to simultaneously solve several problems involving new information	x

The *forward digit-span* assesses short-term auditory memory and simple verbal expression (Hale, Hoeppner and Fiorello, 2002). It is the only test that does not directly involve executive functions. The child was asked to repeat a series of numbers immediately after the enumerator red the series out. The number of digits was continuously increased and the longest list of numbers the child could remember was defined as the digit-span.

The *backward digit-span* test measures the ability to store, use and manipulate new information. Backward digit-span also involves attention, impulse control and shifting from a forward to backward sequence. All of these abilities are commonly considered to be a part of the group of executive functions (Carlson 2005, Hale et al. 2002, Lezak 1995). The administration of the backward digit-span test is the same as in forward digit-span test; however, the child is requested to repeat the digit sequence in its reverse order.

The *Block design* test assesses planning and organizing (Anderson 2001). In this test, children were asked to arrange red and white colored blocks in a way that they match a pattern of a picture. Children received two points if they correctly arranged the blocks on the first try, one point, if they correctly arranged the blocks on the second try, after the test administrator had shown the correct solution to the child, and zero otherwise. They were asked to arrange four different pictures in the baseline survey that increased in difficulty. To account for a general increase in cognitive ability at the endline survey, two more pictures were added.

The *Stroop-like day-and-night* test (Gerstadt, Hong and Diamond, 1994) assesses the ability of inhibition (suppressing a habitual response), which is also considered to be a classical executive function (Anderson 2001, Carlson 2005). We used an extended version of this test, where six pairs of cards, which show pictures of opposites (day–night, boy–girl, large–small, up–down, warm-cold and young-old), were presented to the child. After shuffling the pictures, they were presented to the child one after the other and the child was asked to say the opposite of what they were seeing on the card. The scale for the day-and-night test ranges from 0 to 12. Initial errors that were self-corrected by the child were scored as a half point. Apart from inhibition, this test also requires memorizing two rules simultaneously. First, what the picture on the cards represent and second, to always say the opposite.

Last, we used an abbreviated version of *Raven's Colored Progressive Matrices* (RCPM) (Raven, Raven and Court, 1998) that measures abstract reasoning and the capacity to simultaneously solve several problems involving new information (Carpenter, Just and Shell, 1990). There is some evidence that the RCPM test measures executive functions (Ardila et al. 2005, Giovagnoli 2001).[29] We used six matrices from set A and six matrices from set A_B. In this test, the child was shown an array of pictures with one missing box. Out of the six options, they were to select the picture that fits the missing box. The matrices progressively increased in complexity and abstraction. We score each correct answer with one point, hence the scale for RCPM ranges from 0 to 12. The RCPM are designed for children between 5 and 11 years old.

Since the unit in which cognitive ability is measured is arbitrary, we normalized the test scores of the five cognitive tests. We do this by subtracting the baseline mean of the control group and dividing by the baseline standard deviation of the control group of the given test, both for baseline and endline data. Hence, a normalized cognitive score of 0.5 would mean that the student scored 0.5 standard deviations higher than the mean in the baseline control group (irrespective of the sampling wave in which the score was obtained).

Lastly, we calculate a *cognitive index* out of the five cognitive tests, by using principle component analysis.[30] The advantage of an index is that it reduces the complexity of

[29] In other studies this hypothesis could not be confirmed (Leeds *et al.*, 2001; Blake *et al.*, 2002).

[30] Principal component analysis is a statistical data transformation procedure that reduces the dimension of data. It expresses the maximum amount of variance of a matrix with the fewest number of variables (principle components).

information. In our case, five different numbers for five different cognitive tests are reduced to one number. On the other hand, it reduces transparency. That is why we present the cognitive index in addition to the individual cognitive tests.

Education outcomes – We additionally assessed the children's math and reading skills and recorded their school attendance. For the assessment of the reading skills, we used the Annual Status of Education Report (ASER, 2014) test material developed by the Indian Non-Governmental Organization *Pratham*. Reading skills were assessed on a scale from 0 to 4, ranging from *child does not recognize letters* to *child fluently reads a short story*.

For the math assessment, we also used the ASER math test as basis, but extended it to 13 different math exercises at baseline and 15 at endline with increasing difficulty ranging from *child does not recognizes one-digit number* to *child solves advanced division problems*. Again, because the unit in which the education outcomes are measured is arbitrary, we normalize the test scores using the same procedure described above. The material that was used for the education tests can be found in the appendix figure 3.A.2.

The monthly school attendance was copied from the school's attendance registers. We recorded monthly attendance for the last 12 months before the survey; for the baseline, this was from November 2013 until October 2014 and from August 2015 until July 2016 for the endline. Due to many missing values in the attendance registers at the baseline from November 2013 until March 2014, we only included monthly attendance from April through October 2014 to calculate the baseline attendance rate. The school attendance is calculated by summing up the total number of days the child was present divided by the total number of days the school was open in the above-mentioned periods.

Covariates - In accordance with the existing scientific literature[31], we collected covariates that should be associated with the tested outcomes. For the anemia outcomes this includes socioeconomic characteristics (Rural or urban, block, wealth index[32], father's years of schooling, mother's years of schooling, caste, religion and number of household member), nutritional factors (the children's dietary diversity score, an indicator for household food

[31] For a more detailed description of the covariates included and the literature the selection is based on, see the appendix text 3.A.1 and table 3.A.2.

[32] The asset index was constructed by principle component analysis. The following variables were included: Type of toilet facility (improved and unimproved according to WHO), Source of drinking water (improved or unimproved), type of house, wall, roof and floor, possession of assets like chair, table, radio, pressure cooker etc., the amount of agricultural land owned, amount of different farm animals owned, BPL card holder, MNREGA card holder and the dependency ratio.

security, the number of meals the child eats every day, the average intake of calories and iron from the school meal, an indicator for maternal health knowledge, if the child consumes any meat, poultry or fish and if the child received iron supplements), access to health care (dummy for institutional delivery of the child and if any household member is covered by health insurance), morbidity indicators (if the child suffered from diarrhea in the last 30 days and if the household possesses an improved sanitation facility) as well as one biological factor (sex of the child).

For the cognitive outcomes, we include the same covariates, but another set of nutritional factors (the children's dietary diversity score, an indicator for household food security, the number of meals the child eats every day, the average intake of calories and iron from the school meal, an indicator for maternal health knowledge). In addition, we use indicators for psychosocial stimuli (dummy if the mother helps the child with its homework, the time the mother spends on giving physical care to the child, if parents participate in parent-teacher meetings at school and a dummy if the father lives in the household) and a dummy for the test administrator was included. Further indicators for quality of schooling (total school enrollment, student-teacher ratio, the number of children that attended second-grade at the baseline and the fourth-grade at the endline on the day of the interview and the distance to school) were included as controls for the education outcomes.

3.3.3 Needs Assessment

At the baseline, the average child exhibited a hemoglobin value of 11.53 g/dl, which is just above the cut-off point for mild anemia in this age group (hemoglobin < 11.5 g/dl). 45% of all children in our sample suffered from any form of anemia (hemoglobin < 11.5 g/dl), out of which 19% were mildly anemic (hemoglobin ≥ 11 & < 11.5 g/dl), 25% were moderately anemic (hemoglobin ≥ 8 & < 11 g/dl) and 0.1% were suffering from severe anemia (hemoglobin < 8 g/dl). According to maternal reports, a child, on average, showed 1.08 out of the 5 symptoms for anemia in the last 30 days.

Also in education outcomes, the second grade children performed quite poorly. In the math test, children scored 3.34 on average, which means they were not able to do simple one-digit calculations. 13% were not able to recognize numbers from 1-9, slightly below the national average of 20%. 39% of the students in our sample were unable to recognize numbers till 100. Though this is equivalent to the ASER, which reports this inability for 40%

of the second grade students, it is still very high. At the baseline, 49% of the children in our sample were not able to recognize single letters compared to only 32.5% of second grade students in the ASER sample. The ASER reports an average national attendance rate of 71%, as measured by the attendance at the day of the school visit. The equivalent number in our sample is 78%, hence close but a bit above the national average (Pratham, 2015).[33]

An assessment of baseline cognitive ability is rather difficult since the test administration and scoring varies strongly across studies and numbers are not available at a national level. The only test that is possibly administered in a similar fashion across studies, and for which we found numbers for a population in roughly the same age group as our sample, are the forward digit-span and the backward digit-span. On average, children in our sample could memorize and correctly repeat 4.06 digits in the forward sequence and in the reverse order they could repeat 1.12 numbers. Orsini et al. (1987) reported a forward digit-span of about 4.5 for 7 years old Italian children. For 6 year old children, hence children that are about one year younger than the average in our baseline sample, Chen and Stevenson (1988) found a forward digit-span of 4.5 for American children and of 5.33 for Chinese children. For the backward digit-span test, they found an average score of 2.83 for American and of 2.95 for Chinese children. Banerjee et al. (2016) report much lower numbers for children aged 5 to 7 years, so a slightly different age group, in one district of Bihar. They found a forward digit-span of 2.08 for females and of 2.3 for males. For the backward digit-span the equivalent numbers in the assessment of Banerjee et al. (2016) was 0.34 and 0.48. Children in our sample are between these numbers.

Though a needs assessment is difficult to establish for some of our outcomes of interest, these numbers still show that there was a clear need for an intervention at the baseline.

3.4 Estimation Strategy

3.4.1 Empirical Specification

Due to the random assignment of schools to the treatment and control groups, the causal effect of the DFS intervention can be estimated by comparing the averages of the above mentioned outcome variables between the treatment and control schools after the treatment

[33] For our analysis, we use another count of school attendance, namely the attendance reported in the school records for the year before the school visit.

was given (simple difference, SD). We estimate the SD in means by the following OLS equation

$$y_{ai} = \alpha + \beta\ Treat_a + \eta_{ai} \quad (1)$$

where y_{ai} is an outcome for child *i* at school *a* at the endline. α is the intercept. $Treat_a$ is a dummy, which takes the value of one if a child goes to a treatment school. β is the intent-to-treat ITT effect (ITT), the main coefficient of interest. Standard errors are clustered at the school level. If the outcome variable constitutes a dummy variable (any anemia, mild anemia, moderate and severe anemia), we estimate a linear probability model of the same form as regression (1).

The panel structure of our dataset further allows us to combine the randomized design with a difference-in-differences approach (DD). This approach controls for any remaining differences in outcomes variables between treatment and control group at baseline. We also include child fixed effects and additional time-variant covariates (time-invariant covariates are differentiated out by the child fixed effects). If those covariates are correlated with the outcome, this will increase the precision of our estimates, since the share of unexplained variation in the outcome will be reduced. We estimate the DD as an OLS regression in the form:

$$y_{ait} = \alpha_i + \beta Post_t + \gamma Post_t * Treat_a + X'_{at}\,\delta + W'_{it}\varepsilon + \eta_{ait} \quad (2)$$

where y_{ait} is an outcome for child *i* at school *a* at time *t*. α_i is the child specific intercept. $Post_t$ is a dummy variable, which takes on the value of one for the post-treatment time period and hence captures the time trend. $Treat_a$ is a dummy for going to a treatment school. γ is the ITT effect, the main coefficient of interest. It represents the coefficient for the interaction of $Post_t$ and $Treat_a$, which is equal to one for all observations in the post-treatment period that are in the treatment group.[34] X'_{at} is a vector of time-variant school control variables and W'_{it} is a vector of time-variant child and household control variables. η_{ait} constitutes the error term. Standard errors are clustered at the school level. If the outcome variable constitutes a dummy variable (any anemia, mild anemia, moderate anemia or severe anemia), we estimate a linear probability model of the same form as regression (2).

[34] In the DD estimation $Treat_a$ is differenced by the child fixed effects.

Regression (2) requires a balanced panel such that the data requirements are much higher and the number of children included in the analysis is smaller compared to regression (1). We constructed one balanced panel for the anemia outcomes and one balanced panel for the cognitive and education outcomes. Initially we sampled 2005 children. For the balanced sample of the anemia outcomes, for which medical and household data is required, we had 1791 observations at the baseline. Out of these 1791 children, we collected hemoglobin values for 1406 children at the endline and the required covariates from the household questionnaire (attrition is analyzed in section 3.4.3). Similarly, at the baseline we had 1772 children with all of the required cognitive and education outcomes as well as household covariates. Out of these 1772 children, the cognitive and education data, as well as the required covariates, were collected for 1395 children at the endline. See figure 3.1 for an overview of the sample attrition.

Figure 3.1: Sample deduction for the DD estimates

Cognitive and education sample	Cognitive Data baseline	+ HH data baseline	Missing on cognitive outcome	Missing on covariates	Full dataset at baseline	Children for which the full dataset was not collected at endline (=attrited children)	Final cognitive and education sample
Household data baseline 2008	1917	1914	- 8	- 134	1772	377	1395
	1924	1921	-1	- 129	1791	385	1406
Anemia sample	Medical data baseline	+ HH data baseline	Missing on anemia outcomes baseline	Missing on covariates	Full dataset at baseline	Children for which the full dataset was not collected at endline (=attrited children)	Final anemia sample

Regressions (1) and (2) estimate the ITT, i.e. the effect on those children that went to schools that were assigned to the treatment at the baseline but not the effect on those children that eventually received the DFS, which would be the average treatment effect (ATE). One might think of cases in which, for some reason, the DFS was not always or insufficiently used at the schools, a situation that has very unlikely occurred, as we discuss in section 3.6.1. We measure the ITT and not the ATE because there were children that did not always attend school and hence only partially complied with the treatment. The treatment effect, which includes children with varying levels of school attendance, and hence different levels of compliance with the treatment, is analyzed in section 3.5.2.

3.4.2 Pre-intervention Balance of the Treatment and Control Group

To obtain unbiased estimates of the ITT effect, it must be ensured that going to a treatment school is not correlated with other characteristics that affect the outcomes of interest. Though this condition might not be fulfilled in observational data, the random assignment of schools to the treatment and control groups in this study, makes treatment assignment independent of the potential outcomes. A balancing test, i.e. a comparison of the characteristics of the treatment and control groups before the intervention, gives an indication if the randomization process was successful and if the condition is fulfilled.[35] For the anemia outcomes, we present such a balancing test in panel A in table 3.2 for the observations included in the SD sample for which we have data from the baseline (out of the 1724 observations for which we have endline hemoglobin values we have 1686 observations at the baseline) and in panel A in table 3.3 for the DD balanced panel. The equivalent tables for the cognitive and education outcomes are found in panel A in the appendix tables 3.A.3 and 3.A.4. In the balancing tables, we show the means for all of the outcomes variables and covariates as well as the standard deviation for the continuous variables. We also present the p-values for the difference in means t-tests. Standard errors are clustered at the school level.

Tables 3.2 and 3.3 show that the means of the treatment and control groups for both specifications are very close to each other and that the p-values of the t-test for the difference in the means are mostly not statistically significant. Numbers for the cognitive and education sample (appendix tables 3.A.3 and 3.A.4) are very similar. This strongly indicates that the randomization worked out and that the two groups that have been created are, on average, equal in their observable characteristics. This observation gives us good reason to believe that unobservable characteristics are also balanced between the two groups such that the treatment and control groups only differ with respect to the intervention but not to any other characteristics.

The baseline hemoglobin level is, however, not perfectly equalized between the treatment and control groups. In all samples, the baseline hemoglobin level was a bit higher in the control group and hence the prevalence of anemia was lower. If not taken into account, this would lead to an underestimation of the true treatment effect. We hence control for the

[35] The balance in means is only true on average, meaning if the experiment was conducted many times. It is still likely that due to random chance and especially if the sample size is small (as it is the case for the school level indicators in this study) some means show up to be different between control and treatment group (Altman, 1985; Imai, King and Stuart, 2008).

baseline hemoglobin level (as well as the other factors mentioned in section 3.3.2) in the DD estimations.

Table 3.2: Sample description and balancing test (SD sample for hemoglobin and anemia outcomes)

	Panel A: SD sample								Panel B: Sample including attrited children							
	Total			Control		Treatment			Total			Control		Treatment		
	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test
A. Child level variables																
Anemia																
Hemoglobin	1,670	11.53	1.10	11.62	1.11	11.45	1.08	0.010**	1,918	11.52	1.10	11.60	1.09	11.44	1.11	0.013**
Any anemia	1,670	0.45		0.42		0.48		0.030**	1,918	0.45		0.42		0.48		0.036**
Mild anemia	1,670	0.19		0.17		0.20		0.140	1,918	0.19		0.18		0.20		0.194
Moderate or severe anemia	1,670	0.26		0.24		0.27		0.207	1,918	0.26		0.25		0.28		0.154
Number of anemia symptoms	1,723	1.08	1.11	1.06	1.10	1.11	1.12	0.560	2,005	1.07	1.10	1.06	1.09	1.09	1.11	0.676
Cognitive tests																
Block design	1,669	3.67	2.21	3.61	2.23	3.73	2.19	0.569	1,910	3.73	2.22	3.68	2.24	3.78	2.20	0.626
Digit span forward	1,671	4.05	1.00	4.05	1.01	4.06	0.98	0.859	1,912	4.08	1.00	4.07	1.02	4.08	0.99	0.886
Digit span backward	1,671	1.11	1.29	1.08	1.28	1.13	1.30	0.649	1,912	1.13	1.30	1.11	1.30	1.15	1.31	0.695
Progressive matrices	1,670	4.74	1.65	4.81	1.62	4.67	1.68	0.328	1,910	4.76	1.69	4.84	1.65	4.69	1.72	0.279
Day and night Cognitive score (pca)	1,669	5.30	3.42	5.37	3.46	5.23	3.38	0.583	1,909	5.39	3.43	5.48	3.48	5.30	3.39	0.480
Education outcomes																
Math	1,671	4.70	3.83	4.77	3.87	4.62	3.79	0.710	1,912	4.85	3.85	4.93	3.89	4.77	3.82	0.680
Reading	1,670	0.88	1.12	0.91	1.14	0.85	1.10	0.557	1,911	0.92	1.14	0.95	1.16	0.89	1.12	0.507
School attendance	1,630	0.79	0.16	0.80	0.16	0.79	0.17	0.551	1,887	0.78	0.17	0.79	0.17	0.78	0.17	0.658
Socioeconomic variables																
Muslim HH	1,723	0.03		0.03		0.03		0.936	2,005	0.03		0.02		0.03		0.855
Sc/st	1,723	0.29		0.26		0.33		0.218	2,005	0.28		0.24		0.31		0.220
Block	1,724	0.68		0.73		0.64		0.367	2,005	0.67		0.72		0.62		0.317
Rural HH	1,723	0.98		0.98		0.98		0.604	2,005	0.98		0.98		0.98		0.679
N of HH members	1,723	7.73	3.41	7.77	3.40	7.68	3.41	0.686	2,005	7.77	3.49	7.82	3.52	7.73	3.46	0.667
Years schooling	1,691	5.44	4.80	5.53	4.80	5.35	4.81	0.638	1,967	5.58	4.88	5.49	4.86	5.65	4.90	0.668

father																
Years schooling mother	1,718	1.69	3.14	1.74	3.20	1.64	3.07	0.627	1,999	1.89	3.35	1.89	3.34	1.90	3.36	0.967
Asset index	1,692	-0.04	0.95	-0.02	0.93	-0.05	0.97	0.713	1,965	0.00	1.00	0.01	0.99	-0.01	1.01	0.799
Health care																
Institutional delivery	1,714	0.38		0.39		0.37		0.585	1,994	0.39		0.40		0.38		0.585
Health insurance	1,707	0.39		0.40		0.38		0.720	1,985	0.39		0.40		0.38		0.740
Morbidity																
Diarrhea	1,723	0.03		0.03		0.03		0.502	2,005	0.03		0.03		0.03		0.942
Improved sanitation	1,723	0.08		0.07		0.08		0.616	2,005	0.09		0.08		0.10		0.458
Biological factor																
Male child	1,724	0.45		0.43		0.47		0.187	2,005	0.46		0.44		0.47		0.306
Nutrition																
Diet diversity score	1,723	3.86	1.17	3.89	1.20	3.83	1.14	0.513	2,005	3.88	1.17	3.90	1.20	3.86	1.14	0.646
Number of meals	1,722	3.05	1.02	3.05	1.03	3.06	1.01	0.886	2,004	3.04	1.04	3.02	1.05	3.06	1.03	0.574
Cut meals	1,721	0.80		0.80		0.80		0.991	2,001	0.79		0.79		0.79		0.908
Maternal health knowledge	1,723	0.37		0.35		0.40		0.110	2,005	0.39		0.35		0.41		0.062*
Child eat meat, poultry or fish	1,723	0.53		0.54		0.52		0.577	2,005	0.53		0.54		0.53		0.821
Child got iron supplements	1,697	0.17		0.18		0.17		0.756	1,975	0.17		0.18		0.16		0.585
B. School level variables																
Calories of MDM per child	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739	107	222.21	158.24	222.15	168.29	222.28	149.30	0.997
Amount of iron of MDM per child	107	0.79	0.35	0.84	0.41	0.75	0.28	0.185	107	28.19	16.87	28.87	20.41	27.52	12.63	0.681

This table presents the baseline summary statistics as well as p-values for differences in the means t-tests between children in the treatment and control schools. All variables shown are child level variables from the baseline except for panel B, which shows school level variables. Standard errors are clustered at the school level. Sd: Standard deviation N: Number of observations. MDM: Midday Meal. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 3.3: Sample description and balancing test (DD balanced panel for hemoglobin and anemia outcomes)

	Panel A: DD sample								Panel B: Sample including attrited children							
	Total			Control		Treatment			Total			Control		Treatment		
	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test
A. Child level variables																
Anemia																
Hemoglobin	1,406	11.53	1.09	11.62	1.11	11.44	1.07	0.008***	1,791	11.52	1.10	11.59	1.10	11.45	1.10	0.020**
Any anemia	1,406	0.45		0.41		0.50		0.002***	1,791	0.45		0.42		0.49		0.022**
Mild anemia	1,406	0.19		0.16		0.22		0.003***	1,791	0.19		0.18		0.21		0.089*
Moderate or severe anemia	1,406	0.26		0.25		0.27		0.321	1,791	0.26		0.25		0.28		0.186
Number of anemia symptoms	1,406	1.09	1.12	1.07	1.10	1.11	1.14	0.650	1,791	1.08	1.11	1.05	1.09	1.10	1.12	0.609
Cognitive tests																
Block design	1,375	3.61	2.20	3.53	2.22	3.69	2.19	0.450	1,734	3.73	2.23	3.68	2.26	3.79	2.19	0.591
Digit span forward	1,377	4.05	1.01	4.03	1.01	4.07	1.00	0.611	1,736	4.08	1.01	4.05	1.02	4.10	1.00	0.564
Digit span backward	1,377	1.12	1.29	1.07	1.27	1.16	1.30	0.460	1,736	1.12	1.29	1.09	1.29	1.15	1.30	0.559
Progressive matrices	1,376	4.71	1.67	4.76	1.64	4.67	1.69	0.494	1,734	4.76	1.69	4.82	1.67	4.71	1.71	0.440
Day and night	1,375	5.22	3.44	5.25	3.49	5.19	3.40	0.812	1,733	5.33	3.43	5.39	3.48	5.27	3.38	0.662
Cognitive score (pca)	1,372	-0.05	1.00	-0.06	1.01	-0.03	1.00	0.727	1,729	-0.00	1.00	-0.01	1.02	0.00	0.98	0.885
Education outcomes																
Math	1,377	4.60	3.81	4.66	3.85	4.53	3.77	0.772	1,736	4.82	3.83	4.90	3.86	4.74	3.81	0.694
Reading	1,376	0.86	1.11	0.87	1.12	0.85	1.10	0.868	1,735	0.90	1.13	0.94	1.15	0.87	1.10	0.490
School attendance	1,338	0.80	0.16	0.80	0.15	0.79	0.16	0.438	1,694	0.79	0.16	0.79	0.16	0.78	0.17	0.659
Socioeconomic variables																
Muslim HH	1,406	0.02		0.02		0.03		0.886	1,791	0.03		0.03		0.03		0.898
Sc/st	1,406	0.30		0.25		0.34		0.124	1,791	0.29		0.25		0.32		0.167
Block	1,406	0.66		0.71		0.62		0.383	1,791	0.66		0.71		0.62		0.415
Rural HH	1,406	0.98		0.97		0.98		0.590	1,791	0.98		0.97		0.99		0.412
N of HH members	1,406	7.73	3.38	7.85	3.47	7.62	3.29	0.310	1,791	7.75	3.44	7.84	3.53	7.66	3.35	0.408
Years schooling	1,406	5.38	4.78	5.48	4.76	5.29	4.81	0.628	1,791	5.49	4.87	5.42	4.85	5.56	4.89	0.708

father															
Years schooling mother	1,406	1.64	3.05	1.70	3.15	1.58	2.94	0.600	1,791	1.80	3.24	1.81	3.27	1.80	3.22	0.971
Asset index	1,406	0.00	1.00	0.01	0.95	-0.01	1.04	0.823	1,791	-0.03	0.97	-0.03	0.95	-0.02	1.00	0.908
Health care																
Institutional delivery	1,406	0.38		0.40		0.37		0.423	1,791	0.39		0.39		0.38		0.632
Health insurance	1,406	0.39		0.40		0.37		0.544	1,791	0.39		0.40		0.38		0.683
Morbidity																
Diarrhea	1,406	0.03		0.04		0.03		0.302	1,791	0.03		0.03		0.03		0.574
Improved sanitation	1,406	0.08		0.07		0.08		0.557	1,791	0.08		0.08		0.09		0.526
Biological factor																
Male child	1,406	0.45		0.44		0.46		0.430	1,791	0.46		0.45		0.47		0.390
Nutrition																
Diet diversity score	1,406	3.87	1.16	3.90	1.18	3.83	1.15	0.446	1,791	3.87	1.17	3.89	1.20	3.86	1.14	0.698
Number of meals	1,406	3.06	1.04	3.04	1.06	3.07	1.01	0.773	1,791	3.02	1.05	3.01	1.07	3.04	1.03	0.660
Cut meals	1,406	0.81		0.82		0.81		0.735	1,791	0.80		0.80		0.80		0.796
Maternal health knowledge	1,406	0.37		0.35		0.38		0.416	1,791	0.38		0.36		0.40		0.187
Child eat meat, poultry or fish	1,406	0.53		0.55		0.52		0.447	1,791	0.53		0.53		0.53		0.895
Child got iron supplements	1,406	0.17		0.16		0.17		0.625	1,791	0.17		0.17		0.17		0.931
B. School level variables																
Calories of MDM per child	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739
Amount of iron of MDM per child	107	0.79	0.35	0.84	0.41	0.75	0.28	0.185	107	0.79	0.35	0.84	0.41	0.75	0.28	0.185

This table presents baseline summary statistics as well as p-values for difference in the means t-tests between children in the treatment and control schools. All variables shown are child level variables from the baseline except for panel B, which shows school level variables. Standard errors are clustered at the school level. Sd: Standard deviation N: Number of observations. MDM: Midday Meal. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

3.4.3 Attrition

Attrition describes the loss of participants within a study. It can either emerge when participants fully drop out or when certain outcome or control variables are not collected for some reason. Attrition might threaten internal validity in a case where the characteristics of the participants that are lost systematically differ between the treatment and control groups and those characteristics are correlated with the outcome (so called differential attrition). Whereas a loss of participants that is unrelated to the treatment status, i.e. exogenous, is less of a concern, the systematic loss of observations leads to a selection bias and hence possibly biased impact estimates (Duflo, Glennerster and Kremer, 2007). For instance, in our setting it might have been possible that the iron treatment causes a systematic loss of children that were already at a high cognitive and education level at the baseline because the treatment increased the cognitive ability and education outcomes of such children in a way that many of them qualified for scholarships and moved to higher quality schools. In such a case, the baseline cognitive and education outcomes for the treatment group would have been systematically lower in the samples included in our regressions, compared to the full initial sample, i.e. our estimates would be biased downward. If too many observations are lost it might further challenge the power of the estimates and the representativeness of the results (Fewtrell *et al.*, 2008).

Generally, we reduced the number of lost participants by making all effort in the field to locate the children and their parents. Enumerators were assigned a smaller number of interviews per day such that they had sufficient time to extensively search for parents and children.[36] We returned to the village or schools a second time in case many children or parents were not present on the day of the survey. During the cognitive tests, which were performed at the schools, we searched for children in the village in case they were not present at the school on the day of the survey. Nevertheless, since the data was collected with three different survey teams that visited children and their parents on different days, we have several cases for which we do not have the full dataset (i.e. household, cognitive and education data for the cognitive sample and household and hemoglobin values for the anemia sample, both for the baseline and the endline surveys).

[36] Parental interviews took about 20 minutes and enumerators completed seven interviews per day, on average. Cognition and Education tests took about 15 minutes and enumerators were supposed to do eight to nine interviews per day.

We discuss attrition first for the SD estimates and then for the DD estimates. In our initial sample, 2005 children were listed for which we received parental consent and collected household data.[37] There were only isolated cases of households that did not agree to participate in the survey.[38] Out of these 2005 children, we collected hemoglobin values for 1724 children during the endline survey. These 1724 children are those that were included in the SD analysis for the hemoglobin and anemia outcomes. Hence the attrition rate from the initial sample to the endline was 14.0% (281/2005).[39] The equivalent attrition rate for the SD analysis for the cognitive and education outcomes is 13.2% (265/2005). These attrition rates are low in comparison to other RCTs (e.g. Banerjee et al. (2015) report an attrition rate of 10% between two follow-up surveys Glewwe et al. (2009) report 25% attrition after one year of an intervention and a bit more than 30%, two years after an intervention and Ashraf et al. (2014) report 26% to 28% attrition in their follow-up survey). Of the 14.0% of children that were lost in the anemia sample, 47.4% belonged to the control group and 52.7% to the treatment group; and of the 13.2% lost in the cognitive and education sample, 52% belonged to the control group and 48% belonged to the treatment group. To test if the treatment caused attrition at the extensive margin, we created an attrition dummy that takes on the value of one if the child's data on the respective outcome category was not collected at the endline. In a linear probability model, we regress the attrition dummy on treatment status. In table 3.4, we present the regression results for the different samples that we use in our specifications. None of the coefficients are found to be statistically significant and the point estimates are very small, indicating that the treatment did not cause attrition at the extensive margin.

[37] This maximum number of 2005 observations is only found for household covariates. For each school, the enumerator team received a list of all children that are registered in the second grade. They built a random ranking of the children on this list and visited the households of the first 20 children on the list for the parental interview and parental consent. As some children are registered at the school but no longer live in the village, missing children were replaced by the next child on the list until the parents of 20 children were interviewed. The 2005 children that were identified by this procedure build the basis for our sample and hence the basis for the medical, cognition and education team to collect data. In some schools, less than 20 children go to the second grade. In this case, we interviewed all children that go to the second grade but sampled more than 20 children in the next school to have an average of 20 children per school.

[38] Only in a village that belonged to one school, were households very suspicious and did not allow our team to enter the village, so we replaced this village by another randomly selected school.

[39] 1686 observations out of the 1724 endline observations are the same children as in the baseline (i.e. for which we can show the balancing test in table 3.2, Panel A). The remaining 38 children were found at endline but not at the baseline for hemoglobin testing.

Table 3.4: Attrition (extensive margin) for SD samples

	(1) Anemia sample	(2) N of anemia symptoms sample	(3) Cognitive and education sample	(4) Attendance sample
Treat	0.011	-0.024	-0.013	0.002
Standard error	0.018	0.023	0.016	0.010
P value	0.528	0.289	0.431	0.861
N	2,005	2,005	2,005	2,005

N: Number of observations. All estimates are based on OLS regression where treatment status is regressed on attrition. Standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Though attrition was equal in terms of the numbers of attrited children, it might still be the case that the lost participants differ in their characteristics causing an imbalance between the treatment and control groups. The balance of the baseline characteristics between the treatment and control groups of the samples used in the SD regressions has already been discussed in section 3.4.2. In the balancing test, attrited children are already excluded and as stated above, the characteristics are balanced between the treatment and control groups in these samples. Still, we compare the balancing tests without attrited children (Panel A, table 2) to the balancing test of the full initial sample where attrited children are included, i.e. all data that was collected at the baseline (Panel B, table 3.2). The balancing tests without attrited children and the balancing test with attrited children look nearly identical, indicating that differential attrition did not happen in the experiment. In the appendix table 3.A.3, we perform the same exercise for the cognitive and education SD sample and the picture looks very much the same.

Due to the higher data requirements, attrition is higher for the DD estimates. For the balanced panel of the anemia outcomes, we could collect the required data for 1791 children at the baseline and for 1406 children at the endline, resulting in 599 attrited observations from the initial and final samples or 29.9% (599/2005) and 385 attrited children between the baseline and endline surveys, which is equivalent to an attrition rate of 21.5% (385/1791), respectively. Though these attrition rates are much higher than for the SD estimates, they are still in the range of other RCTs (examples reported above). Attrition rates were very similar across the treatment and control groups. Out of the 29.9% of the children that were lost between the initial sample and the final sample, 49.4% were from the control group and 50.6% were from the treatment group. Of the 21.5% of the children lost between the baseline and endline surveys, 48% were from the treatment group and 52% from the control group. For

the balanced panel of cognitive and education outcomes we had collected the required data for 1772 children at the baseline and 1395 children at the endline, meaning 610 attrited children between the initial and final samples (610/2005 = 30.4%) and 377 attrited children between the baseline and endline (377/1772 = 21.8%), respectively. For the 30.4% of the attrited children between the initial sample and the sample included in the DD analysis, 48.4% were lost in the control and 51.6% in the treatment group. For the 21.8% of attrited children between the baseline and the endline, 46% were in the treatment group and 54% in the control group. To test if the treatment caused attrition, we regress the attrition dummy on the treatment status in a linear probability model. In table 3.5 we present the regression results for attrition from the initial sample to the endline sample and in table 3.6 for attrition from the baseline sample to the endline sample. None of the coefficients are found to be statistically significant at conventional levels and the point estimates are very small, showing that treatment did not cause attrition at the extensive margin in the DD estimates.

Table 3.5: Attrition (extensive margin) for DD samples (from the initial sample to endline)

	(1) Anemia sample	(2) N of anemia symptoms sample	(3) Cognitive and education sample	(4) Attendance sample
Treat	-0.031	-0.058*	-0.024	-0.047
Standard error	0.027	0.032	0.028	0.031
P value	0.254	0.073	0.384	0.138
N	2,005	2,005	2,005	2,005

N: Number of observations. All estimates are based on OLS regression where treatment status is regressed on attrition. Standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 3.6: Attrition (extensive margin) for DD samples (from baseline to endline)

	(1) Anemia sample	(2) N of anemia symptoms sample	(3) Cognitive and education sample	(4) Attendance sample
Treat	-0.024	-0.059*	-0.037	-0.042
Standard error	0.028	0.031	0.030	0.031
P value	0.394	0.061	0.219	0.174
N	1,791	1,869	1,772	1,713

N: Number of observations. All estimates are based on OLS regression where treatment status is regressed on attrition. Standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

How far attrition in the DD estimates imbalanced the samples used for our analysis is reflected in table 3.3. The balance of the characteristics between the treatment and control groups has already been discussed in section 3.4.3. In the balancing test, attrited children are already excluded and as stated above, the characteristics are balanced between the treatment and control groups in these samples. Still, we compare the balancing tests without attrited children (Panel A) to the balancing test of the full baseline sample (Panel B), as well as the full initial sample. The full initial sample has already been show in the SD attrition tables and can hence be found in Panel B of table 3.2. We therefore compare panel A to panel B in table 3.3 to analyze the attrition between the baseline and endline and Panel A in table 3.3 to Panel B in table 3.2 to analyze the attrition between the initial sample and the sample included in the DD estimates. The balancing tests without attrited children and the balancing test with attrited children look nearly identical, indicating that differential attrition did not happen in the experiment. In the appendix table 3.A.4, we perform the same exercise for the cognitive and education sample and the picture looks very much the same. We therefore conclude that differential attrition is not an issue in our study and that the sample used in the analysis is identical to the one that was initially sampled. This means that results of our analysis are still representative for students going to government-funded schools in the two blocks named above that were in the second grade at the baseline.

3.5 Results

3.5.1 Main Results

Table 3.7, 3.8 and 3.9 show the regression results of the ITT effect, based on specification (1) and (2).

Anemia – Looking at the SD specification we do not find any effect of DFS treatment on any of the anemia outcomes. As already mentioned in section 3.4.2, these results are biased due to statistically significant differences in baseline hemoglobin values between the treatment and control groups. The DD estimates account for this baseline imbalance and hence produce unbiased results. According to the DD estimated, the usage of DFS for the preparation of the MDM increased the average hemoglobin level by 0.136 g/dl (P-value: 0.076, SE: 0.076) (Table 3.7). However, this coefficient is only significant at the 10% level. The size of the effect found in our study is a bit smaller than the effect found in other DFS intervention studies. Haas et al. (2014) found an increase in hemoglobin of 0.24 g/dl among Indian tea

pickers. In a household level DFS intervention, Nairl et al. (2013) found an increase in hemoglobin due to a DFS intervention among Indian school-aged children of 0.5 g/dl. Zimmermann et al. (2003) found an increase of 1.4 g/dl in their DFS intervention in Moroccan school children; however their DFS formula contained 2 mg of iron per gramm of salt. For the subgroup of adolescents, Banerjee et al. (2016) found an increase in the average hemoglobin value of 0.41 g/dl.

With a statistical significance at the 1% level, a child going to a treatment school is, on average, 9.3 percentage points less likely to suffer from any form of anemia (hemoglobin < 11.5 g/dl) (P-value: 0.006, SE: 0.033). This means a reduction in prevalence of anemia of more than 20% (9.3/45.2) due to the treatment. The usage of DFS in the MDM reduced the likelihood that a child suffers from mild anemia (hemoglobin ≥ 11 & < 11.5 g/dl) by 6.0 percentage points, on average, which is equivalent to a reduction in the prevalence of mild anemia by nearly 30% (6.0/19.3). This effect is statistically significant at the 5% level (P-value: 0.032, SE: 0.027). Coefficients of moderate and severe anemia and the number of anemia symptoms go into the expected direction; however, they are not statistically significant. The reduction in the prevalence of anemia is similar to the reduction found in other studies. Haas et al. (2014) found that DFS decreases the prevalence of anemia among Indian tea pickers by about 25%. Nairl et al. (2013) report a reduction in anemia by about 20% and Banerjee et al. (2016) of about 26% for a subgroup of adolescents. Zimmermann et al. (2003) report a higher reduction in anemia of about 43% and in the multi-micronutrient-fortified biscuit school-intervention, van Stuijvenberg (2005) also found a decrease of the prevalence of anemia as high as 46%.

Table 3.7: ITT effects on hemoglobin level and anemia

	(1) Hemoglobin		(2) Any anemia		(3) Mild anemia		(4) Moderate or severe anemia		(5) N of anemia symptoms	
	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD
Treat	0.022	0.136*	-0.020	-0.093***	-0.001	-0.060**	-0.019	-0.034	-0.022	-0.137
Standard error	0.057	0.076	0.023	0.033	0.014	0.027	0.018	0.031	0.083	0.099
P value	0.705	0.076	0.394	0.006	0.936	0.032	0.311	0.286	0.787	0.171
Mean dependent var	11.522	11.529	0.453	0.452	0.190	0.193	0.263	0.260	1.074	1.091
Child fixed effects		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes
N	1,724	2,812	1,724	2,812	1,724	2,812	1,724	2,812	1,768	2,748

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. For the DD estimates we use a balanced sample for hemoglobin and anemia outcomes, except for (6) N of anemia symptoms, which is based on its own balanced sample as it only contains variables from the household dataset. Any anemia is defined as a hemoglobin value < 11.5 g/dl, mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dl, moderate/severe anemia is defined as a hemoglobin value < 11 g/dl. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered at the school level.

Table 3.8: ITT effects on cognitive tests

	(1) Block design		(2) Digit span forward		(3) Digit span backward		(4) Progressive matrices		(5) Day and night SD		(6) Cognitive Index	
	SD	DD	SD	DD	SD	DD	SD	DD		DD	SD	DD
Treat	0.068	0.012	-0.014	-0.105	0.135	0.009	0.000	0.070	0.098	0.116	0.086	0.028
Standard error	0.068	0.082	0.064	0.074	0.091	0.084	0.050	0.095	0.091	0.098	0.090	0.080
P value	0.318	0.880	0.828	0.159	0.142	0.920	0.992	0.464	0.284	0.241	0.340	0.725
Mean dependent var	0.024	0.048	0.008	0.030	0.016	0.033	-0.041	-0.020	-0.024	0.000	0.000	-0.038
Child fixed effects		Yes		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes		Yes
N	1,740	2,790	1,740	2,790	1,740	2,790	1,740	2,790	1,740	2,790	1,740	2,790

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. For the DD estimates we use a balanced sample for cognitive and education outcomes except for attendance, which is based on its own balanced sample as it only contains variables from the household and attendance datasets. All outcomes are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Standard errors are clustered at the school level.

Table 3.9: ITT effects on education

	(1) Math		(2) Reading		(3) School attendance	
	SD	DD	SD	DD	SD	DD
Treat	0.112	0.129	0.129	0.104	-0.006	-0.005
Standard error	0.117	0.090	0.120	0.081	0.023	0.022
P value	0.341	0.154	0.284	0.201	0.795	0.811
Mean dependent var	-0.019	-0.007	-0.026	-0.019	0.782	0.798
Child fixed effects		Yes		Yes		Yes
Controls		Yes		Yes		Yes
N	1,739	2,790	1,740	2,790	1,964	2,715

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. For the DD estimates we use a balanced sample for cognitive and education outcomes except for attendance, which is based on its own balanced sample as it only contains variables from the household and attendance datasets. All outcomes, except attendance, are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Standard errors are clustered at the school level.

Overall, the usage of DFS in the MDM resulted in a small increase in the hemoglobin values and a meaningful reduction in the prevalence of anemia. One reason why we only observe a significant reduction in mild, but not in moderate and severe anemia, might be due to the different sizes of the intervals and the distribution of the baseline hemoglobin values that is close to a normal distribution with a mean of 11.5 g/dl. Whereas mild anemia ranges from a hemoglobin value ≥ 11 to < 11.5 g/dl, and thus has a span of 0.5, moderate and severe anemia is defined as having hemoglobin values < 11 g/dl and hence has a span of more than 3.0. Due to the normal distribution the number of observations per discrete hemoglobin value is higher to the left of the mild anemia threshold than to the left of the moderate anemia threshold. Thus, if the treatment evenly improves hemoglobin values, then more children will be lifted over the threshold from mild to no anemia than from moderate or severe to mild anemia.

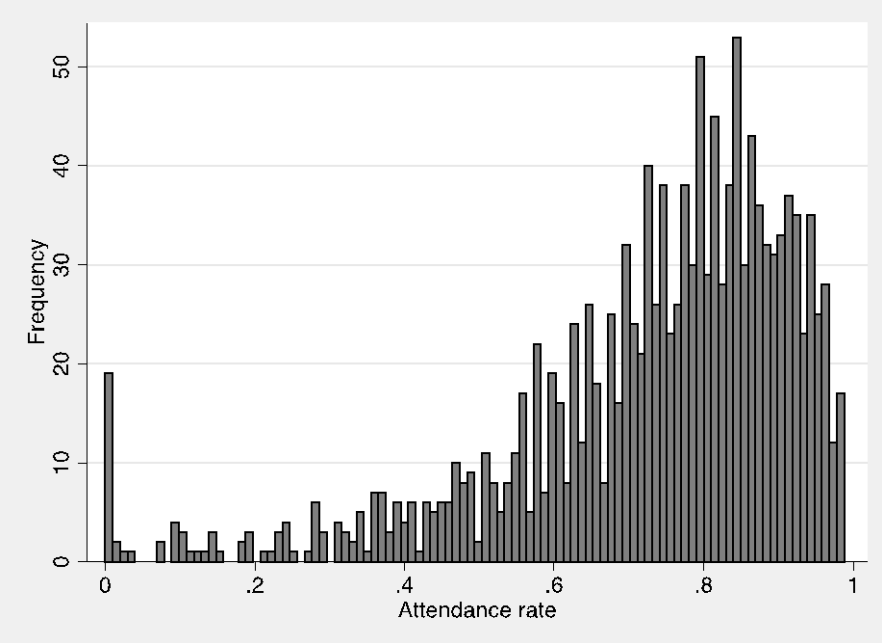
Cognition and education - Table 3.8 shows the estimation results for cognitive indicators and Table 3.9 for the education outcomes. SD and DD estimates are roughly similar in magnitudes, but none of the coefficients are statistically significant at conventional levels. Even though the treatment decreased the prevalence of anemia, on average the one-year intervention was not effective in improving the cognitive ability of second grade students. This is true even though cognitive indicators were chosen in a way that they test abilities that are supposed to experience spurts of development precisely at the age of the children in our sample. However, it might be the case that the effects of DFS needs more than one year to materialize in more indirect outcomes such as education and cognition. Another explanation of the absence of any effect on cognition and education might be that the ITT estimates include all children that were included in the baseline sample and that were found during the follow up survey, irrespective of their school attendance. Hence, these estimates include children with low or no school attendance, who therefore only partially complied with the treatment.

3.5.2 Heterogeneous Treatment Effects

School attendance and partial compliance – To get a better idea of how effective the treatment could potentially be with higher compliance, we investigate the treatment effect depending on school attendance. Figure 3.2 shows the distribution of school attendance during the treatment period. Though the reported average school attendance during the

treatment period was already relatively high, with 73% in the DD hemoglobin sample and 71% in the SD hemoglobin sample[40], we would expect a higher treatment effect at higher levels of school attendance.

Figure 3.2: Distribution of school attendance during treatment period (0 to 100%)



To investigate the treatment effect at higher rates of school attendance we use the variable *School attendance*$_i$, that runs from 0 to 1 and centered it at 70%, 80% and 90% respectively, i.e. we construct three new variables where we subtract 0.7, 0.8 and 0.9 respectively. We than interact each centered attendance rate once with the treatment dummy. For the SD we than run three times the following regression, once for each centered variable

$$y_{ai} = \alpha + \delta_1 \text{ School attendance}_{ai} + \delta_2 \text{ Treat}_a + \delta_3 \text{ Treat}_a \text{ x School attendance}_{ai} + \eta_{ai} \quad (3)$$

where abbreviations are the same as described below equation (1). We are only able to do so because as shown in Table 3.9, the treatment did not have any effect on school attendance. For the three different regressions, δ_2 gives the treatment effect for a child with 70%, 80%,

[40] The calculation of the school attendance rate is based on school records, a book where teachers note down the daily attendance of each child enrolled in the school. Since certain benefits for headmasters are attached to the attendance rate, for instance the amount of money they receive to buy the ingredients for the MDM, these registers are prone to over reporting. The true school attendance is hence likely to be lower than reported in those registers, which means that the estimated effects are likely to be lower than the effects of estimated from the true school attendance.

90% school attendance during the treatment period, respectively. δ_3 is same for all three estimations and shows the extra effect of a one unit increase in attendance, i.e. as school attendance runs for 0 to 1, a 100 percentage point increase in school attendance. The SD estimates in table 3.10-3.12 report the three coefficients δ_2 and once coefficient δ_3.

We perform the same exercise for the DD analysis. The regression equation is as follows

$$y_{ait} = \alpha_i + \delta_1 Post_t + \delta_2 School\ attendance_{ai} + \delta_3 Treat_a x\ Post_t + \delta_4 Post_t\ x\ School\ attendance_{ai} + \delta_5\ Treat_a\ x\ School\ attendance_i + \delta_6\ Post_t\ x\ Treat_a\ x\ School\ attendance_{ai} + X'_{at}\ \varepsilon + Y'_{it}\ \zeta + \eta_{ait} \quad (4)$$

where abbreviations are the same as describes below equation (2). We again center *School attendance*$_i$ at 70%, 80% and 90% attendance and run regression (4) three times, once for each centered variable. The DD estimates in tables 3.10-3.12 report the three coefficients δ_3, i.e. the effect for a child with 70%, 80% and 90% school attendance, respectively, and once for the coefficient δ_6, the extra effect of a 100 percentage point increase in school attendance.

Table 3.10: Treatment effect for hemoglobin and anemia (heterogeneous treatment effects for different rates of school attendance)

	(1) Hemoglobin		(2) Any anemia		(3) Mild anemia		(4) Moderate or severe anemia		(5) N of anemia symptoms	
	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD
Treat (70% attendance)	0.023	0.138*	-0.018	-0.089***	0.001	-0.060**	-0.019	-0.029	-0.023	-0.130
Standard error	0.059	0.075	0.024	0.032	0.014	0.027	0.019	0.031	0.084	0.098
P value	0.696	0.069	0.461	0.007	0.931	0.028	0.317	0.347	0.785	0.190
Treat (80% attendance)	0.040	0.144	-0.020	-0.093**	-0.006	-0.058*	-0.015	-0.036	-0.012	-0.143
Standard error	0.063	0.089	0.028	0.039	0.015	0.030	0.021	0.034	0.090	0.112
P value	0.525	0.107	0.460	0.019	0.705	0.061	0.493	0.301	0.895	0.203
Treat (90% attendance)	0.057	0.151	-0.023	-0.098*	-0.013	-0.056	-0.011	-0.042	-0.001	-0.157
Standard error	0.079	0.113	0.036	0.051	0.021	0.039	0.027	0.042	0.108	0.139
P value	0.470	0.184	0.520	0.059	0.546	0.158	0.693	0.314	0.994	0.261
Treat*Attendance	0.172	0.065	-0.029	-0.046	-0.070	0.021	0.041	-0.067	0.110	-0.137
Standard error	0.296	0.363	0.132	0.175	0.093	0.139	0.088	0.130	0.352	0.443
P value	0.561	0.857	0.828	0.793	0.453	0.883	0.643	0.609	0.754	0.757
Mean dependent var	11.515	11.530	0.453	0.451	0.189	0.191	0.265	0.260	1.077	1.093
Child fixed effects		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes
N	1,703	2,780	1,703	2,780	1,703	2,780	1,703	2,780	1,746	2,780

SD: Simple difference. DD: Double-Difference, N: Number of observations. All coefficients are from separate regression, except for the last coefficients that are from one regression where treatment and attendance are interacted. All DD estimates use a balanced sample for hemoglobin and anemia outcomes, except for (5) number of anemia symptoms, which is based on its own balanced sample as it only contains variables from the household dataset. Any anemia is defined as a hemoglobin value < 11.5 g/dl, mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dl, moderate/severe anemia is defined as a hemoglobin value < 11 g/dl. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Standard errors are clustered at the school level.

Table 3.11: Treatment effect for cognitive outcomes (heterogeneous treatment effects for different rates of school attendance)

	(1) Block design		(2) Digit span forward		(3) Digit span backward		(4) Progressive matrices		(5) Day and night		(6) Cognitive Index	
	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD
Treat (70% attendance)	0.053	0.005	-0.020	-0.117	0.122	0.001	-0.005	0.065	0.094	0.109	0.074	0.017
Standard error	0.069	0.081	0.065	0.073	0.094	0.090	0.052	0.099	0.091	0.100	0.091	0.083
P value	0.438	0.946	0.754	0.112	0.197	0.992	0.919	0.513	0.301	0.276	0.421	0.834
Treat (80% attendance)	0.088	0.025	-0.003	-0.128	0.148	-0.011	0.010	0.089	0.109	0.160	0.106	0.038
Standard error	0.069	0.086	0.068	0.081	0.093	0.084	0.051	0.097	0.096	0.102	0.094	0.084
P value	0.204	0.770	0.960	0.117	0.114	0.895	0.846	0.365	0.257	0.120	0.266	0.653
Treat (90% attendance)	0.123	0.045	0.014	-0.139	0.174*	-0.023	0.025	0.112	0.124	0.210	0.137	0.058
Standard error	0.078	0.104	0.079	0.099	0.102	0.095	0.058	0.110	0.108	0.114	0.106	0.096
P value	0.119	0.665	0.864	0.163	0.090	0.808	0.667	0.311	0.254	0.068	0.198	0.550
Treat*Attendance	0.350	0.198	0.170	-0.108	0.259	-0.120	0.151	0.235	0.147	0.504	0.318	0.203
Standard error	0.255	0.353	0.258	0.316	0.308	0.392	0.212	0.386	0.283	0.329	0.300	0.330
P value	0.173	0.576	0.511	0.732	0.402	0.761	0.477	0.544	0.605	0.129	0.292	0.541
Mean dependent var	0.017	0.049	0.012	0.035	0.016	0.036	-0.044	-0.023	-0.020	-0.002	-0.000	-0.039
Child fixed effects		Yes		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes		Yes
N	1,718	2,766	1,718	2,766	1,718	2,766	1,718	2,766	1,718	2,766	1,718	2,766

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. All coefficients are from separate regression, except for the last coefficients that are from one regression where treatment and attendance are interacted. For the DD estimates we use a balanced sample for cognitive and education outcomes except for attendance, which is based on its own balanced sample as it only contains variables from the household and attendance datasets. All outcomes are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 3.12: Treatment effect for education outcomes (heterogeneous treatment effects for different rates of school attendance)

	(1) Math		(2) Reading	
	SD	DD	SD	DD
Treat (70% attendance)	0.108	0.124	0.122	0.096
Standard error	0.121	0.088	0.119	0.081
P value	0.374	0.162	0.306	0.235
Treat (80% attendance)	0.155	0.161*	0.182	0.139
Standard error	0.118	0.092	0.128	0.087
P value	0.189	0.084	0.158	0.111
Treat (90% attendance)	0.203	0.197*	0.242	0.182*
Standard error	0.126	0.105	0.150	0.104
P value	0.111	0.063	0.110	0.083
Treat*Attendance	0.472	0.364	0.601	0.431
Standard error	0.385	0.305	0.440	0.342
P value	0.223	0.235	0.175	0.211
Mean dependent var	-0.026	-0.008	-0.036	-0.019
Child fixed effects	Yes		Yes	
Controls	Yes		Yes	
N	1,717	2,766	1,718	2,766

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. All coefficients are from separate regression, except for the last coefficients that are from one regression where treatment and attendance are interacted. For the DD estimates we use a balanced sample for cognitive and education outcomes except for attendance, which is based on its own balanced sample as it only contains variables from the household and attendance datasets. All outcomes, except attendance, are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level respectively.

For the hemoglobin outcome (table 3.10), the point estimates for children with higher attendance indeed increased but are not statistically significant. For a child with 70% school attendance, the treatment effect is 0.138 g/dl (P-value: 0.069, SE: 0.075), for a child with 80% school attendance it is 0.144 g/dl (not statistically significant) and for a child with 90% school attendance, the point estimate is 0.151 g/dl (not statistically significant). For any level of anemia and mild anemia, the point estimates stay nearly the same with higher school attendance. The interaction *Treat*Attendance* shows the global effect of an increase in school attendance. For the hemoglobin outcome, the coefficient would be interpreted in the following way. On average, an increase in school attendance of 100 percentage points increases the treatment effect by 0.172 g/dl over and above the effect of the treatment effect with 0 school attendance.[41] However, the coefficient is not statistically significant for any of the outcomes, such that globally no extra effect of the treatment depending on school attendance can be found in the data. The sizes of the coefficients for the different forms of anemia stay very similar across different levels of school attendance. For the cognitive and education outcomes, the picture looks very much the same (tables 3.11 and 3.12). None of the coefficients are statistically significant at conventional levels, but the sizes of the coefficients increase as expected with higher school attendance. For the education outcomes, some of the coefficients for the DD estimates are statistically significant at the 10% level, which might be interpreted as weak evidence that with higher school attendance, i.e. more compliance with the treatment, math and reading scores increased further. For a child with 90% school attendance, the treatment increased math scores by 0.197 standard deviations (P-value: 0.063, SE: 0.105) and reading scores by 0.182 standard deviations (P-value: 0.083, SE: 0.104).

Scheduled caste and Scheduled tribe – From a welfare perspective it is also interesting to investigate if children with disadvantaged backgrounds benefit more from the intervention. Scheduled caste and scheduled tribes (SC/ST) are historically disadvantages groups in India. In our sample 29% of the children belong to SC/ST households (Table 3.2). We ran regression (3) and (4) but substitute *School attendance*$_i$ by the dummy variable *SC/ST*$_i$ that takes on the value of one if the child's parents report to belonging to a SC/ST and 0 otherwise.[42] We hence estimate the following regression for the SD estimates:

[41] Similarly, but maybe more intuitively, a 10 percentage points increase in school attendance increases hemoglobin values by 0.0172 g/dl.

[42] There are about 140 observations for which the SC/ST status changed between the baseline and endline, meaning that certain castes have either been incorporated in or excluded from the SC/ST category. We code the

$$y_{ai} = \alpha + \delta_1 SC/ST_i + \delta_2 Treat_a + \delta_3 Treat_a \times SC/ST_i + \eta_{ai} \quad (5)$$

We report coefficient δ_2 that gives the treatment effect for children from non-SC/ST households and δ_3 that gives the extra effect of the treatment for children from SC/ST households.

For the DD estimates the regression is as follows:

$$y_{ait} = \alpha_i + \delta_1 Post_t + \delta_2 SC/ST_i + \delta_3 Treat_a x\ Post_t + \delta_4 Post_t \times SC/ST_i + \delta_5 Treat_a \times SC/ST_i + \delta_6 Post_t \times Treat_a \times SC/ST_i + X'_{at}\ \varepsilon + Y'_{it}\ \zeta + \eta_{ait} \quad (6)$$

We report coefficient δ_3 that gives the treatment effect for children from non-SC/ST households and δ_6 that gives the extra effect of the treatment for children from SC/ST households.

Anemia rates are higher among children from SC/ST households. On average 44.1% of the non-SC/ST children were suffering from any form of anemia, whereas within the group of SC/ST children 47.6% were suffering from any form of anemia. Mild anemia rates are similar across the two groups. With a prevalence of 28.5% SC/ST children were more affected by moderate and severe anemia. The prevalence among non-SC/ST children was only 24.7% at baseline. Looking at the DD estimates, i.e. our preferred specification for the anemia outcomes (table 3.13) we do not find a statistically significant additional treatment effect for children from SC/ST households, however point estimates indicates that SC/ST benefitted slightly more. Whereas the treatment reduced any form of anemia for a child from a non-SC/ST household by 8.5 percentage points (P-value: 0.028, SE: 0.038) it additionally reduced any form of anemia in the group of SC/ST children by 1.8 percentage points, the latter point estimate is however not statistically significant at conventional levels. Point estimates for mild, moderate and severe anemia and number of anemia symptoms point into a similar direction are however also not statistically significant.

observations as belonging to SC/ST if either at the baseline or endline or in both surveys, parents reported to belonging to SC/ST.

Table 3.13: Treatment effect for SC/ST (anemia outcomes)

	(1) Hemoglobin		(2) Any anemia		(3) Mild anemia		(4) Moderate or severe anemia		(5) N of anemia symptoms	
	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD
Treat	0.009	0.138*	0.006	-0.085**	0.019	-0.052	-0.013	-0.033	-0.000	-0.082
Standard error	0.067	0.082	0.029	0.038	0.019	0.034	0.021	0.033	0.092	0.111
P value	0.892	0.095	0.839	0.028	0.304	0.132	0.534	0.313	0.996	0.460
Mean dependent var	11.563	11.570	0.436	0.441	0.191	0.194	0.245	0.247	1.110	1.134
Treat*SC/ST	0.065	-0.030	-0.093*	-0.018	-0.072**	-0.024	-0.021	0.006	-0.051	-0.120
Standard error	0.127	0.149	0.051	0.061	0.035	0.052	0.039	0.054	0.164	0.189
P value	0.614	0.839	0.069	0.774	0.042	0.652	0.597	0.912	0.755	0.528
Mean dependent var	11.419	11.443	0.492	0.476	0.186	0.191	0.306	0.285	0.988	0.977
Child fixed effects		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes
N	1,694	2,812	1,694	2,812	1,694	2,812	1,694	2,812	1,766	3,068

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. For all of the DD estimates we use a balanced sample for hemoglobin and anemia outcomes, except for (6) N of anemia symptoms, which is based on its own balanced sample as it only contains variables from the household dataset. Any anemia is defined as a hemoglobin value < 11.5 g/dl, mild anemia is defined as a hemoglobin value ≥ 11 & < 11.5 g/dl, moderate/severe anemia is defined as a hemoglobin value < 11 g/dl. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Standard errors are clustered at the school level.

Table 3.14: Treatment effect for SC/ST (cognitive outcomes)

	(1) Block design		(2) Digit span forward		(3) Digit span backward		(4) Progressive matrices		(5) Day and night		(6) Cognitive Index	
	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD	SD	DD
Treat	0.095	-0.002	0.037	-0.031	0.253***	0.103	0.091	0.031	0.164*	0.167	0.189	0.077
Standard error	0.082	0.093	0.071	0.075	0.088	0.096	0.060	0.112	0.096	0.101	0.096	0.087
P value	0.248	0.980	0.601	0.679	0.005	0.285	0.129	0.784	0.089	0.102	0.051	0.381
Mean dependent var	0.053	0.083	0.064	0.082	0.112	0.113	-0.010	0.016	0.011	0.025	0.077	0.025
Treat*SC/ST	-0.106	-0.033	-0.128	-0.111	-0.341***	-0.264*	-0.248***	0.038	-0.183	-0.048	-0.291**	-0.121
Standard error	0.121	0.120	0.116	0.139	0.129	0.152	0.093	0.176	0.142	0.157	0.139	0.131
P value	0.380	0.784	0.272	0.428	0.009	0.085	0.009	0.829	0.200	0.761	0.039	0.359
Mean dependent var	-0.052	-0.025	-0.129	-0.078	-0.204	-0.137	-0.127	-0.098	-0.111	-0.052	-0.175	-0.181
Child fixed effects		Yes		Yes		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes		Yes		Yes
N	1,709	2,790	1,708	2,790	1,708	2,790	1,708	2,790	1,708	2,790	1,708	2,790

SD: Simple-Difference, DD: Double-Difference, N: Number of observations. For the DD estimates we use a balanced sample for cognitive and education outcomes except for attendance, which is based on its own balanced sample as it only contains variables from the household and attendance datasets. All outcomes are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 3.15: Treatment effect for SC/ST (education outcomes)

	(1) Math		(2) Reading		(3) School attendance	DD
	SD	DD	SD	DD	SD	
Treat	0.229**	0.157*	0.258*	0.160*	-0.001	-0.009
Standard error	0.113	0.092	0.133	0.091	0.026	0.024
P value	0.045	0.092	0.055	0.082	0.971	0.713
Mean dependent var	0.080	0.086	0.059	0.067	0.791	0.805
Treat*SC/ST	-0.300	-0.048	-0.315*	-0.133	-0.020	0.013
Standard error	0.186	0.150	0.166	0.130	0.028	0.034
P value	0.109	0.750	0.060	0.311	0.467	0.698
Mean dependent var	-0.250	-0.206	-0.225	-0.199	0.762	0.768
Child fixed effects	Yes		Yes		Yes	
Controls	Yes		Yes		Yes	
N	1,707	2,790	1,708	2,790	1,877	2,896

SD: Simple Difference. DD: Double-Difference, N: Number of observations. For all DD estimates we use a balanced sample for cognitive and education outcomes. Standard errors are clustered at the school level. All outcomes, except attendance, are normalized with reference to the baseline mean in the control group. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Standard errors are clustered at the school level.

Regarding the cognitive and education outcomes (Tables 3.14 and 3.15), it is first remarkable that at the baseline, children from SC/ST households scored much worse than children from non-SC/ST households. Most of the interaction effects are negative indicating that the extra effect for SC/ST children is negative, i.e. SC/ST children benefitted less than children from non-SC/ST households with respect to cognitive outcomes. However the picture is not uniform across indicators as only the coefficients for the SD and DD estimates for the *block design backwards* as well as the SD estimates for the *progressive matrices* and for the *cognitive index* are statistically significant. This might still reflect that some initial level of cognitive ability is necessary such that the DFS can have an effect on those outcomes. The same reason might apply to the education outcomes (Table 3.15), where the mostly negative point estimates indicate that SC/ST children benefited less from the DFS treatment. However only the interaction effect for reading is statistically significant and only at the 10% level. Estimates for children from the non-SC/ST for the math and reading indicate to a positive and statistically significant treatment effect for this subgroup. At a statistical significance of 10%, the treatment increased reading scores by 0.160 (P-value: 0.091, SE: 0.082) and math scores by 0.229 standard deviations at a statistical significance of 5% (P-value: 0.113, SE: 0.045) for a child that does not belong to SC/ST household. For a children from a SC/ST the additional treatment effect was 0.109 and 0.027 standard deviations (DD estimates, not statistically significant).

3.6 Challenges to Internal Validity

3.6.1 Partial Compliance

No or partial compliance, a situation where individuals assigned to the treatment group do not or only partly take up the treatment, or where individuals in the control group (partly) take up the treatment, can threaten the internal validity, if this behavior results in the self-selection of individuals into the treatment or control group. Partial compliance in our setting might occur at the school or the individual level. Even if partial compliance is not systematic, it reduces the power of the estimates (Duflo, Glennerster and Kremer, 2007).

At the school level, it potentially might have been the case that DFS was not or insufficiently used. This risk was strongly reduced since the intervention was approved and supported by the local government, which obliged headmasters to purchase the DFS for the

usage of the MDM. Local authorities even intervened whenever an issue of non-compliance emerged (which was only once the case). The monthly or bi-monthly delivery of DFS to the treatment schools further functioned as a monitoring system. Our team confirmed that DFS was being used as instructed. Also during the unannounced endline survey, DFS was found in all treatment schools, except for two schools where the schools had run out of stock just the day before the survey and the headmasters were not able to contact our team to deliver DFS before the next delivery date. Another potential source of imperfect compliance at the school level might have occurred, if the delivery system had not worked perfectly. At endline, we asked headmasters and cooks in treatment schools if they ever experienced a situation where not sufficient DFS was available at the school for the preparation of the MDM. Seven out of 54 headmasters reported that this situation had ever occurred; however, in two schools this was only the case for one day, in three schools for two to five days and in the other two schools this was the case for ten days.[43] Since only a few schools, and only for a limited amount of time compared to a whole year of treatment, were affected by the shortage of DFS, treatment intensity can still be considered high. We also argue that this situation is close to a real-world policy. The lack of DFS emerged either because the stock of DFS was exhausted before the next delivery date and headmasters were not able to contact our team for earlier delivery, or because schools could not be reached due to monsoon flooding, which is typical of the study region and many other parts of South Asia.

One could further imagine that DFS from the treatment schools was resold to other schools, including control schools and households. However, this is very unlikely. First, because the financial incentive of reselling DFS is very small since salt is a relatively cheap product. Furthermore, we know from anecdotal evidence that awareness of the benefits of DFS, and hence the demand for it, is very low – at least among the rural households in our study region. In the parental questionnaire from the endline survey, we asked for the salt package used at the household, and if it was not available, we asked for the last salt brand that was bought. A negligible number of four households in the DD sample (0.28%) had *Tata Salt Plus* or reported that they had bought it, indicating that the reselling of salt to households did not happen. However, it would have been more likely that the headmasters in the treatment schools sold DFS to headmasters in the control schools, since headmasters in both control and treatment schools were informed of the study for ethical reasons. We told headmasters in the

[43] These numbers broadly match with the report of the cooks that were asked the same question in order to verify answers.

treatment and control schools that a lottery would take place in which they would have a 50% chance of being supplied with DFS to be used in the preparation of the MDM. They were later informed of the results of the lottery. Headmasters were further informed that the government supported this project and that they had to obey with the results of the lottery. Furthermore, headmasters in control schools were told that if the given study would yield positive results, they would also have the possibility to purchase DFS. In the endline survey, headmasters in the control schools were asked about their awareness of DFS. 29 out of 53 headmasters had heard about iron-fortified iodized salt, but only two stated that they knew that the salt was available in other schools. None of the control headmasters reported to have ever used DFS in the preparation of the MDM.[44] Also, during the unheralded endline survey, no DFS was found in any of the control schools. Apparently, the prospect of the offer of DFS after the completion of the study and the order from the government to obey the study design, provided enough reason for headmasters in control schools to postpone their demand for DFS. In combination with the non-existence of an incentive for headmasters in treatment schools to sell DFS, we are very certain that partial compliance at the school level is not an issue.

Non-compliance at the child level due to low school attendance has been discussed in section 3.5.2. Non-compliance at the individual level could have potentially also occurred, if children that were enrolled in control schools at the baseline went to treatment schools to receive the treatment and the other way around. This potential bias was encountered in several ways. First, only headmasters, but not parents and children, knew about the interventions. Anecdotal evidence further confirms that awareness of the benefits of DFS was very low, such that the incentive of parents to send their child to a treatment school to receive the DFS was non-existent. In our study region, every village generally has one government-funded school and going to a school in another village is very uncommon and the larger travel distance would have been an additional disincentive.

3.6.2 Attenuation Bias

When conducting research with human subjects, and especially with vulnerable populations such as children, it has to be ensured that the benefits of the research outweigh the risks for ethical reasons (Medical Research Council, 2004). In order to maximize the benefit for the

[44] Initially two headmasters reported that they had used DFS. This was verified later on and it turned out that the headmasters had confused DFS with conventional TATA iodized salt.

children involved in the survey, medical personnel was instructed to tell parents in case their child was moderately or severely anemic and to advise them to feed their children more diversely and particularly more food items with high iron content (green leafy vegetables and meat in case they were non-vegetarians), in cases where a child was diagnosed as moderately anemic. In cases of severe anemia, parents were instructed to consult a doctor (which was only the case for 14 children). In general, this additional intervention does not bias our results, since this information intervention affected the treatment and control groups equally and its effect is therefore balanced between these two groups. However, in case the information intervention did indeed provoke a change in feeding practices or medical treatment, a *saturation effect* might have occurred (i.e. decreasing returns to scale of iron-interventions). It might have been the case that the usage of DFS in the school's MDM could only add little to the already higher hemoglobin level, the reduced anemia prevalence and the improved cognitive and education outcomes. Compared to an exclusive DFS intervention, the estimated coefficients in this study could therefore be downward biased and might constitute a lower bound. To encounter this potential thread we included the dietary diversity score of the child as a control variable for feeding practices in the DD estimates. We further analyzed the effect of the nutrition information intervention in essay 3 of this dissertation (chapter 4). We did not find any robust effect of the nutrition information intervention on any of the tested outcomes, such that we are very certain that the nutrition information did not provoke attenuation bias.

3.6.3 Hawthorne Effect

The Hawthorne effect describes a change in the behavior of the individuals in the control group as a reaction of being part of an experiment. In our experiment, individuals in the treatment as well as in the control group, were surveyed once at the baseline and once at the endline; hence, any behavioral change that results from the survey itself (being monitored or the evaluation of the education level of the students) is balanced in the treatment and the control group. It might still have been the case that individuals in the treatment group changed their behavior due to the treatment itself (regular delivery of DFS to the school) and that this change affected the outcomes. Since only headmasters, but not parents, knew of the intervention, this behavioral change is limited to headmasters (and maybe some teachers who knew about the intervention). A change in hemoglobin values due to a behavioral change by the headmasters (e.g. spending more money on the MDM to improve dietary diversity) has very unlikely occurred since

headmasters generally did not know about the hemoglobin testing as it took place in the villages or at the households but not in the schools. Nevertheless, we also control for the average calories and the average iron content of the MDM as measured on the survey day. We also believe that the expected benefits of a behavioral change were too small in comparison to the effort needed to manipulate the outcomes. A change in the components of the MDM would involve additional costs and a change in cognitive or education outcomes would need a large quality improvement in teaching (e.g. more teachers, more training material etc.). This is opposed to the benefits of being offered DFS at a subsidized price for a longer time period and maybe to please the researchers over and above the level that control schools already do because of the baseline and endline survey itself.

3.7 Cost-Effectiveness Analysis

Having found positive impacts from the DFS intervention does not indicate that this intervention should become a policy priority. It might be the case that other interventions have a better cost-effectiveness ratio and thus, in a world of scarce resources where all interventions with positive impact results cannot be implemented, various interventions must be compared through cost-effectiveness analysis to determine the most cost-effective choice.

The costs of the DFS intervention are composed of the cost of the DFS subsidy (13 rupees per kilogram of DFS) and the costs of delivering DFS to schools (the wages of the person delivering the salt and transportation costs). We calculate the yearly costs per child to be 32.61 INR, which is equivalent to 0.48 USD[45] or 2.08 International Dollar (I$)[46].[47] The details of how we arrived at this number are outlined in Table 3.16. If the intervention were to be implemented on a large scale, these costs – and especially the delivery costs – could possibly be much lower due to economics of scale.

[45] Using the average exchange rate from 2016.

[46] An international $ has the same purchasing power as the USD has in the USA. It is derived by dividing the local currency by the purchasing power parity exchange rate 2005 (PPP). PPP for India is 15.66 (WHO, 2017a).

[47] There is no uniform methodological approach for standardizing costs across cost-effectiveness studies. Whereas the WHO promotes the usage of I$, other authors use the local currency, exchange rates at the time of the analysis or standardize costs to USD in 2012 (Black *et al.*, 2016).

Table 3.16: Calculating the cost of the intervention

	Monthly costs per child (13.794 children were present at school at the endline survey)	**Daily costs per child** (4g of DFS per day)	**Yearly costs per child** (244 school days from August 2015 till July 2016)
Costs of DFS Subsidy per kg: 13 INR		4g*0.013 INR=0.052 INR	0.052 INR*244 days = 12.69 INR
Cost of DFS delivery 6.800 INR: Wages for distributor 12.000 INR: Vehicle and fuel	18.800 INR/13.794 children = 1.36 INR		1.36 INR*12 month = 16.36
Total costs per child per year			**32.61 INR (0.48 USD[1], 2.08 I$[2])**

[1] Average exchange rate in 2016.
[2] In purchasing power parity exchange rate 2015.

We calculate the effects in terms of the number of disability-adjusted life years (DALYs) saved (Table 3.17). DALYs are a widely accepted metric for the burden of disease.[48] DALYs are calculated as the sum of the years lived with disability or disease (YLD) and the years of premature death in comparison to the life expectancy (YLL) (Murray and Lopez, 2013). Since in our survey we only had a very low number of children with severe anemia, but the risk of death is generally only attributed to severe anemia (Stein *et al.*, 2005), we only include YLD in the effect calculation but no YLL. For 100 children, the intervention saved 11.267 DALYs if the effect on mild and the (not statistically significant) effect on moderate and severe anemia is taken into account (Table 3.17).

Table 3.17: Calculating the number of disability-adjusted life years (DALYs) averted

	(1) No. of children affected	(2) Duration until remission or death[1]	(3) Dis-ability weight[2]	(4) (1)*(2)*(3) = YLD = DALYs	(5) **Difference Control and Treat = DALYs saved**	(6) **Cost per DALYs saved I$ (USD)[3]**
Control group						
Mild anemia	19	59	0.004	4.484		
Moderate and severe anemia	26	59	0.052	79.768		
			SUM	84.252		
Treatment group						
Mild anemia	15	59	0.004	3.587	0.897	**231.94 (53.52)**

[48] In the DALYs metric each healthy year lived without disease or disability is assigned a value of 0 (no disease burden at all) and the factor one is counted for each year of premature death. Years lived in disease or disability are weighted based on the severity of the disability (Murray and Lopez, 2013).

Moderate and severe anemia	23	59	0.052	69.398	10.370	20.06 (4.63)
		SUM		72.985	**11.267**	**18.46 (4.26)**

Notes: I$: International Dollar. Calculation per 100 children. At the baseline, 19% of the children were mildly anemic and 26% were moderately or severely anemic. The impact of the DFS treatment: A 20% reduction in the prevalence of mild and 13% reduction in the prevalence of moderate and severe anemia (not statistically significant).
[1] Life expectancy in India in 2014: 68 years (World Bank, 2017b), age of the children in this sample at the endline: 9 years; 68 years minus 9 years = 59 years
[2] Disability weights from the global burden of disease study (Salomon, Haagsma and Davis, 2015). Since the proportion of severely anemic children is very small, we take the disability weight for moderate anemia.
[3] Costs of the intervention as calculated in table 3.16. Cost per year for 100 children: 208 I$ (51 USD), divided by DALYs saved for 100 children (column 5).

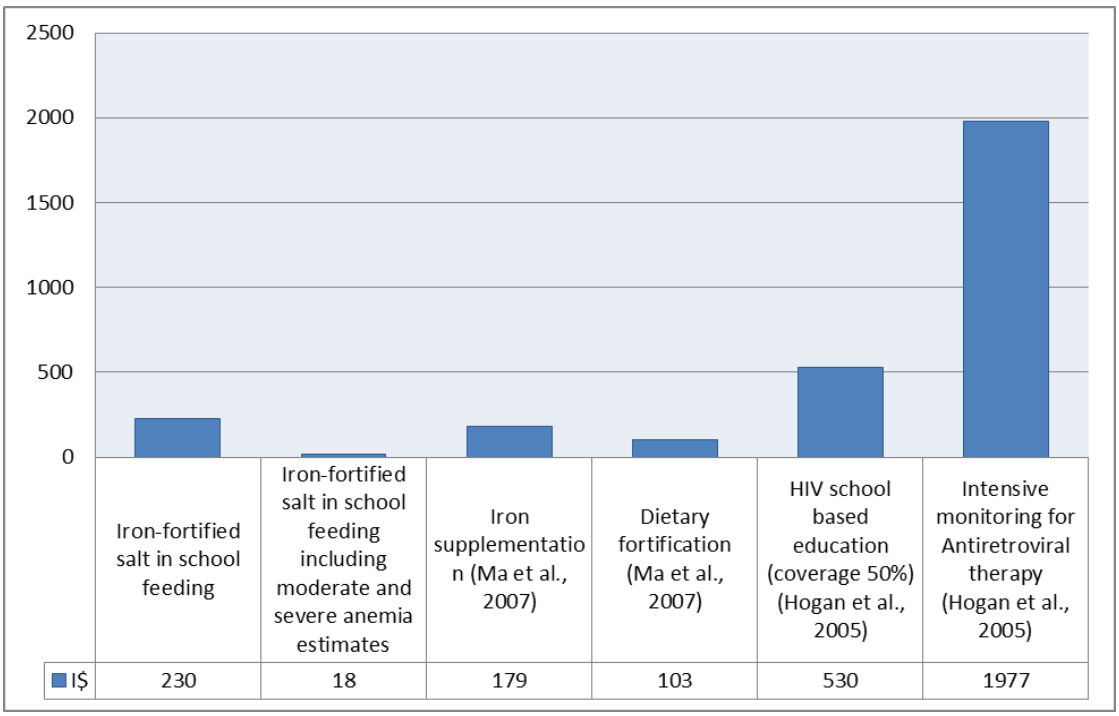
Because DFS was consumed by all children attending government-funded schools, costs also apply to all children attending these schools; however, we only know the prevalence of anemia for the second-grade students. As such, we assume that the prevalence of anemia is the same across all grades. Per DALY averted, the total cost of the intervention was 4.26 USD or 18.46 I$ (Table 3.17). We have found only a few other evaluations that were as cost-effective as our intervention. For example, Baird et al. (2003) calculated a cost-effectiveness of 5 USD[49] per DALY for childhood deworming. Black et al. (2016) compare the cost-effectiveness of different reproductive, maternal, new-born and child health interventions and prove that pneumococcus vaccinations costing between 1 and 30 USD per DALY averted, and community based Vitamin A supplementations costing between 1 and 9 USD per DALY averted, as the most cost-effective interventions.[50] These numbers indicate that using DFS, in the existing school-feeding program, is extremely cost-effective.

Most of the reduction in DALYs is based on moderate and severe anemia, as the adverse health consequences from moderate and severe anemia are much higher and thus assigned higher disability weights (Table 3.17). If only the (statistically significant) effect of the intervention on mild anemia is taken into account, for 100 children only 0.897 DALYs were averted. However, even when considering the cost per DALY saved for the reduction in mild anemia only, the intervention remains very cost-effective, with a cost of 53.52 USD or 231.94 I$.

[49] The authors do not specify how they exactly derived at this number. We assume that they are referring to USD at the time of publication, 2004, and not to I$. 5 USD in 2014 would be 5.07 USD in 2016.
[50] Black et al. (2016) report costs in 2012 USD, the inflation rate between 2012 and 2014 was 4.5%, i.e. one USD in 2012 was 1.05 USD in 2016.

Figure 3.3 illustrates the cost-effectiveness for different interventions. It shows the costs in I$ per DALY adverted for different interventions. The first bar shows the costs of the DFS school-intervention when only the effect on mild anemia is taken into account and the second bar shows the costs when also the non-statistically significant effect on moderate and severe anemia is taken into account. These costs are compared to the costs of two micronutrient intervention assessments from Ma et al. (2007) (bars three and four) and two HIV interventions. Ma et al. (2007) calculated costs of 66 I$ for food fortification with iron, 179 I$ for iron supplementation and 103 I$ for dietary diversification.[51] Hogan et al. (2005) report costs for different HIV interventions with a very wide range, where the most expensive ones go up to 2000 I$. It becomes clear from this figure that the interventions evaluated in this essay is in the middle range of costs when only the statistically significant effect on mild anemia is considered but moves to an extremely cost effective intervention when adverse effects from moderate and severe anemia are also taken into account. Hence also in terms of cost-effectiveness the longer term effects, where an impact on moderate and severe anemia is more likely to occur, are of particular interest.

Figure 3.3: Cost-effectiveness of the DFS school-level intervention



[51] Although DALYs allow the comparison of costs cross interventions and diseases, the methodologies in the computation of the cost per DALYs averted vary to some degree. Cost can also vary between regions. Therefore, these numbers should only be taken as a benchmark.

The WHO assesses interventions as very cost-effective if the cost per DALY averted is less than the GNI per capita of the country where the intervention is going to be implemented and cost-effective if it is less than three times the GNI per capita (Sachs, 2001). India's GNI per capita in 2015 was 6,030 USD (World Bank, 2017a), which means that any intervention costing less than 6,030 USD per DALY averted would be considered very cost-effective under the WHO definition. The intervention evaluated in this study strongly falls below this threshold.

3.8 Conclusion

In this study, we assessed if the Indian MDM has the potential to be used as an effective channel to provide school-aged children with iron, administered by salt as iron vector. We tested the effectiveness of such an intervention by running a randomized controlled trial with an analytic sample of 107 schools, and depending on the econometric specification, 1400 to 1700 children from two blocks in Bihar, India. This was the first time the effect of the usage of iron-fortified salt in a school-feeding program was tested. The second innovation of this study was the investigation of productive outcomes, such as cognition and education, in addition to anemia outcomes.

The treatment was able to reduce the prevalence of any form of anemia by 20% and of mild anemia by 30%. There is no statistically significant interaction effect between treatment and school attendance, i.e. the treatment effect does not change significantly with increasing compliance. However, despite the short treatment period, there is weak evidence that with higher school attendance (80% or 90% school attendance) reading and math scores were improved by a little less than 0.2 standard deviations. Differential effects for children from SC/ST households are not statistically significant. Point estimates indicate that children from SC/ST households benefited slightly more from the intervention in the form of a reduction in anemia but less with respect to cognition and education outcomes, which we explain by the requirement of a minimum education level such that the DFS treatment can have an effect on those outcomes. We further find that the intervention is cost-effective.

These results are encouraging with respect to using iron-fortified salt in the MDM. They are also particularly interesting in light of the results from the Banerjee et al. (2016) study, which did not find any effect of DFS when sold at subsidized prices through local

shops or the PDS system, and also in light of the study by Banerjee et al. (2011), which offered decentralized flour iron-fortification and did not find any positive impact at the endline due to a rapid drop in take-up after six months. We contribute to the literature by proving a positive impact on second grade students through an alternative delivery channel: a school-feeding program. One central difference between the delivery channel of our study and the studies by Banerjee et al. (2016) and Banerjee et al. (2011) is that in our study, the DFS was mandatorily used, such that the end consumer was given no choice in purchase or consumption, whereas Banerjee et al. (2016) and Banerjee et al. (2011) tested the effect of a voluntary DFS policy, where the end consumer was given the option to buy the fortified product or a conventional one or to fortify their flour or not. The positive effects found in our study therefore provide evidence of the potential advantage of mandatory fortification where a behavioral change in dietary patterns at the household level, which seems to be a central challenge, is not required.

There are some limitations of our study and potentials for further research. First, since our results only apply to the two blocks in Bihar from which we have drawn a random sample of schools (and maybe to some other areas with similar conditions), we propose replications of this study in other contexts and to include measures on productive outcomes such as cognition and education. Furthermore, though we encountered small irregularities in supplying schools with DFS due to weather conditions and whenever headmasters did not communicate their need for a new batch of DFS, the presented results mimic a real world situation of a sufficiently well-functioning supply system of DFS to government-funded school. In our supply system, a strong and trusting relationship between the headmasters and DFS distributors was maintained and the intervention was strongly supported by the local government. We therefore cannot make any statement on the effectiveness of other delivery channels, i.e. if the DFS would be available in local shops and the government would enforce usage or if it would be supplied by an already existing delivery system, such as the Food Corporation of India (that already delivers rice to the schools). Evaluating other and maybe more integrated and sustainable delivery channels for the DFS, would be of high policy relevance. Furthermore, we would like to point to the duration of the intervention. Though we could not find positive effects for the cognitive outcomes and only weak evidence for positive effects on education for subgroups after one year of the intervention, it might very well be the case that these effects indeed exist but need longer time periods to materialize. We therefore propose to study the longer run effects of the intervention. In addition, we would like to

emphasize that an increase in the availability of iron in the body, strongly depends on the joint consumption of iron with other products that inhibit or facilitate the absorption. Whereas Vitamin C rich items facilitate the absorption of iron, milk, rice and whole grain wheat products work as inhibitors. The absorption of iron in a rice and wheat-rich diet such as the one in South Asia is therefore particularly challenging. Testing the joint effects of DFS with products that facilitate absorption and that can easily be included in the MDM would be advisable, but possibly constitutes an even more challenging task than the usage of iron-fortified salt.

3.A Appendix

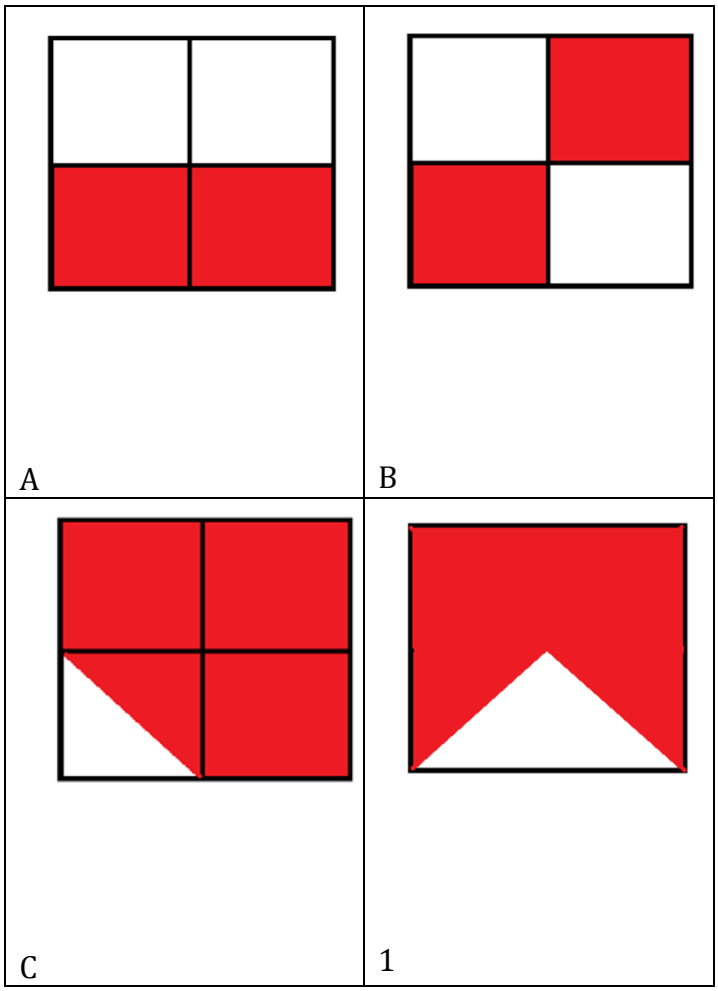
Table 3.A.1: MDE for different outcomes and different % of take-up

	Sd	Baseline ICC	MDE (c = 0.7)	MDE (c = 0.8)	MDE (c = 0.9)
Anemia					
Hemoglobin	1.10342	0.03130	0.23313	0.20399	0.18132
Any anemia	0.49803	0.02208	0.09889	0.08653	0.07691
Mild anemia	0.39225	0.00000	0.06438	0.05633	0.05007
Moderate anemia	0.43686	0.02087	0.08599	0.07524	0.06688
Severe anemia	0.08848	0.00000	0.01452	0.01271	0.01129
N anemia symptoms	1.09933	0.08387	0.29981	0.26233	0.23318
Perceived child health	0.40007	0.09271	0.11272	0.09863	0.08767
Cognition and education					
Block design	0.98478	0.13069	0.31275	0.27366	0.24325
Digit span forwards	0.98601	0.06170	0.24519	0.21454	0.19070
Digit span backwards	1.00564	0.10957	0.29987	0.26239	0.23323
Progressive Matrices	1.01746	0.11758	0.31103	0.27215	0.24191
Day and night	0.98825	0.09899	0.28460	0.24902	0.22135
Education					
Math	0.98212	0.13875	0.31888	0.27902	0.24802
Reading	0.99259	0.20815	0.37754	0.33035	0.29364
School attendance	0.17110	0.19551	0.06345	0.05552	0.04935

Sd: Standard deviation. ICC: Inter-cluster correlation. MDE: Minimal detectable effect. c: Take up rate. Assumptions: Sample size per cluster 22, number of clusters: 108 schools, division of observations between treatment and control: 50:50. Hemoglobin in expressed in g/dl and the different forms of anemia represent percentage points). Cognitive and education outcomes are normalized with respect to the control group mean and standard deviation.

Figure 3.A.1: Material cognitive tests

A. Block design



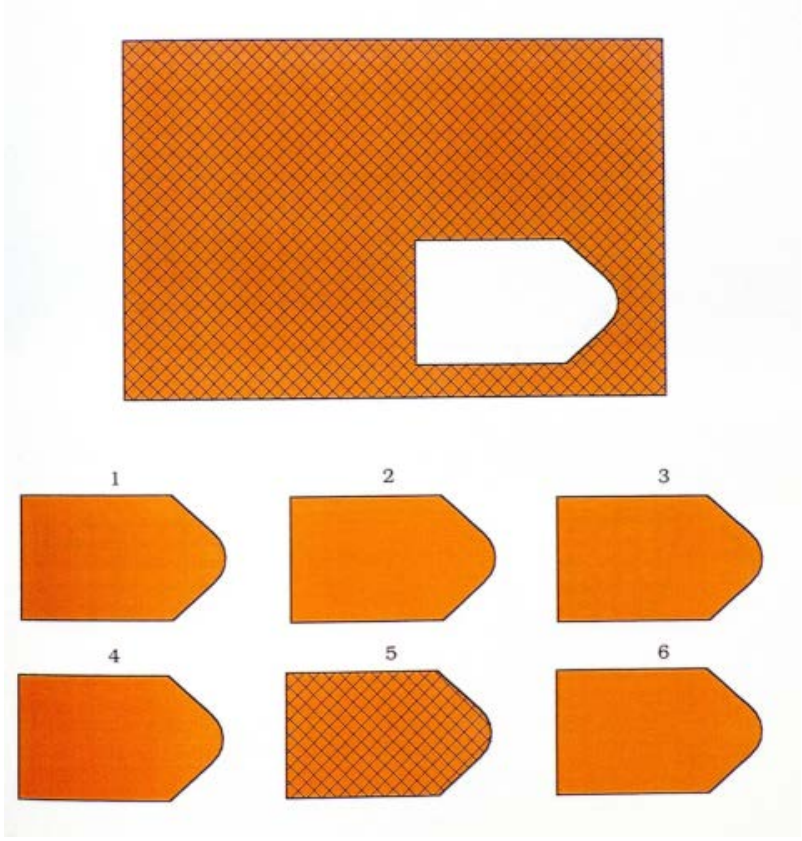
B. Forward and Backward digit span

	Column 1	Column 2
Forward test	(3) 2-6-5 (4) 1-5-2-3 (5) 2-4-7-6-1 (6) 4-2-1-9-3-7 (7) 3-6-4-8-5-2-9 (8) 7-5-8-2-9-6-1-3 (9) 5-8-6-4-2-7-3-9-1	(3) 2-8-1 (4) 1-9-5-2 (5) 5-2-1-4-3 (6) 8-5-3-1-4-7 (7) 6-8-1-4-7-2-5 (8) 2-8-5-9-7-3-1-4 (9) 4-2-5-8-1-3-9-7-6
Backward test	(2) 2-1 (3) 5-8-4 (4) 4-8-9-1 (5) 6-8-7-2-1 (6) 5-8-1-7-4-6 (7) 8-5-3-6-7-2-9 (8) 1-7-4-3-8-9-5-2	(2) 2-8 (3) 3-2-8 (4) 2-9-4-1 (5) 3-5-9-7-6 (6) 4-3-1-9-2-5 (7) 5-3-2-4-1-6-8 (8) 6-8-4-7-5-3-9-2

Picture is taken from Larsen et al. (2000)

C. Stroop-like day and night test

D. Colored progressive matrices



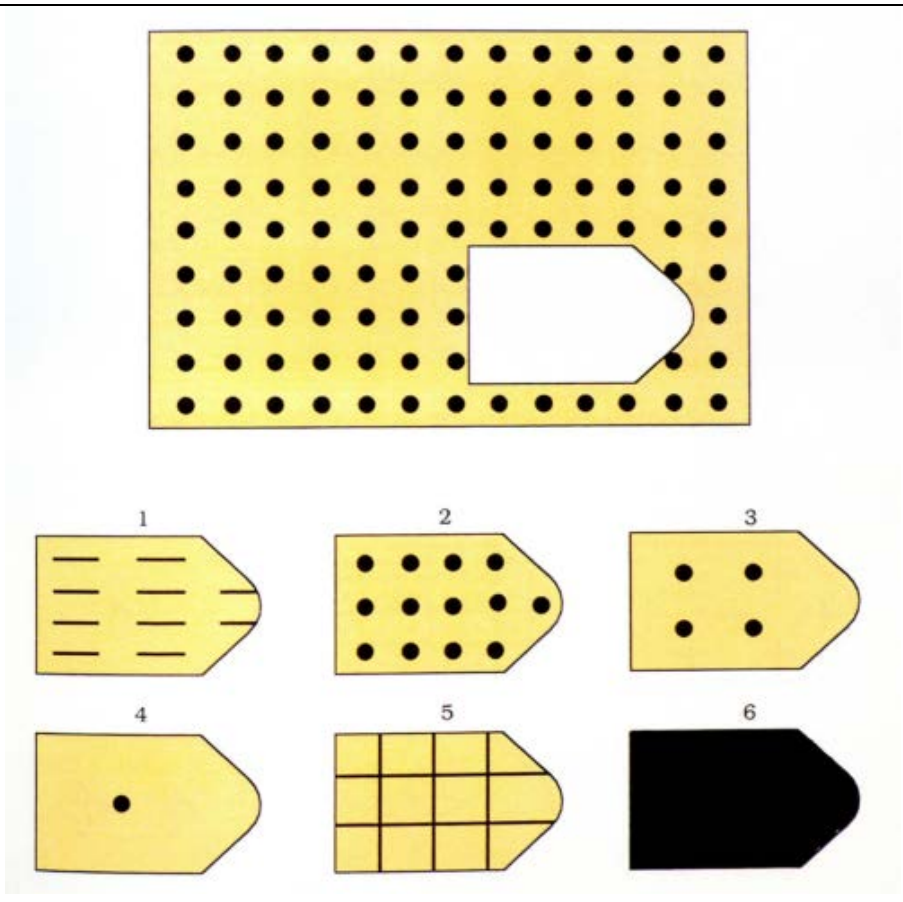


Figure 3.A.2: Material education tests

A. Math test

1. Digit Recognition (1-9)	2. Quantity Recognition (10-20)	3. Quantity Recognition (21-99)
5 9 8 1 2	12 16 18 19 11	71 24 92 86 48
	Trial A	Trial B
Single digit		
1. Addition (single-digit)	2 + 3 =	4 + 5 =
2. Subtraction (single-digit)	8 - 3 =	9 - 4 =
3. Multiplication (single-digit)	2 x 4 =	3 x 3 =
4. Division (single-digit)	9 ÷ 3 =	8 ÷ 4 =
Two digit		
5. Addition (2-digit)	42 + 14	11 + 68
6. Addition (2) (2-digit)	35 + 18	23 + 67
7. Subtraction (2-digit)	29 - 14	34 - 12
8. Subtraction (2) (2-digit)	47 - 18	33 - 15
9. Multiplication (2 digit)	23 x 4 =	16 x 5 =
10. Division (2 digit)	39 ÷ 3 =	44 ÷ 4 =

A. Reading test

A big tree stood in a garden. It was alone and lonely. One day a bird came and sat on it. The bird held a seed in its beak. It dropped the seed near the tree. A small plant grew there. Soon there was another tree. The big tree was happy.

Rani likes her school.
Her class is in a big room.
Rani has a bag and a book.
She also has a pen.

e	d	w
s	c	
g	h	z
i	q	

hand	star
bus	
cat	book
day	few
old	
sing	bold

Text 3.A.1: Choice of Covariates

We developed a set of covariates separately for the three outcome categories (hemoglobin, cognitive ability, education outcomes). The scientific literature that was used as a basis for the choice of covariates was selected by the following search strategy. "Predictors/Determinates of anemia (in childhood/school-aged children)" "Predictors/Determinates of cognitive ability/cognition in childhood anemia (in childhood/school-aged children)" "Predictors/Determinates of education outcomes/test scores" were used as key words on Science Direct, Google scholar, Google and the Electronic journal library of the library of the University of Goettingen. Articles were chosen based on the suitability of the title and the abstract. Focus was given to more recently published articles and systematic reviews, as those articles are most adequate in reflecting the current state of the art. If studies on India or low- and middle-income countries were available, we also gave priority to those studies as they reflect the country context of our study. We do not include controls that constitute outcomes in other regressions, as this might prevent us from detecting the true effect of the intervention if those outcomes are indeed affected by the treatment (e.g. hemoglobin as determinates for cognitive ability). Sometimes indicators had to be adjusted or different indicators had to be used for specific determinates due to data availability (i.e. we did not conduct an analysis of the children's stool for hookworm infections but we substitute this determinate by the imperfect measure of diarrhea incidences of the child as reported by the mother). We categorized the determinants into socioeconomic characteristics, nutritional factors, access to healthcare and morbidity indicators for the anemia outcomes. For the cognitive outcomes, additional indicators for psychosocial stimuli were collected. Further indicators for quality of schooling were collected for the education outcomes.

Regarding the determinates of *anemia* in childhood, we choose covariates in accordance with the comprehensive reviews of Osório (2002) and Goswmai & Das (2015) and Pasricha et al. (2010). As socioeconomic covariates, we include if the household is rural or urban, the wealth of the household,[52] parental education level, caste, religion and the total number of household members. Since closeness to the next city and weather conditions are different in the two blocks included in the study, we additionally include a control for these two blocks. Socioeconomic factors are supposed to influence hemoglobin level mostly

[52] The asset index is calculated using the following variables: Type of toilet facility (improved and unimproved according to the WHO), source of drinking water (improved or unimproved), type of house, wall, roof and floor, possession of assets such as a chair, table, radio, pressure cooker etc., the amount of agricultural land owned, amount of different farm animals owned, BPL card holder, MNREGA card holder and the dependency ratio. The asset index is calculated using principal component analysis.

through the accessibility and availability of iron-rich foods and to health care. As hemoglobin is mostly determined by food intake, we include the dietary diversity score of the child, a dummy if the child consumes any meat, poultry or fish, if the child receives iron supplements, an indicator for household food security and the number of meals the child eats every day. We also include the average iron and calorie intake from the MDM at the survey day per child. We further theorize that maternal health knowledge might be correlated with the child's hemoglobin level and include a maternal health knowledge indicator. We do not control for breastfeeding, which would generally protect the child from anemia, as this variable suffers from a strong recall error. We also do not control for maternal BMI due to too many missing values. Prenatal health care visits and institutional deliveries, where mothers might be informed about adequate nutrition and nutrition supplementation, might further be associated with a child's hemoglobin level. We include a covariate for institutional delivery, but not for prenatal health care visits as the latter variable exhibits several missing values. We further include a dummy for health insurance coverage since health insurance coverage might be related to prevention and detection of anemia. Some evidence exists that links anemia to increased morbidity. We control for the incidence of diarrhea for each child (in the last 30 days) and if the household possesses improved sanitation facilities. Lastly, as biological factors we include the sex of the child and its birth order. We do not include the age of the child due to a strong recall error.

The choice of the covariates for the *cognitive outcomes* are based on Santos et al. (2008) and Paxson & Schady (2007). We include the same socioeconomic, morbidity and biological factors as mentioned above in the anemia outcome regressions. Except for nutritional factors, where we include the dietary diversity score of the child, an indicator for household food security, the number of meals the child eats every day and maternal health knowledge. We also include the average calorie intake from the MDM on the survey day per child. In addition, we include if the father lives in the household as a socioeconomic variable. We further include three indicators for interaction and psychosocial stimuli, which are if the mother helps the child with their homework and how much time the mother spends on giving physical care to their child as well as how often parents participate in the parent-teacher meetings. To account for systematic differences in test administration across enumerators, we include the interviewer id as control variable.

For *education outcomes* we follow the overview of empirical evidence by Considine

& Zappalà (2002) and White et al. (2015). In addition to the covariates included in the regressions for cognitive outcomes, we include indicators for education quality and accessibility: total school enrollment, number of children in second/fourth grade, distance to school and the student-teacher ratio.

Table 3.A.2: Overview control variables

	Hemoglobin (H1)	Cognition (H2)	Education (H3)
Socioeconomic	- HH rural or urban - Block - Wealth index - Parental education level (years of schooling) - Caste - HH religion - N HH members	- HH rural or urban - Block - Wealth index - Parental education level (years of schooling) - Caste - HH religion and - N HH members	- HH rural or urban - Block - Wealth index - Parental education level (years of schooling) - Caste - HH religion and - N HH members
Food intake	- Child's dietary diversity score - Indicator for household food security - The number of meals the child eats every day - Average calorie intake from the MDM at survey day per child - Indicator for maternal health knowledge - Dummy if the child consumes any meat, poultry or fish - Dummy if the child get iron supplements - Average iron intake from the MDM at survey day per child - Dummy if the child get iron supplements	- Child's dietary diversity score - Indicator for household food security - The number of meals the child eats every day - Average calorie intake from the MDM at survey day per child - Indicator for maternal health knowledge	- Child's dietary diversity score - Indicator for household food security - The number of meals the child eats every day - Average calorie intake from the MDM at survey day per child - Indicator for maternal health knowledge
Health care	- Dummy for institutional delivery - Dummy for health insurance coverage	- Dummy for institutional delivery - Dummy for health insurance coverage	- Dummy for institutional delivery - Dummy for health insurance coverage
Morbidity	- Diarrhea incidence of child (in last 30 days) - Improved sanitation facilities	- Diarrhea incidence of child (in last 30 days) - Improved sanitation facilities	- Diarrhea incidence of child (in last 30 days) - Improved sanitation facilities
Biological factors	- Sex of child	- Sex of child	- Sex of child
Interaction and		- If the mother helps the	- If the mother helps the child

psychosocial stimuli		child with its homework - Time the mother spends on giving physical care to the child - Parental participation in parent-teacher - Father lives in HH	with its homework - Time the mother spends on giving physical care to the child - Parental participation in parent-teacher - Father lives in HH
Test administration		- Interviewer ID	- Interviewer ID
Quality of schooling			- Total school enrollment - Number of children in second/fourth grade - Distance to school meetings - Student teacher ratio (school level)

Table 3.A.3: Sample description and balancing test (SD sample for cognitive and education outcomes)

	Panel A: SD sample								Panel B: Sample including attrited children							
		Total		Control		Treat				Total		Control		Treat		
	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means t-test
A. Child level variables																
Anemia																
Hemoglobin	1,686	11.52	1.10	11.61	1.10	11.43	1.10	0.007***	1,918	11.52	1.10	11.60	1.09	11.44	1.11	0.013**
Any anemia	1,686	0.45		0.42		0.48		0.059*	1,918	0.45		0.42		0.48		0.036**
Mild anemia	1,686	0.19		0.19		0.20		0.628	1,918	0.19		0.18		0.20		0.194
Moderate or severe anemia	1,686	0.26		0.24		0.28		0.084*	1,918	0.26		0.25		0.28		0.154
Number of anemia symptoms	1,736	1.08	1.12	1.06	1.10	1.09	1.13	0.738	2,005	1.07	1.10	1.06	1.09	1.09	1.11	0.676
Cognitive tests																
Block design	1,682	3.69	2.21	3.60	2.23	3.77	2.19	0.393	1,910	3.73	2.22	3.68	2.24	3.78	2.20	0.626
Digit span forward	1,683	4.07	0.98	4.06	0.99	4.07	0.98	0.819	1,912	4.08	1.00	4.07	1.02	4.08	0.99	0.886
Digit span backward	1,683	1.10	1.28	1.08	1.27	1.12	1.30	0.694	1,912	1.13	1.30	1.11	1.30	1.15	1.31	0.695
Progressive matrices	1,681	4.74	1.65	4.80	1.61	4.69	1.68	0.430	1,910	4.76	1.69	4.84	1.65	4.69	1.72	0.279
Day and night	1,681	5.27	3.39	5.29	3.42	5.25	3.36	0.901	1,909	5.39	3.43	5.48	3.48	5.30	3.39	0.480
Cognitive score (pca)	1,678	-0.03	0.98	-0.03	0.99	-0.02	0.96	0.846	1,905	0.00	1.00	0.01	1.02	-0.00	0.98	0.874
Education outcomes																
Math	1,683	4.66	3.77	4.73	3.81	4.59	3.73	0.743	1,912	4.85	3.85	4.93	3.89	4.77	3.82	0.680
Reading	1,682	0.88	1.12	0.90	1.15	0.85	1.10	0.556	1,911	0.92	1.14	0.95	1.16	0.89	1.12	0.507
School attendance	1,650	0.79	0.16	0.80	0.16	0.78	0.17	0.524	1,887	0.78	0.17	0.79	0.17	0.78	0.17	0.658
Socioeconomic variables																
Muslim HH	1,736	0.03		0.03		0.03		0.966	2,005	0.03		0.02		0.03		0.855
Sc/st	1,736	0.29		0.25		0.32		0.170	2,005	0.28		0.24		0.31		0.220
Block	1,738	0.67		0.72		0.62		0.289	2,005	0.67		0.72		0.62		0.317
Rural HH	1,736	0.98		0.97		0.98		0.545	2,005	0.98		0.98		0.98		0.679
N of HH members	1,736	7.76	3.41	7.82	3.45	7.69	3.38	0.555	2,005	7.77	3.49	7.82	3.52	7.73	3.46	0.667
Years schooling father	1,704	5.42	4.80	5.38	4.77	5.45	4.83	0.870	1,967	5.58	4.88	5.49	4.86	5.65	4.90	0.668

Years schooling mother	1,731	1.69	3.14	1.72	3.18	1.66	3.11	0.759	1,999	1.89	3.35	1.89	3.34	1.90	3.36	0.967
Asset index	1,703	-0.03	0.96	-0.02	0.94	-0.04	0.97	0.809	1,965	0.00	1.00	0.01	0.99	-0.01	1.01	0.799
Health care																
Institutional delivery	1,727	0.38		0.40		0.37		0.370	1,994	0.39		0.40		0.38		0.585
Health insurance	1,721	0.38		0.39		0.38		0.700	1,985	0.39		0.40		0.38		0.740
Morbidity																
Diarrhea	1,736	0.03		0.03		0.03		0.445	2,005	0.03		0.03		0.03		0.942
Improved sanitation	1,736	0.08		0.07		0.08		0.580	2,005	0.09		0.08		0.10		0.458
Biological factor																
Male child	1,738	0.45		0.43		0.47		0.173	2,005	0.46		0.44		0.47		0.306
Care taking																
Help with homework	1,722	0.16		0.17		0.15		0.475	1,984	0.17		0.18		0.17		0.762
Time physical care	1,736	45.17	25.41	43.56	24.66	46.69	26.03	0.148	2,005	45.13	25.01	43.81	24.22	46.42	25.70	0.206
School meetings	1,730	0.63		0.64		0.62		0.637	1,999	0.62		0.63		0.61		0.687
Father at home	1,735	0.87		0.87		0.86		0.698	2,004	0.85		0.85		0.85		0.802
Distance to school																
Distance to school	1,736	10.26	6.15	10.12	6.00	10.40	6.29	0.606	2,005	10.33	6.20	10.24	5.98	10.41	6.42	0.748
Nutrition																
Diet diversity score	1,736	3.87	1.18	3.89	1.21	3.84	1.14	0.568	2,005	3.88	1.17	3.90	1.20	3.86	1.14	0.646
Number of meals	1,735	3.04	1.04	3.03	1.06	3.06	1.03	0.666	2,004	3.04	1.04	3.02	1.05	3.06	1.03	0.574
Cut meals	1,733	0.80		0.80		0.80		0.796	2,001	0.79		0.79		0.79		0.908
Maternal health knowledge	1,736	0.38		0.36		0.41		0.134	2,005	0.39		0.35		0.41		0.062*
B. School level variables																
Total enrollment	107	222.21	158.24	222.15	168.29	222.28	149.30	0.997								
Class size	107	28.19	16.87	28.87	20.41	27.52	12.63	0.682								
Student teacher ratio	107	35.68	11.63	37.52	12.55	33.87	10.46	0.105								
Calories of MDM per child	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739								

This table presents baseline summary statistics as well as p-values for the difference in means tests between children in the treatment and control schools. All variables shown are child level variables from the baseline except for panel B, which shows school level variables. Standard errors are clustered at the school level. Sd: Standard deviation N: Number of observations. MDM: Midday Meal.

Table 3.A.4: Sample description and balancing test (DD balanced panel for cognitive and education outcomes)

	Panel A: DD sample								Panel B: Sample including attrited children							
		Total		Control		Treat				Total		Control		Treat		
	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means test	(1) Number of children/ schools	(2) Mean	Sd	(3) Mean	Sd	(4) Mean	Sd	(5) P Value of Difference in means test
A. Child level variables																
Anemia																
Hemoglobin	1,368	11.52	1.11	11.62	1.10	11.42	1.10	0.004***	1,729	11.52	1.10	11.60	1.10	11.44	1.10	0.010**
Any anemia	1,368	0.46		0.41		0.50		0.007***	1,729	0.45		0.42		0.49		0.019**
Mild anemia	1,368	0.20		0.18		0.21		0.140	1,729	0.19		0.18		0.21		0.133
Moderate or severe anemia	1,368	0.26		0.24		0.29		0.068	1,729	0.26		0.24		0.28		0.130
Number of anemia symptoms	1,395	1.09	1.11	1.09	1.10	1.08	1.13	0.891	1,772	1.07	1.10	1.05	1.08	1.09	1.11	0.667
Cognitive tests																
Block design	1,395	3.65	2.20	3.54	2.22	3.75	2.19	0.308	1,772	3.74	2.22	3.68	2.25	3.80	2.19	0.551
Digit span forward	1,395	4.06	0.99	4.03	0.99	4.09	0.99	0.462	1,772	4.08	1.01	4.07	1.02	4.09	0.99	0.697
Digit span backward	1,395	1.12	1.29	1.07	1.27	1.15	1.30	0.492	1,772	1.12	1.30	1.10	1.29	1.14	1.30	0.703
Progressive matrices	1,395	4.71	1.66	4.75	1.62	4.68	1.69	0.670	1,772	4.75	1.68	4.81	1.66	4.69	1.71	0.386
Day and night	1,395	5.23	3.41	5.23	3.45	5.24	3.37	0.994	1,772	5.36	3.43	5.46	3.49	5.27	3.37	0.489
Cognitive score (pca)	1,395	-0.04	0.99	-0.07	1.00	-0.01	0.97	0.569	1,772	0.00	1.00	0.00	1.02	-0.00	0.98	0.980
Education outcomes																
Math	1,395	4.58	3.78	4.61	3.83	4.55	3.73	0.894	1,772	4.83	3.83	4.90	3.86	4.75	3.80	0.698
Reading	1,395	0.87	1.12	0.89	1.15	0.85	1.09	0.698	1,772	0.91	1.13	0.95	1.16	0.87	1.10	0.453
School attendance	1,334	0.80	0.16	0.80	0.15	0.79	0.16	0.539	1,680	0.79	0.16	0.79	0.16	0.79	0.17	0.653
Socioeconomic variables																
Muslim HH	1,395	0.03		0.03		0.03		0.941	1,772	0.03		0.03		0.03		0.965
Sc/st	1,395	0.29		0.26		0.33		0.181	1,772	0.28		0.25		0.31		0.233
Block	1,395	0.65		0.71		0.59		0.251	1,772	0.67		0.71		0.63		0.417
Rural HH	1,395	0.98		0.97		0.98		0.606	1,772	0.98		0.97		0.98		0.541
N of HH members	1,395	7.66	3.23	7.75	3.34	7.57	3.13	0.434	1,772	7.73	3.35	7.78	3.40	7.68	3.31	0.618
Years schooling father	1,395	5.42	4.78	5.43	4.75	5.42	4.82	0.973	1,772	5.44	4.84	5.37	4.82	5.52	4.87	0.678

Years schooling mother	1,395	1.72	3.15	1.78	3.22	1.66	3.08	0.616	1,772	1.79	3.24	1.80	3.26	1.78	3.22	0.935
Asset index	1,395	-0.00	1.00	0.02	0.99	-0.02	1.01	0.635	1,772	-0.02	0.97	-0.01	0.96	-0.04	0.99	0.707
Health care																
Institutional delivery	1,395	0.39		0.40		0.37		0.387	1,772	0.38		0.39		0.38		0.695
Health insurance	1,395	0.39		0.40		0.38		0.701	1,772	0.39		0.40		0.38		0.592
Morbidity																
Diarrhea	1,395	0.03		0.04		0.03		0.258	1,772	0.03		0.03		0.03		0.490
Improved sanitation	1,395	0.08		0.08		0.08		0.732	1,772	0.08		0.08		0.09		0.584
Biological factor																
Male child	1,395	0.45		0.44		0.46		0.456	1,772	0.45		0.45		0.46		0.729
Car taking																
Help with homework	1,395	0.16		0.18		0.14		0.264	1,772	0.16		0.17		0.16		0.459
Time physical care	1,395	45.59	25.83	44.15	24.93	46.95	26.60	0.222	1,772	45.22	24.95	44.00	24.27	46.43	25.56	0.251
School meetings	1,395	0.63		0.63		0.63		0.937	1,772	0.62		0.63		0.61		0.732
Father at home	1,395	0.88		0.89		0.87		0.514	1,772	0.87		0.88		0.86		0.548
Distance to school																
Distance to school	1,395	10.41	6.24	10.18	6.00	10.63	6.46	0.434	1,772	10.41	6.29	10.30	6.03	10.52	6.55	0.688
Nutrition																
Diet diversity score	1,395	3.87	1.18	3.89	1.21	3.84	1.15	0.598	1,772	3.87	1.17	3.90	1.21	3.84	1.14	0.566
Number of meals	1,395	3.04	1.08	3.00	1.11	3.07	1.04	0.444	1,772	3.02	1.06	3.01	1.08	3.04	1.05	0.652
Cut meals	1,395	0.82		0.82		0.81		0.610	1,772	0.80		0.80		0.80		0.949
Maternal health knowledge	1,395	0.38		0.36		0.40		0.361	1,772	0.38		0.36		0.40		0.218
B. School level variables																
Total enrollment	107	222.21	158.24	222.15	168.29	222.28	149.30	0.997	107	222.21	158.24	222.15	168.29	222.28	149.30	0.997
Class size	107	28.19	16.87	28.87	20.41	27.52	12.63	0.682	107	28.19	16.87	28.87	20.41	27.52	12.63	0.682
Student teacher ratio	107	35.68	11.63	37.52	12.55	33.87	10.46	0.105	107	35.68	11.63	37.52	12.55	33.87	10.46	0.105
Calories of MDM per child	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739	107	68.80	23.55	69.57	21.45	68.04	25.62	0.739

This table presents baseline summary statistics as well as p-values for difference in the means tests between children in the treatment and control schools. All variables shown are child level variables from the baseline except for panel B, which shows school level variables. Standard errors are clustered at the school level. N: Number of observations. MDM: Midday Meal.

4. Nutrition Information, Anemia Testing and Feeding Practices - A Regression Discontinuity Analysis Using Data from Rural India[†]

WITH SANTOSH KUMAR AND SEBASTIAN VOLLMER

Abstract

This essay studies if revealing the anemia status of a child and informing the child's parents about the need to feed the child more iron-rich food items is effective in changing their feeding practices, the child's hemoglobin level and subsequently cognitive and education outcomes. Using a dataset from rural India we exploit the discontinuity in the provision of the nutrition information that resulted from the ethical need to inform parents about the anemia status of their child in a randomized trial, if their child's hemoglobin level was below a clinical threshold. This circumstance allows us to apply a regression discontinuity design, which has the potential to identify the causal effect of the intervention. We do not find any treatment effect on any of the tested outcomes, that is robust across specifications and indicators. Information alone, even when combined with revealing the anemia status of a child, does hence not seem to be effective in changing nutritional behavior. This might indicate that other forces than information constrain people in making rational nutrition investments. We speculate that the unavailability and inaccessibility of iron-rich food items, as well as other social and psychological factors, might be the underlying causes for why we did not find robust effects. A drawback of this study is that the given dataset only provides the statistical power to detect relatively large effect sizes.

[†] Ethical clearance: The Medical Ethics commission of the University of Göttingen approved the study. Informed written consent of the survey, including the medical tests, was taken from the parents. Acknowledgments: We thank the Foundation *fiat panis* for providing funding for this study, as well as the German Research Foundation (DFG), which provided funding within the scope of the RTG 1666.

4.1 Introduction

Despite the recognition of the importance of nutrition for human capital formation, i.e. health and education, undernutrition and particularly micronutrient deficiencies are still widespread in the world and even more so in low- and middle-income countries. Nutritional disorders at any life stage, but particularly in childhood, translate into impaired cognitive and physical development and result in a high-risk of low productivity. Therefore, it might constrain economic development (Dasgupta & Ray, 1986 and Strauss & Thomas 1998).

For a long-time, researchers have put emphasis on developing low-cost technologies to improve health and nutrition. These technologies range from oral rehydration solutions and dissemination of deworming pills to insecticide-treated bed nets, condoms to prevent HIV infections and chlorine stations at wells to purify the water. In the field of nutrition, these technologies include nutritional supplements, micronutrient-fortified products or biofortified seeds. Innovative mechanisms to distribute these technologies to the population in need have also partly been tested. Distribution channels include public facilities such as schools, health camps or agricultural extensions. Due to the high benefits and the low costs of these health technologies, a rational agent[54] would be expected to adopt these technologies from a neoclassical point of view. Though some progress has been made, adoption of these technologies is not as high as one would expect (Dupas 2011, Banerjee & Duflo 2012).

This led researchers to the hypothesis that the supply of or access to health and nutrition technologies, i.e. availability and monetary feasibility, might not be the only constraints poor households in low-income countries are facing. There is a growing strand of literature that investigates how far the lack of information constitutes an additional constraint, which limits the demand for and the proper use of these technologies (Dupas 2011b, Karlan, Ratan & Zinman 2014). The theory being that if the information constraint could be loosened by providing individuals with the required information, they would make better health and nutritional investments.

In this essay, we present new evidence on the impact of informing parents about the anemia status of their child and the need to feed their child more iron-rich food items. We investigate how this intervention affects changes in feeding practice, the children's hemoglobin

[54] Early neoclassical models define rationality in terms of utility maximization. Individuals maximize their utility given their preferences and the constraints they are facing. It is generally assumed that individual utility maximization is exclusively based on self-interest. A rational agent performs an action, e.g. makes an investment or buys a certain good, if the marginal utility is higher than the marginal costs. An individual behaving according to this pattern is called *homo oeconomicus* (Mankiw and Taylor, 2011).

levels as well as cognitive ability and education outcomes. The conveyance of this information resulted from the ethical need to inform parents about the anemia status of their child in a randomized trial and the means to eliminate the nutritional disorder if their child's hemoglobin level was below a clinical threshold. By applying a regression discontinuity design (RDD), we are able to identify the causal effect of the information intervention on the mentioned outcomes.

Anemia refers to a situation where the level of hemoglobin in the blood is low. Though anemia can have different causes[55], iron deficiency is the most common one (WHO, 2001). Iron deficiency emerges from a diet that is low in iron or when iron cannot be properly absorbed from the diet (McLean *et al.*, 2009). Anemia not only leads to low levels of physical activity (fatigue and loss of energy), but it also impairs cognitive development and work productivity. In economic terms, iron deficiency is considered to be the *costliest* micronutrient deficiency (Halterman et al. 2001, Bobonis et al. 2006). According to Horton & Ross (2003), who used data from 10 low-income countries, physical and cognitive impairment due to iron deficiency causes a loss of 4.05% at the median of a country's GDP. Globally, more than 20% of the world's population (about 1.62 billion people) are anemic (WHO, 2008).[56] The low-income population is at a high-risk for iron deficiency due to a lack of dietary diversity. Moreover, their diet generally includes a large amount of rice and wheat, which inhibits the absorption of iron due to the high concentrations of phytate in these products, and the low consumption of meat from which iron can more easily be absorbed (FAO and WHO, 2002). Depending on age and gender, anemia in India ranges from 23% to 58% (NFHS, 2015b). In the Jehanabad district of Bihar, where our study took place, the prevalence of anemia is even higher and ranges from 26% for adult males to 63% for children below the age of 5 (NFHS, 2015a).

Rigorous impact evaluation of health and nutrition information on health-related behavioral change are still rare. With respect to general health information, the findings are generally positive. Jalan & Somanathan (2008) found that households in a suburb of New Delhi started purifying their water after they were informed that their drinking water was

[55] E.g. excessive bleeding, hookworm infections or malaria (WHO, 2001). Since all children in the data set used for the analysis are dewormed at school once a year, we are quite sure that most of the anemia observed in our study comes from iron deficiency.

[56] For the identification of an anemic individual (WHO, 2008), the authors used the age and gender specific WHO hemoglobin cutoffs of mild anemia. This is for children under 5 years: < 11.0 g/dl; children 6–11 years: < 11.5 g/dl; children 12–14 years: < 12.0 g/dl; adult males: < 13.0 g/dl; adult females (non-pregnant): < 12.0 g/dl; adult females (pregnant): < 11.0 g/dl.

contaminated. Thornton (2008) studied changes in sexual practices after individuals were informed of their HIV status and finds that HIV-positive individuals are three times more likely to purchase condoms compared to HIV-positive individuals who were not informed of their HIV status. Dupas (2011a) also reported positive effects for providing teenagers in Kenya with information on the relative risk of HIV infection according to the relationship between their partner's relative age and their sexual risk behavior. In contrast, Kremer & Miguel (2007) found no improvement in worm prevention behavior after an intensive intervention on health-education at the school level. The literature on nutrition information in the context of anemia is less encouraging than the literature on general health information presented above. In an RCT conducted by Childs et al. (1997), existing doctor-parent contacts were used to convey information about breastfeeding and the link between iron and diet to parents of newborns in the UK. At the child's age of 18 months, they did not find any effect of the intervention on the prevalence of anemia in the sample. Using a relatively small sample of about 250 newborns in Brazil, Bortolini & Vitolo (2011) found the same results. They evaluated the effect of systematic dietary home counseling. However, they found a longer duration of breastfeeding and a higher consumption of iron-rich foods in the treatment group.

There is, however, one particularity of the intervention evaluated in this study that leads us to the assumption that it might be more effective than the other health and nutrition information interventions. While much of the literature relates to enhancing preventive healthcare, this study addresses the adoption of remedial behavior. In our treatment the anemia status of a child is revealed by a diagnostic test, which distinguishes our study from Childs et al. (1997) and Bortolini & Vitolo (2011). The hemoglobin testing makes the disorder explicit and the need for action immediate. In contrast, with preventive healthcare interventions, there is always the hope or assumption that only the others fall ill. Hence, the need for action might not be perceived as acute and is likely to be further limited if the required preventive actions are costly. Related to this argument, people in low-income settings typically spend a large part of their budget on remedial healthcare but the adoption of preventive healthcare is limited (Dupas, 2011b).

Our study adds to the limited literature that combines nutrition information with revealing an individual's health status. Thornton (2008) studied the effect of revealing an individual's HIV status; however, she examined the adoption of health behavior that reduces the transmission of the disease rather than individual treatment. She found that individuals buy more condoms when they are diagnosed as HIV positive compared to HIV positive

individuals that did not learn their HIV status, indicating that individuals change their health behavior to adapt to their health conditions after learning about their status. Cohen et al. (2015) examined the effect of purchasing anti-malaria medicine after using a rapid diagnostic test for malaria. They found that some patients did indeed respond to revealing their own health status, however, half of the patients were buying anti-malaria medicine, despite having tested negative for malaria.[57] The paper that relates best to our study, and that also reveals the anemia status of a child, is from Luo et al. (2012). In a RCT, Luo et al. (2012) informed the parents of Chinese elementary school children about the anemia status of their child and present strategies in addressing their child's nutritional deficiency (eating balanced meals, including iron-rich products, counseling a doctor or taking iron supplements). The information was either conveyed by letter, by a single or by multiple face-to-face information session(s). The different information interventions did not have any impact on hemoglobin levels or anemia rates.[58] Using a sample of rural Indian households, we add external validity to the existing literature. An additional innovation of this study is not only the assessment of health and nutrition, but also of productive outcomes such as cognition and education. To our knowledge, this is also the first time that RDD was applied in the context of anemia and nutrition information.

For a similar, but somewhat simpler information intervention and a similar age group, we can confirm the findings from Luo et al. (2012) that information alone does not seem to change nutrition related behavior, even when combined with revealing a nutritional disorder of a child. Neither the hemoglobin levels nor the cognitive and education outcomes were affected by the treatment, which is little surprising since any effect on this outcomes would be based on a change in feeding practices in the first place. Tough our study might be suffering from too little statistical power, with is a drawback of the study, this finding might indicates that other forces, despite the lack of information, limit people in making rational health

[57] The case of anemia evaluated in this essay is different from HIV and malaria infections, as anemia is a non-communicable disease and hence not directly transmittable. This means that the adverse effects, though being potentially strong for the individual, do not involve externalities. Malnutrition only indirectly exhibits externalities, because it weakens the immune system, which is then less able to resist communicable diseases.

[58] The revealing of the individual health status also partly relates to the debate on the effectiveness of screening. Screening is the counseling and testing for certain diseases or body disorders of a large population before the individual notices any symptoms. The purpose of screening is the early diagnosis and treatment of a disease and hence the prevention of more severe health consequences (Wilken *et al.*, 2012). Screening is commonly done for diabetes, hypertension, tuberculosis, cancer and HIV but more so in high-income countries. It is generally directly linked to the administration of medication plus the promotion of a change in lifestyle. In our study, the intervention after screening is limited to nutrition information, i.e. a change in lifestyle. We are not aware of any studies that test the effect of screening on behavioral changes towards the adoption of health technologies and better health behavior. The screening debate focuses more on the costs (financial, health psychological) of the screening process itself.

investments. We speculate that the unavailability and inaccessibility of iron-rich food items might be one reason for not detecting a treatment effect. Comparing our findings with the existing literature on human health behavior, other social and psychological factors are also likely to play a role. We therefore make a case for incorporating a complex set of factors that influence human behavior, including insights from sociology and psychology, into the design of policies that aim at changing health and nutrition behavior.

The remainder of the paper is structured as follows: in section 4.2, we describe the treatment, the dataset and the methodological approach. In section 4.3, we present the empirical specification and we describe the results in section 4.4. In section 4.5, we present robustness checks and we discuss our findings and conclude in section 4.6.

4.2 Treatment, Data and Methodological Approach

4.2.1 Treatment

In the scope of the data collection described in essay 2 of this dissertation (chapter 3), second grade school-children were tested for their hemoglobin value. The testing was performed with an on-site hemoglobin measurement device directly in the village or at the children's homes. Irrespective of the medical results, parents were informed about the hemoglobin level of their child. Only in cases where a child's hemoglobin value was below a defined threshold, parents were additionally informed about adequate treatment. The hemoglobin thresholds applied are the official WHO cutoffs for moderate and severe anemia for children aged between 5 and 11 years (WHO, 2011). Following the recommendation of a local physician, the information about adequate treatment given to parents was as follows:

> **8 ≥ Hemoglobin level ≤ 10.9 g/dl (moderate anemia):**[59] Recommendation of a more diverse diet for the child, especially the consumption of green leafy vegetables and meat if the household consumes meat.
>
> **Hemoglobin level < 8 g/dl (severe anemia):** In addition to the nutritional advice, it was highly recommended to bring the child to the next healthcare facility. If the hemoglobin level was below 6 g/dl, the interviewer team ensured that the child was quickly taken to the next healthcare facility.

[59] The hemoglobin measurement is accurate to one decimal place. Similarly, the WHO cutoffs are also defined to one decimal place.

We found very few observations (13 children, less than 1% of the sample) with a hemoglobin score below 8 g/dl and only one child with a hemoglobin score below 6 g/dl. That is why we limit the impact analysis to the information given for a hemoglobin value ≤ 10.9 g/dl. We exclude observations with a hemoglobin value of < 8g/dl at the baseline.

Informing parents about their child's anemia status and recommending dietary changes were driven by the ethical need to share with the parents any findings from our study that might indicate a direct risk to the health of a child in our sample.[60] The evaluation of this treatment can thus be considered a *natural experiment*.[61] [62]

4.2.2 The Discontinuity in Treatment Assignment

The causal identification of the effects of revealing the presence of a nutritional disorder and of giving nutritional advice is complicated in most of the settings due to the potential presence of a selection bias: Generally it is not exogenously determined who received the nutrition information, but rather individuals ask for diagnostic tests and seek nutritional advice either because they suffer from very poor health conditions or they are particularly concerned and careful with their own health. These characteristics might lead to biased estimates in simple regressions due to the potential presence of a selection bias. To overcome the selection bias, in this study we make use of the discontinuity in the conveyance of the nutrition information and of the anemia status that is a result of the data generating process described above.

The sharp cutoff point for moderate anemia, which clearly determines if the information treatment was given, creates room for the application of a *regression discontinuity design* (RDD).[63] If the cutoff point is indeed arbitrary, was not manipulated by the entity who assigns the treatment and the forcing variable (also called *assignment* or

[60] When conducting research with human subjects and especially with vulnerable populations, such as children, it must be ensured that the benefits of the research outweigh the risks (Medical Research Council, 2004). The treatment of providing information on nutrition and recommending dietary changes was implemented to maximize the benefits of the children involved in the survey.

[61] Natural experiments are categorized as a subgroup of quasi-experiments and are defined as situations where individuals are assigned to the treatment and control groups because of a natural event. The researcher does not influence the assignment to the treatment or control groups, as is the case in a RCT, but merely functions as an observer (DiNardo, 2008).

[62] For our main analysis, we include both the control and treatment groups of the RCT described in essay 2, since we were expecting to lose too much power if we restrict our analysis to the control group alone. It might be the case that our estimates are downward biased due to the inclusion of the treatment group of the RCT. We therefore perform the analysis with the control group only in a robustness check in chapter 3.5.

[63] The idea of the RDD method dates back to Thistlethwaite & Campbell (1960).

running variable), i.e. the hemoglobin score in this study, was not manipulated by the target individuals – assumptions that are investigated in the next section – RDD is a valid estimation method which will causally identify the treatment effect at the cutoff points.[64] If the assumptions are fulfilled, the stochastic error component is continuously distributed over the forcing variable, such that around the thresholds the assignment to the treatment is as good as random. Observations further away from the cutoff are, however, likely to be different from each other and hence do not constitute a valid comparison group. This is because socio-economic characteristics and the general health status of children are possibly correlated with their hemoglobin level. Unhealthy and poorly nourished children in poor households are more likely to have low hemoglobin levels, whereas healthier and better nourished children tend to be systematically better off and have higher hemoglobin levels. However, individuals close to the threshold only differ in treatment status but not in other characteristics including their underlying health. Thus, average outcomes just above the cutoff can be used as valid counterfactuals for average outcomes just below the cutoff (Lee and Lemieux, 2010).

4.2.3 Validity of the RDD

A valid RD design depends on two conditions. First, the cutoff, which determines the treatment, needs to be arbitrary and second, neither the individual who assigns the treatment nor the targeted individual should be able to precisely manipulate the forcing variable (Hahn et al. 2001, Lee & Lemieux 2010, Imbens & Lemieux 2008).

Arbitrary cutoff – The cutoffs for anemia were set in terms of standard deviations from the mean of a hemoglobin distribution of a reference population. The reference population refers to a healthy population of the same gender and age group, living at the same altitude (WHO, 2001). Cutoffs are thus not based on a physiological phenomenon but rather on a statistical convention. Which implies that a child is not exposed to a sharp health risk increase between a hemoglobin value of 11 g/dl and 10.9 g/dl. This arbitrariness is explicitly pointed out in the report of the study group on iron deficiency anemia of the WHO in 1959, when the anemia cutoffs were defined.

> *„To detect and evaluate the anemia problem of a community, it is necessary to have standards of reference, even if they be somewhat arbitrary,...“* (WHO 1959, p. 4)

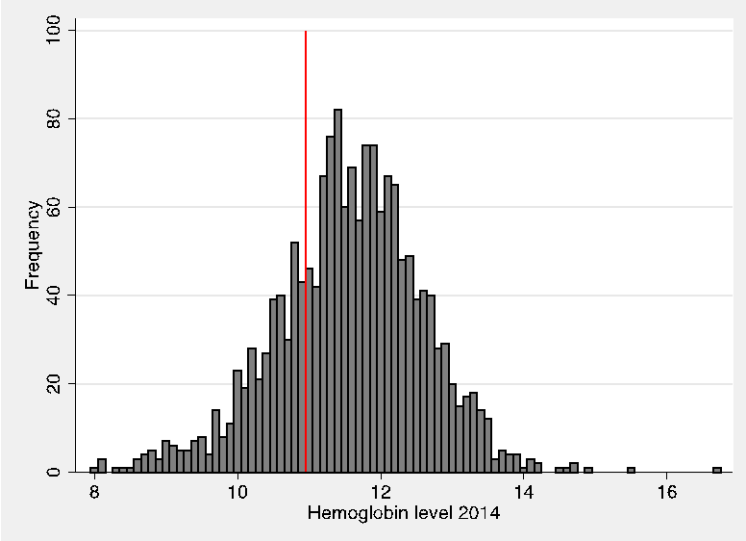
[64] The identifying assumptions have been formalized by Hahn et al. (2001).

Moreover, as investigated by Rippmann et al. (1997), the on-site device by which the hemoglobin level was measured, the HemoCue® Hb 301 (AB Leo Diagnostics, Helsinborg, Sweden), exhibits some variability. Hence, imprecision of the measurement device results in some degree of randomness.

Inability to manipulate assignment to the treatment – The second condition, the inability of manipulation, is a requirement that ensures that there is no selection bias. A selection bias could have been induced either by the entity that assigns the treatment, i.e. the medical staff or by the targeted individual, i.e. the parents of the tested children. However, both possibilities have very unlikely occurred. The cutoff levels and also the occurrence of the information treatment were unknown to parents and children and they did not have any incentive to manipulate the hemoglobin value. Furthermore, it is impossible to adjust feeding practices in a way that hemoglobin levels can be precisely determined. Since no benefits for the medical staff were involved in the conveyance of the treatment and the required effort for the communication of the nutrition information was very low, we also do not see any incentive for the medical staff to have manipulated a child's assignment to the treatment. Medical personnel were additionally supervised through spot-checks and no irregularities were observed. However, we cannot fully rule-out the possibility that medical staff sometimes forgot to convey the nutrition information. At the same time, we can also not rule-out the possibility that parents of non-anemic children sometimes took-up the treatment, when the nutrition information was addressed to the parents of anemic children. These circumstances imply that assignment to the treatment might not be a deterministic function of the forcing variable but instead, the probability of assignment to the treatment increases at the cutoff. In such a setting, a *fuzzy* RDD can potentially be applied. Unfortunately, we do not know which individuals were affected by this potential imprecision in the conveyance of the treatment, such that we are not able to perform a *fuzzy* RDD and conduct a *sharp* RDD analysis. This issue is further discussed in section 4.5.2.

Though randomness around the thresholds cannot be fully tested, there are some empirical tests that can provide suggestive evidence. First, the non-manipulation of the hemoglobin level is supported by the histogram (Figure 4.1). If individuals had precisely manipulated the forcing variable, one would see a gap in frequencies around the cutoff (marked with a vertical red line in figure 4.1), e.g. we would observe very few children with a hemoglobin level of 10.9 but many with a level of 11 g/dl. There is no such gap in frequencies in figure 4.1, indicating that manipulation around the cutoff is not an issue.

Figure 4.1: Distribution of baseline hemoglobin values



Second, if the treatment was indeed as good as randomly assigned around the threshold, baseline covariates should be equally distributed just above and below the cutoff (Lee and Lemieux, 2010). [65] This balancing test is presented in table 4.1. It shows the baseline means of the outcome and control variables (variables are described in the next section) for all observations with hemoglobin values in 2014 between 10.5 and 10.9 g/dl (left side of the cutoff) and between 11.0 and 11.4 g/dl (right side of the cutoff), i.e. just above and just below the cutoff. The last column of the table provides the p-values for the t-test of equality of the means, clustered at the school level. Except for the hemoglobin value, which by construction is lower below and higher above the threshold, and the share of mothers that help their child with their homework, which is more likely to have differed by coincidence, all means are similar and none of the other p-values of the t-test for differences in means above and below the cutoff are statistically significant, indicating randomness around the cutoff.

[65] Hahn et al. (2001) show that continuity in the assignment variable is sufficient to obtain unbiased estimates. Therefore, the equality in means of individuals above and below the threshold is not required, however, it is likely to be the case within a small bandwidth around the cutoff point.

Table 4.1: Balancing table (Hemoglobin sample)

	Left side 10.5 till 10.9 g/dl hemoglobin			Right side 11.0 till 11.4 g/dl hemoglobin			
	(1) Mean	SD	N	(2) Mean	SD	N	(3) P-value
Feeding practices							
Dietary Diversity Score	3.80	1.20	204	3.88	1.17	313	0.446
Frequency of meat consumption	1.83	0.91	203	1.77	0.81	313	0.393
Frequency of green veg consumption	3.38	1.00	200	3.44	1.06	309	0.558
Hemoglobin							
Hemoglobin	10.71	0.14	204	11.23	0.14	313	0.000***
Number of anemia symptoms	1.01	1.05	204	1.03	1.09	313	0.858
Cognition							
Block design	3.76	2.21	199	3.46	2.19	307	0.133
Digit span forward	3.98	1.00	200	4.01	0.95	307	0.666
Digit span backward	0.96	1.29	200	0.98	1.22	307	0.879
Progressive matrices	4.87	1.73	200	4.84	1.46	307	0.881
Day and night	5.04	3.28	200	5.09	3.33	307	0.865
Cognitive index	-0.07	0.96	199	-0.10	0.92	307	0.721
Education							
Math	4.30	3.78	200	4.38	3.57	307	0.807
Reading	0.73	0.99	200	0.75	1.04	307	0.809
School attendance	0.78	0.16	195	0.80	0.16	295	0.196
Covariates							
Treatment group from intervention of essay 2	0.52	0.50	204	0.55	0.50	313	0.552
Muslim HH	0.03	0.17	204	0.02	0.14	313	0.451
Sc/st	0.32	0.47	204	0.31	0.46	313	0.804
Block	0.71	0.45	204	0.66	0.47	313	0.270
Rural HH	0.98	0.16	204	0.98	0.13	313	0.492
N of HH members	7.39	3.04	204	7.53	3.22	313	0.618
Years schooling father	4.84	4.86	201	5.13	4.73	307	0.510
Years schooling mother	1.44	2.89	203	1.46	2.93	311	0.940
Asset index	-0.15	0.97	200	-0.10	0.89	309	0.558
Institutional delivery	0.37	0.48	201	0.33	0.47	313	0.272
Health insurance	0.42	0.49	202	0.35	0.48	310	0.143
Diarrhea	0.04	0.21	204	0.05	0.21	313	0.841
Improved sanitation	0.07	0.26	204	0.07	0.26	313	0.998
Male child	0.40	0.49	204	0.45	0.50	313	0.260
Help with homework	0.10	0.30	201	0.19	0.39	311	0.006***
Time physical care	47.25	28.92	204	44.83	22.62	313	0.287
School meetings	0.66	0.48	203	0.65	0.48	313	0.878
Father at home	0.88	0.33	203	0.89	0.31	313	0.613
Distance to school	10.03	6.28	204	10.33	6.02	313	0.583
Number of meals	3.09	0.92	204	3.07	1.10	313	0.874
Cut meals	0.82	0.38	204	0.82	0.38	313	0.944
Iron supplementation	0.16	0.37	201	0.21	0.41	309	0.151
Maternal health knowledge	0.36	0.48	204	0.37	0.48	313	0.914
Total enrollment	252.83	153.70	204	262.88	153.56	313	0.467
Class size	32.62	16.37	204	34.26	17.50	313	0.286
Student teacher ratio	38.89	11.76	204	38.46	11.52	313	0.679
Calories of MDM per child	66.96	20.15	204	69.05	22.85	313	0.288
Iron in MDM per child	0.75	0.29	204	0.78	0.30	313	0.297

This table presents baseline summary statistics as well as p-values for difference in means t-tests between children just above and just below the cutoff of 10.9 g/dl. All variables shown are child level variables from the baseline.

Standard errors are clustered at the school level. SD: Standard deviation, N: Number of observations, MDM: Midday Meal.

4.2.4 Sampling and Data

Data collection - From November 2014 until January 2015, a health survey, which included a diagnostic test for hemoglobin, was carried out among 2000 school-aged children in the two blocks of Modanganj and Kako, district Jehanabad, located in the Indian state of Bihar, an area where the prevalence of anemia is high. In 2014, we found a prevalence of anemia among second grade students of about 45%. From a list of 228 government-funded schools that exist in the two blocks, a simple-random sample of 108 schools was drawn and on average, 20 children per school from the second grade were chosen for anemia testing. The sample is therefore representative of second grade students in government-funded schools in the two blocks. We specifically sampled children from the second grade because they are at the beginning of a phase of rapid brain development, since the frontal lobes experience spurts of development between the ages of 7 and 9 (Anderson 2002, Hudspeth & Pribram 1990, Thatcher 1991).

After parental consent was taken and they were interviewed on child health and feeding practices and data on their household socioeconomic characteristics was collected, trained medical personnel performed medical checks with the children, including a diagnostic test for hemoglobin values.[66] By taking a small drop of blood from a child's finger, hemoglobin levels were assessed using an on-site hemoglobin measurement device called HemoCue® Hb 301 (AB Leo Diagnostics, Helsinborg, Sweden). Additionally, cognitive and education tests were performed with the children at the school level. From August until October 2016, i.e. about two years after the intervention, a follow-up survey was conducted that collected the same data for the same children. Since two years had passed between the treatment and the collection of the follow-up data, we can only investigate medium-term changes in response to the intervention but no immediate reactions.

Outcome variables – As direct outcomes, data on *feeding practices* was collected by a food frequency table. The information in the food frequency table are based on parental reporting. Three different indicators for feeding practices were developed from this table. First, a dietary diversity score (DDS), which was calculated by summing up the number of

[66] There were a negligible number of six children (less than 1%) that refused the hemoglobin test. In contrast, there was a high demand for the medical check, also from the parents of children that were not in our sample, as many households perceived the health survey as a free healthcare service.

food groups represented in the child's diet. Similar to Torheim et al. (2003) and Kennedy et al. (2010) and based on data availability, the following food groups were included: Legumes, fruits, vegetables/green leafy vegetables, eggs, meat/poultry/fish, milk/dairy products.[67] If parents reported that their child consumes an item from one of these food groups, at least several times per month, the food group was assigned the value of one. Values for all food groups were summed up, such that the DDS ranges from zero (no item from any food group is consumed) to six (at least one item from each food group is consumed). FAO (2007) reviews studies that show that DDSs are valid indicators for the adequacy of micronutrient (and macronutrient) intake. Since the hemoglobin level was always measured and the nutrition information was always given after the parents were interviewed for their feeding practices, the possibility that the feeding practices were reported biased is minimized. Second, feeding practices are measured by the frequency of the consumption of food items that are available in the study region and that are supposed to contain a relatively high level of iron. These food items are green leafy vegetables (one indicator) and meat (second indicator). We create categorical variables for the frequency of the consumption of these two food groups that range from 1 (*the child never consumes an item from this food group or less than once per month)* to 5 (*the child consumes an item from this food group daily).*

As a secondary outcome, the *anemia* of the child was measured in two ways. First, the change in hemoglobin level was used as a continuous variable. Second, we collected data on the presence of common anemia symptoms (easily fatigued and loss of energy, shortness of breath particularly with exercise, a bluish color of their lips and nails, headache and shaking chills) as reported by the child's mother. The number of anemia symptoms was summed up. Hence the variable ranges from 0 to 5.

As a more indirect outcome, *cognitive ability* was measured by five different cognitive tests (forward digit-span, backward digit-span, block design, Stroop-like day-and-night test and progressive matrices). For a detailed description of the cognitive tests, please refer to chapter 3.3.2. These tests specifically assess *executive functions*, which are needed for purposeful goal-directed activities including inhibition, planning and organizing and working memory. Executive functions are supposed to be mediated in the frontal lobes, the brain region that, as stated above, experiences spurts of development at the age of the sampled children (Lezak 1995, Anderson 2001, Salimpoor & Desrocher 2006).

[67] There exists no international consent on which food groups should be included in a DDS and how these food groups are defined (FAO, 2007).

Education outcomes are also categorized as more indirect outcomes. They were assessed by math and reading tests as well as the child's school attendance. Please refer to chapter 3.3.2 for a detailed description of the education outcomes. Reading skills were tested on a scale from 0 to 4, ranging from *child does not recognize letters* to *child fluently reads a short story*. For the reading assessment, the materials from the Annual Status of Education Report (ASER, 2014), developed by the Indian non-governmental Organization *Pratham,* were used. For the math assessment, the material from ASER (2014) was used as basis, but extended to 13 different exercises at the baseline and 15 at the endline, ranging from *child does not recognizes one-digit numbers* to *child is able to solve advanced division problems*. Finally, the school attendance of the child for the year before the follow-up survey was recorded from the official school attendance register.

To ensure similar test conditions for all children, cognitive and education assessments were performed one by one by prospective female teachers. In order to minimize disturbance, they were performed in a separate room whenever possible.

Since the unit in which cognitive ability and education outcomes are measured is arbitrary, we standardized the test scores of these outcomes by subtracting the mean from the score at the baseline and dividing by the standard deviation at the baseline for the whole sample for each test. Hence, a standardized cognition score of 0.5 would mean that the student scored 0.5 standard deviations higher than the mean in 2014.

4.3 Empirical Specification

To estimate the discontinuity at the cutoff point, we follow the standard procedures suggested by Lee & Lemieux (2010), Imbens & Lemieux (2008) and the recent publication by Cattaneo et al. (2016).

Functional form and regression equation – Estimating unbiased treatment effects using RDD, crucially depends on the functional form chosen to model the relationship between the forcing and the outcome variable. In our preferred specification, we estimate the regression function using a *pooled normalized local linear regression*. In a local linear regression, a straight line is fitted to the data within a predefined window with bandwidth h (i.e. locally) around the cutoff point. The choice of the window width – h – is described below. The treatment effect is modeled by a jump in the function at the cutoff point. We allow

the regression function to differ at both sides of the cutoff by including an interaction term between X, the forcing variable, and D, the treatment dummy, but estimate both regression lines simultaneously, i.e. pooled. For convenience in the interpretation, we subtract the values of the forcing variable from the value of the cutoff point – c – (i.e. we normalize the forcing variable), thereby the treatment dummy – D – yields the treatment effect. We impose a triangular kernel, which gives more weight to the observations close to the cutoff. The normalized pooled regression function is as follows

$$\Delta Y_i = \alpha_i + \tau D_i + \beta_1(X_i - c) + \beta_2 D_i(X_i - c) + \varepsilon_i \quad \text{where } c - h \leq X \leq c + h \quad (1).$$

ΔY represents the change from 2014 to 2016 of the different outcome variables (feeding practice, hemoglobin, cognitive ability and education outcomes). We use the change in the outcome to control for the initial level of the outcome variable. In case there is high correlation in the data over time, it also increases the precision of our estimates. α_i is the intercept of the function on the right side of the cutoff. β_1 is the slope of the function on the right side and β_2 is the difference between the slopes on the left and right side of the cutoff. ε_i represents the error term. D_i is a dummy that takes on the value of one if a child's hemoglobin level was $\leq$ 10.9 g/dl in 2014 and 0 otherwise and indicates that the parents probably received the nutrition information. Hence D_i is defined as

$$D_i = \begin{cases} 0 & \text{if } X_i > 10.9. \\ 1 & \text{if } X_i \leq 10.9. \end{cases}$$

τ represents the treatment effect, e.g. the size of the discontinuity at the cutoff point and hence the main coefficient of interest. We estimate equation (1) within a narrow window – h –, also called bandwidth, around the cutoff point. We apply robust standard errors clustered at the school level.

Choosing the *bandwidth* – h – is essentially a trade-off between precision and bias. Intuitively, one might want to include only observations very close to the cutoff, assuming that those are indeed random and hence provide unbiased estimates. Only including values very close to the cutoff, however, reduces the number of observation and thereby the precision of our estimation. The bandwidth must therefore be chosen in a way that is wide enough to get precise estimates, but narrow enough to minimize the bias (Lee and Lemieux, 2010). In the last 50 years, a number of different approaches for bandwidth choices have been

developed and applied (see Cattaneo & Vazquez-Bare 2016, for an overview of the different methods). [68] In our preferred specification, we use the data driven method that minimizes the mean squared error (MSE) for the local linear regression point estimator originally proposed by Imbens & Kalyanaraman (2012) for independent and identically distributed data. This method was further developed by Calonico et al. (2014) and Calonico et al. (2016) for clustered data (henceforth CCT). Because our data is clustered at the school level we apply the CCT approach for the case of local liner regression.

Since only data points close to the cutoff are included in the analysis, estimates only apply to individuals with hemoglobin values in 2014 that are close to the 10.9 g/dl cutoff. Hence we measure the average treatment effect for the population close to the cutoff. Only if the treatment effect would be homogeneous, i.e. it does not vary across hemoglobin values, would the estimates apply to the sample as a whole.

As a robustness check, we test our results for different bandwidths. We also impose a rectangular kernel that weighs all observations within the bandwidth equally and is hence equal to estimating a conventional linear regression line. In another specification, we include control variables (control variables are described in Text 4.A.1 in the appendix and are listed in the balancing test table 4.1). If randomness at the threshold indeed applies, our results should not be sensitive to the inclusion of covariates. Furthermore, we impose a polynomial of order two on all data points within the bandwidth selected by the CCT procedure. Finally, we show results for different order polynomials for all data point, i.e. globally and not only for a small bandwidth around the cutoff.

4.4 Results

4.4.1 Graphical Illustration

Figure 4.2 illustrates the potential discontinuities by plotting the change in our outcome variables from 2014 to 2016, against the normalized hemoglobin values in 2014. Due to the normalization of the forcing variable, point 0 at the x-axis is equal to a hemoglobin value of 10.9 g/dl. Section A of figure 2 shows discontinuity graphs for the feeding practice indicators

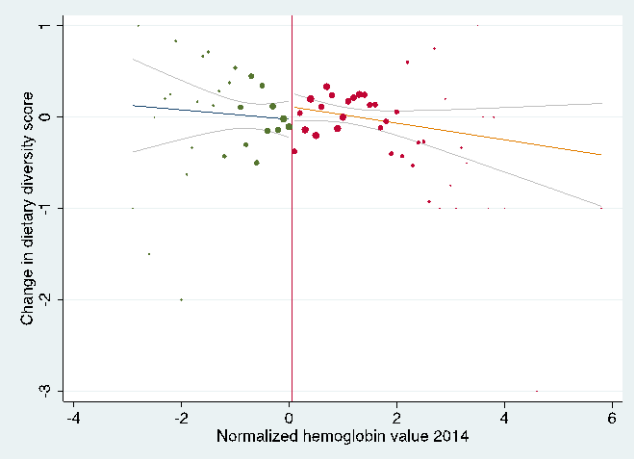
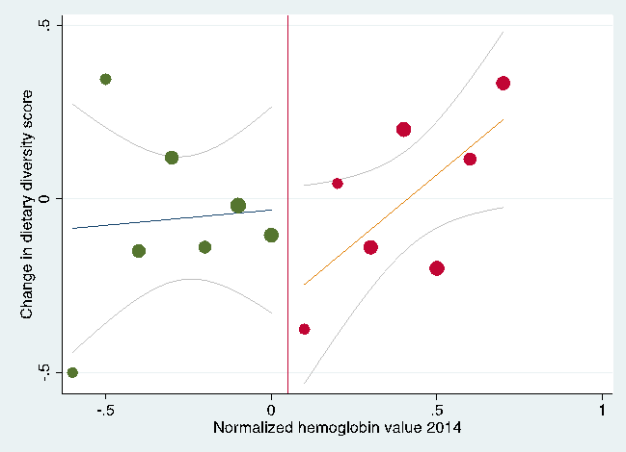
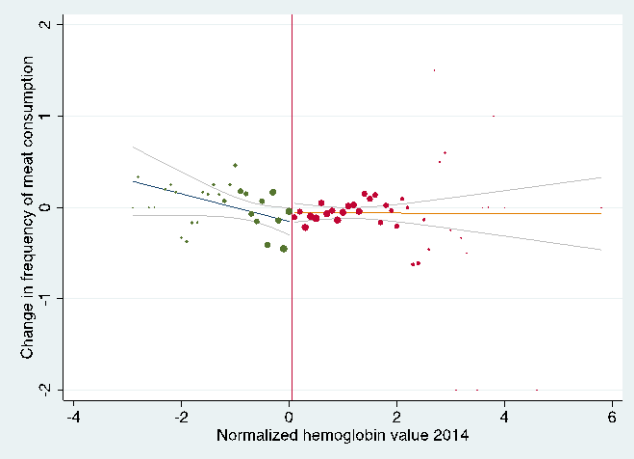
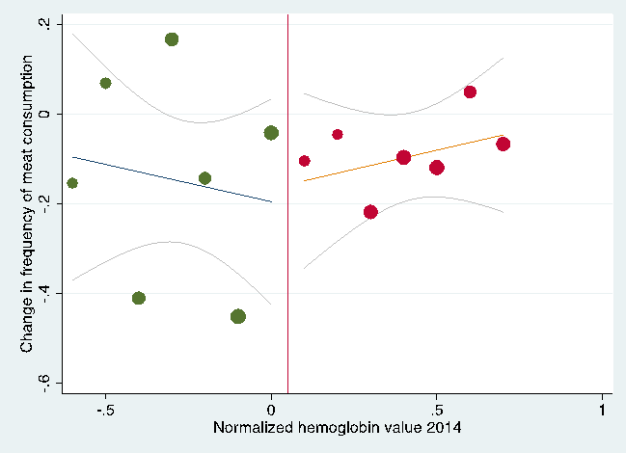
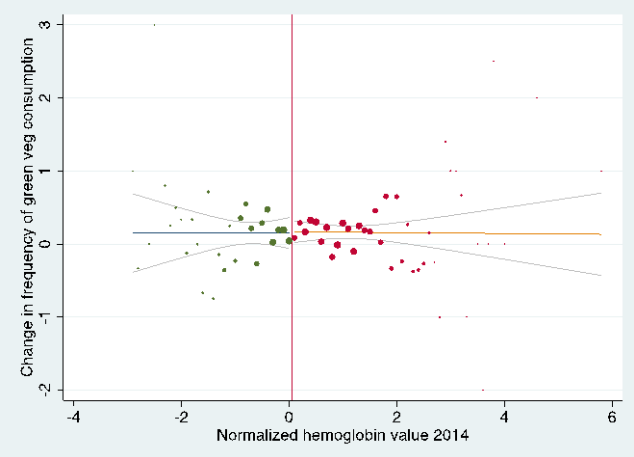
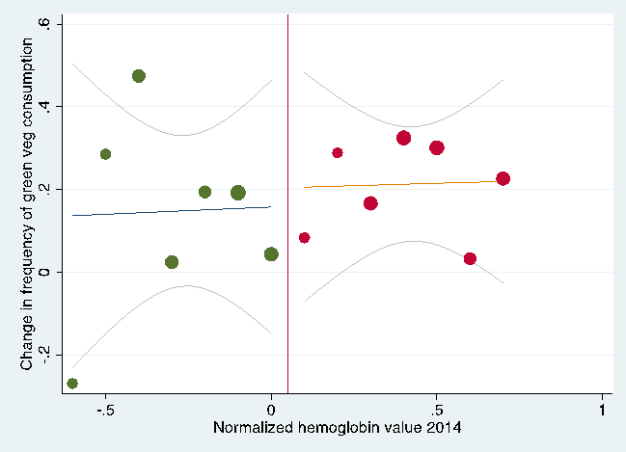
[68] Ad hoc methods use either the full support of the data or choose a smaller bandwidth arbitrarily. Ad hoc methods hence lack objectivity and comparability across studies. Apart from the approach used in our analysis, another data driven approach, proposed by Ludwig & Miller (2007), can be found in the literature. This approach is the "leave-one-out" cross validation procedure. It has been criticized because it uses the whole support of the data to the left and the right of the cutoff (Imbens & Kalyanaraman 2012, Imbens & Lemieux 2008).

(the dietary diversity score, the frequency of meat consumption and the frequency of consuming green leafy vegetables), section B for anemia outcomes (hemoglobin levels and the number of anemia symptoms) and section C for cognitive and education outcomes (5 different cognitive tests, math and reading test scores and school attendance rate). For illustrative reasons, changes in outcomes are averaged over each discrete value of the forcing variables and plotted against the respective discrete values of the normalized hemoglobin values from 2014. To represent the density of the observations, the size of the dots in the graphs represents the number of observations within each discrete hemoglobin value. A linear regression line is fitted to the data points and the grey line shows the confidence intervals. In panel A, we show graphs for all data points (globally) and in Panel B for observations within the bandwidth that is selected by the CCT procedure.

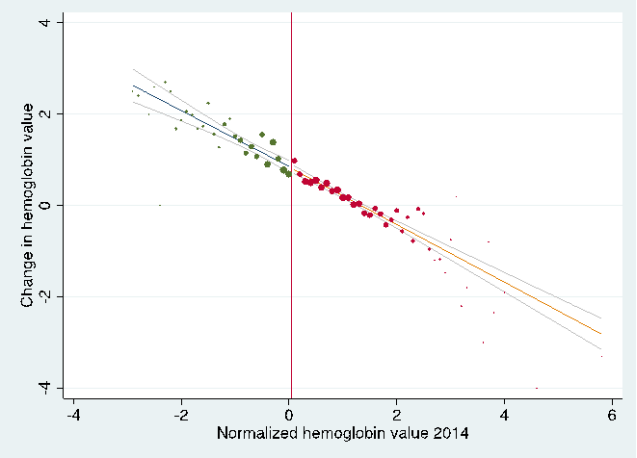
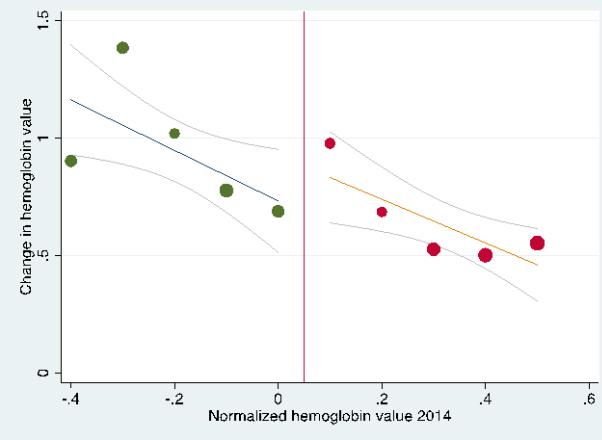
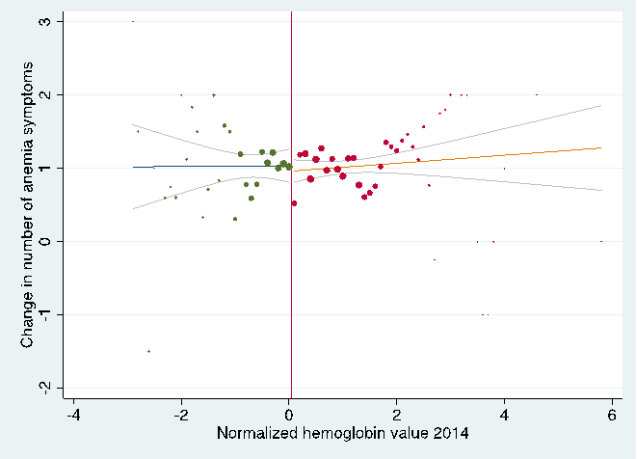
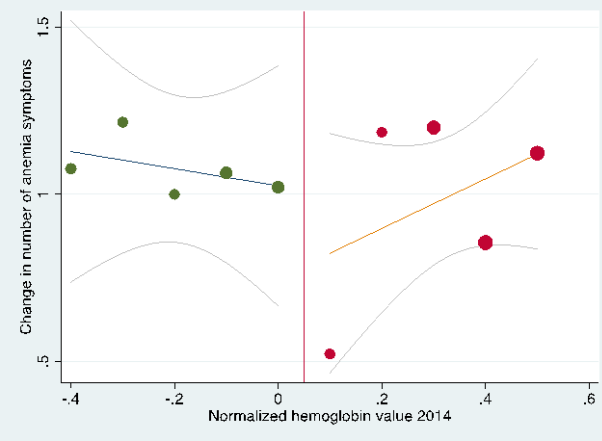
If the information treatment was effective in improving the tested outcomes, one would see a jump at point 0 of the x-axis. For outcomes where we expect an increase due to the treatment (all outcomes except the number of anemia symptoms), the regression line to the left of the cutoff would then be above the regression line to the right of the cutoff, and vice versa for outcomes where we would expect a decrease due to the treatment (anemia symptoms). In panel A in none of the graphs can a discontinuity at the cutoff be detected, instead all data points evolve smoothly at the cutoff, indicating that the information treatment did not affect any of the tested outcomes. When focusing on observations close to the cutoff point (Panel B), no jump can be detected for most outcomes. There might be discontinuity for some of the cognitive and education outcomes such as in the backward digit-span test, the block design test, the cognitive index test and school attendance. Furthermore, there is quite some variability in the data in that the confidence intervals are relatively large and the observed discontinuities in cognitive outcomes also point in an unexpected direction.

Figure 4.2: Discontinuity graphs

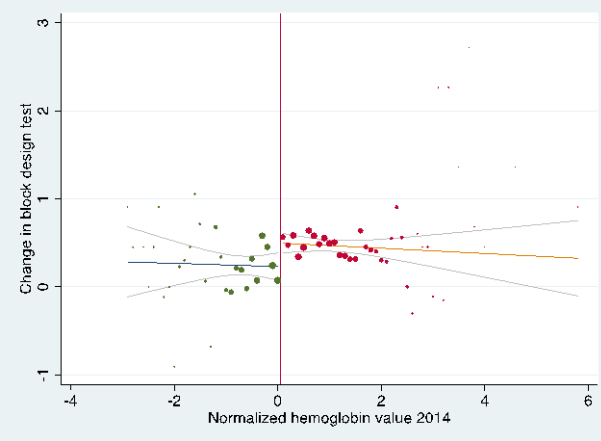
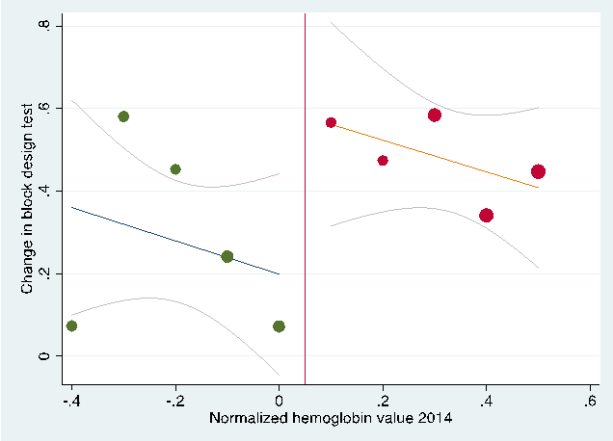
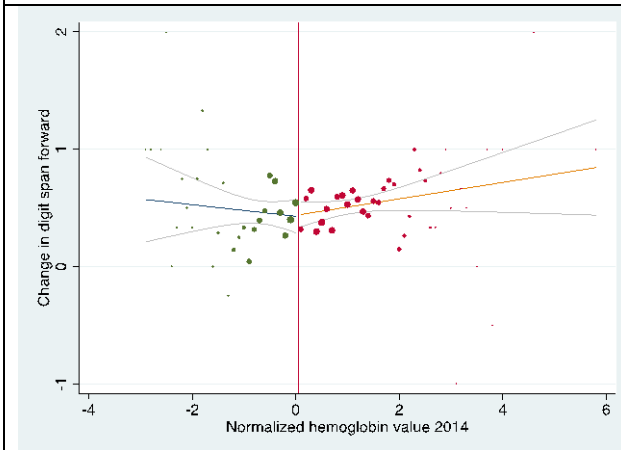
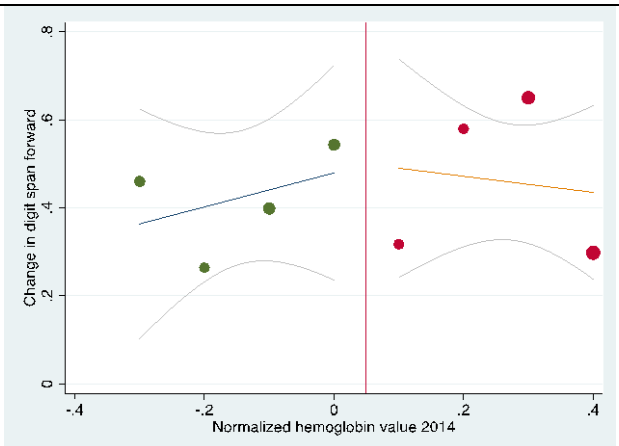
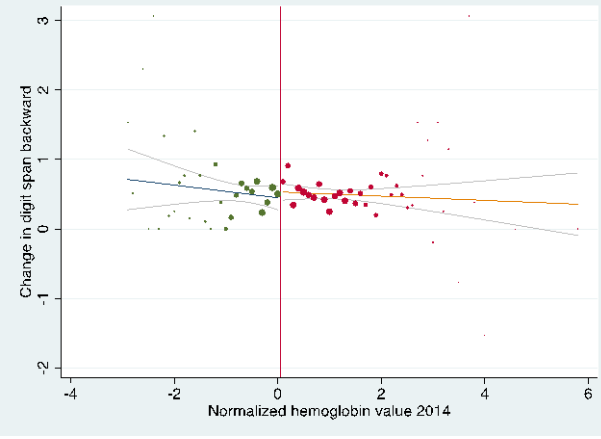
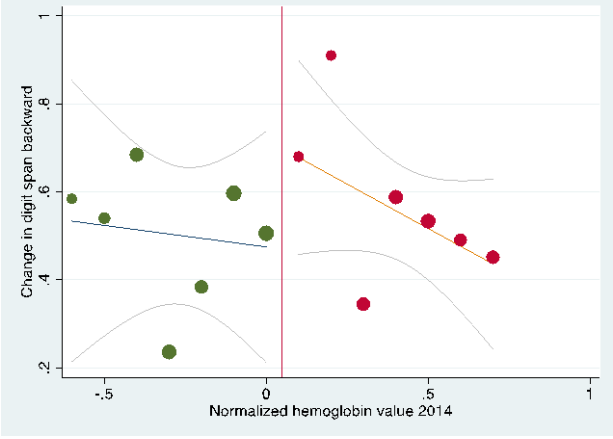
A. Feeding practices

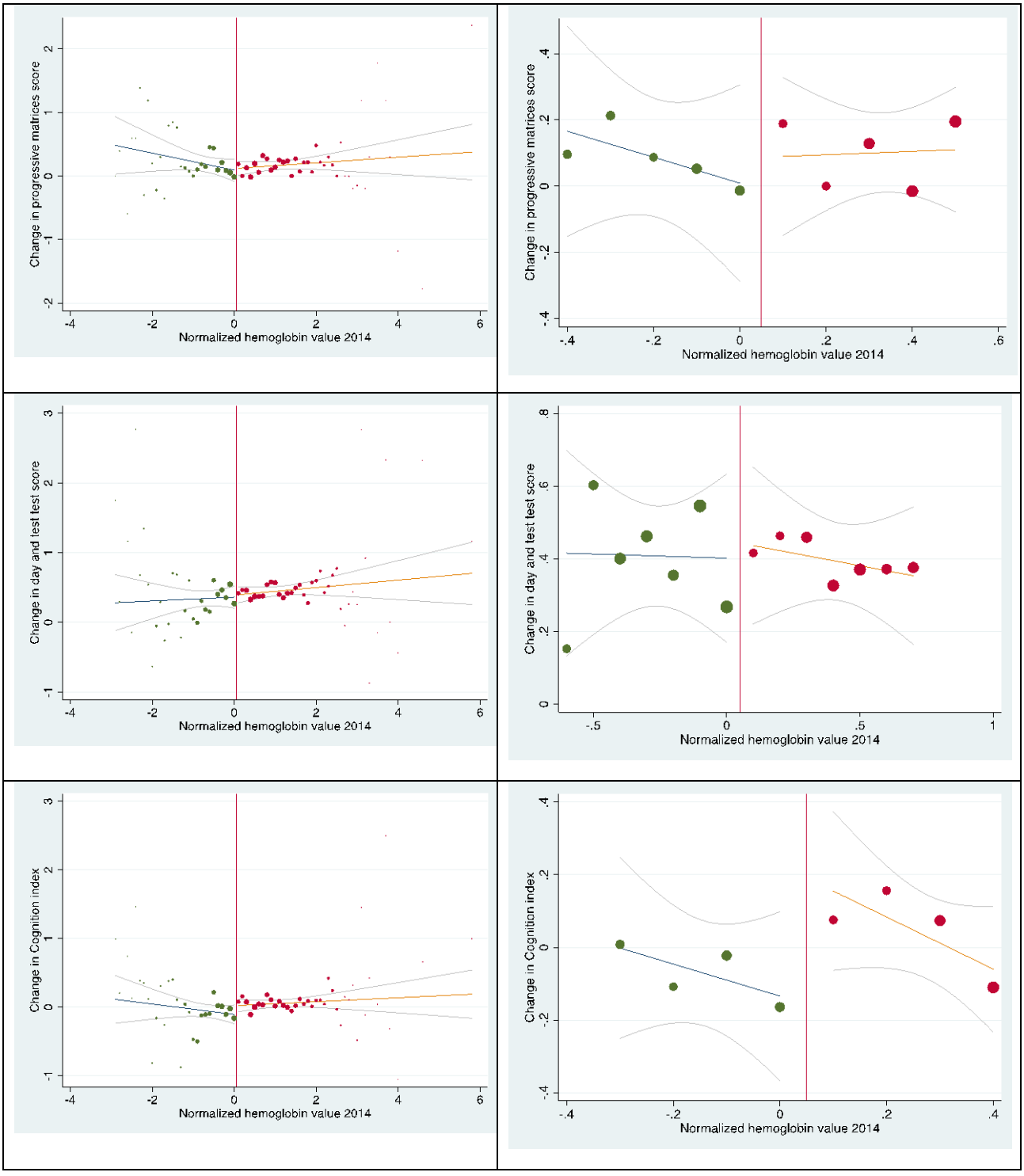
Panel A: Globally	**Panel B: Close to the cutoff**
	
	
	

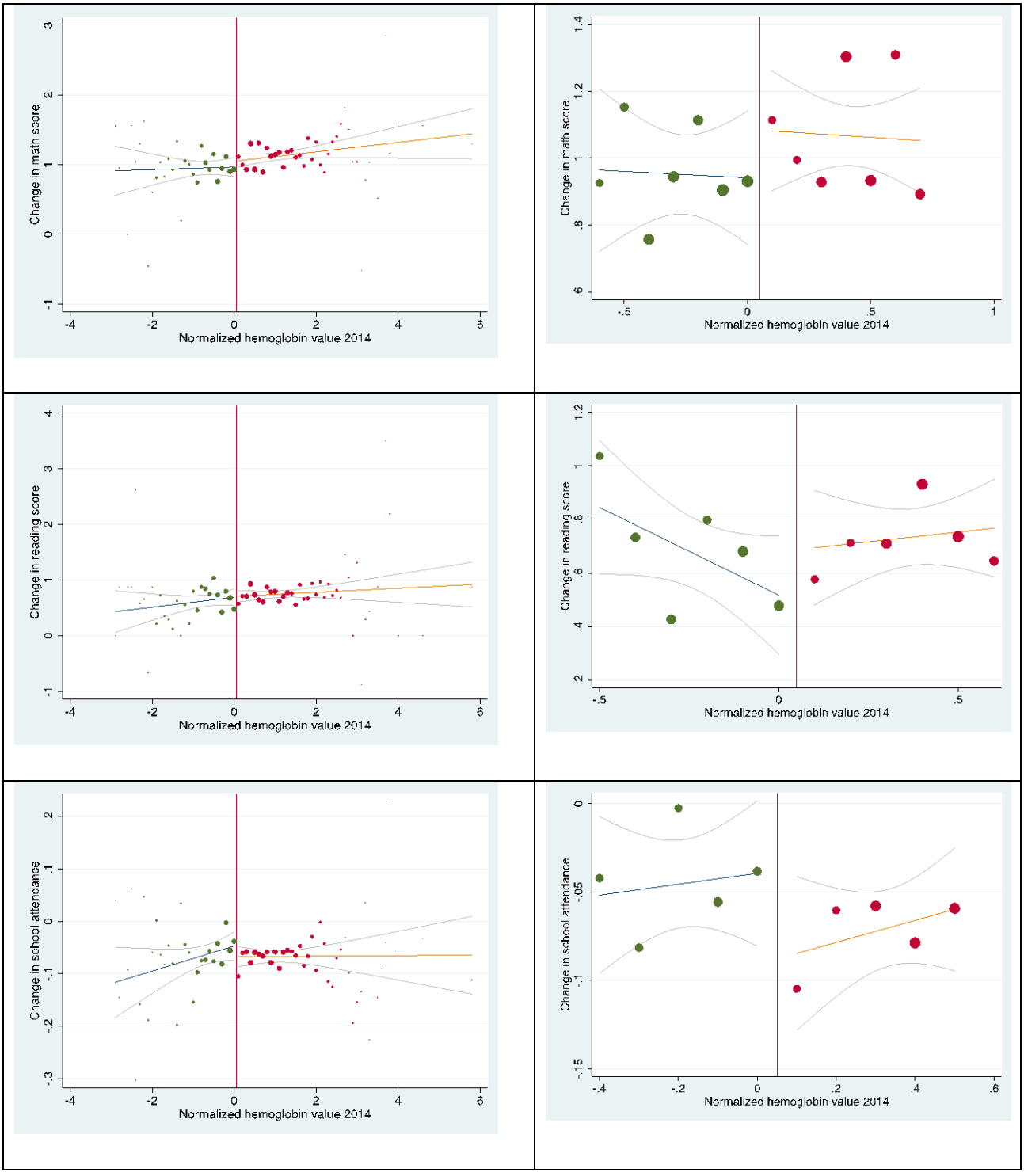
B. Anemia outcomes

Panel A: Globally	Panel B: Close to the cutoff
	
	

C. Cognitive and education outcomes

Panel A: Globally	Panel B: Close to the cutoff
	
	
	


Change in progressive matrices score
Normalized hemoglobin value 2014
Change in progressive matrices score
Normalized hemoglobin value 2014
Change in day and test test score
Normalized hemoglobin value 2014
Change in day and test test score
Normalized hemoglobin value 2014
Change in Cognition index
Normalized hemoglobin value 2014
Change in Cognition index
Normalized hemoglobin value 2014



4.4.2 Main Results

Estimation results for regression (1) are presented in table 4.2 for feeding practices and anemia and in table 4.3 for cognitive and education outcomes. In Panel A, estimation results are presented for the data driven bandwidth selected by the method proposed by CCT. Panel B shows estimates for different arbitrarily chosen bandwidths (0.3, 0.5, 1.0, 1.5, 2.0 and 2.5) and Panel C for estimates with a rectangular kernel. In panel D, we include a set of control

variables, and in panel E, we show results for the application of a polynomial of order two on observations with the CCT bandwidth. In Panel F, results for global estimates for different higher order polynomials are presented.[69]

The results from the discontinuity graphs can broadly be confirmed by the regression analysis. For the feeding practice outcomes (columns 1-3, table 4.2), none of the estimated coefficients are statistically significant and for the frequency of meat and green vegetable consumption, the coefficients display the unexpected sign. Regarding anemia outcomes (columns 4-5, table 4.2) there is a statistically significant effect of the nutrition information intervention on hemoglobin. Using CCT bandwidth, the estimate predicts that the information treatment on average provoked a negative change in hemoglobin scores by the size of 0.469 g/dl (P-value: 0.034, SE: 0.218). However, the effect is only statistically significant for very small bandwidths (0.3, 0.4 and 0.5) and does not stay robust across specifications.[70] The coefficient also displays the unexpected sign. Across all specifications, the coefficient for number of anemia symptoms displays the expected sign. They are, however, not statistically significant.

[69] Regression underlying panel G are described and discussed in the robustness checks.

[70] One would expect standard errors to get larger with smaller bandwidths, as estimates get more imprecise, and coefficients might change because of the bias inherent in a larger bandwidth. This pattern is, however, not observed in table 4.2.

Table 4.2: Average treatment effect for the population close to the cutoff of nutrition information on feeding practices and hemoglobin for different bandwidth and functional forms

	(1) Dietary diversity score	(2) Frequency of meat consumption	(3) Frequency of green veg consumption	(4) Hemoglobin	(5) N anemia symptoms
A CCT Bandwidth	0.229	-0.062	-0.052	-0.469**	0.418
Standard error	0.249	0.188	0.211	0.218	0.313
P-value	0.360	0.741	0.807	0.034	0.184
Bandwidth	0.7	0.7	0.8	0.4	0.5
N	733	733	818	517	475
B Alternative bandwidth					
Bandwidth 0.3	0.360	-0.109	0.012	-0.524**	0.840*
Standard error	0.400	0.308	0.390	0.244	0.461
P-value	0.372	0.725	0.976	0.035	0.072
N	300	300	300	280	255
Bandwidth 0.5	0.263	-0.106	0.051	-0.335*	0.418
Standard error	0.294	0.226	0.284	0.200	0.313
P-value	0.372	0.640	0.858	0.096	0.184
N	543	543	543	517	475
Bandwidth 1.0	0.126	-0.079	-0.098	-0.105	0.262
Standard error	0.197	0.150	0.190	0.145	0.189
P-value	0.525	0.600	0.606	0.473	0.169
N	1,022	1,022	1,022	969	887
Bandwidth 1.5	0.057	-0.088	-0.048	-0.068	0.137
Standard error	0.163	0.132	0.166	0.122	0.157
P-value	0.725	0.507	0.775	0.580	0.382
N	1,379	1,379	1,379	1,302	1,194
Bandwidth 2.0	-0.010	-0.077	0.016	-0.046	0.080
Standard error	0.145	0.120	0.150	0.109	0.144
P-value	0.946	0.520	0.914	0.671	0.582
N	1,606	1,606	1,606	1,509	1,390
Bandwidth 2.5	-0.044	-0.076	0.016	-0.015	0.099
Standard error	0.130	0.112	0.140	0.102	0.138
P-value	0.736	0.502	0.910	0.885	0.475
N	1,708	1,708	1,708	1,609	1,482
C Rectangular kernel	0.274	-0.037	-0.126	-0.193	0.335
Standard error	0.235	0.172	0.198	0.193	0.266
P-value	0.247	0.830	0.526	0.321	0.211
Bandwidth	0.7	0.7	0.8	0.4	0.5
N	733	733	818	517	475
D With controls	0.071	-0.079	-0.039	-0.434	0.413
Standard error	0.259	0.185	0.224	0.235	0.262
P-value	0.783	0.669	0.863	0.067	0.117
Bandwidth	0.7	0.7	0.8	0.4	0.5
N	686	686	762	485	447
E Local polynomial					
2nd order	0.152	-0.040	-0.073	-0.320	0.366
Standard error	0.236	0.211	0.265	0.205	0.301
P-value	0.520	0.850	0.783	0.121	0.227

Bandwidth	0.9	0.7	0.7	0.6	0.6
N	918	733	733	604	552
F Global polynomial regressions					
Polynomial 1st order	-0.170	-0.095	0.000	0.011	0.136
Standard error	0.109	0.100	0.126	0.093	0.142
P-value	0.123	0.345	0.998	0.906	0.340
N	1,758	1,758	1,758	1,656	1,526
Polynomial 2nd order	-0.033	-0.104	-0.014	0.004	0.163
Standard error	0.166	0.134	0.172	0.115	0.153
P-value	0.843	0.443	0.934	0.970	0.289
N	1,758	1,758	1,758	1,656	1,526
Polynomial 3rd order	0.123	-0.068	-0.031	-0.099	0.073
Standard error	0.205	0.160	0.219	0.143	0.199
P-value	0.549	0.674	0.889	0.491	0.714
N	1,758	1,758	1,758	1,656	1,526
G Donut (excluding Hb value 10.9 and 11)	-0.001	-0.014	-0.137	-0.047	-0.122
Standard error	0.251	0.202	0.295	0.241	0.272
P-value	0.996	0.946	0.643	0.844	0.655
Bandwidth	0.9	0.7	0.7	0.6	0.6
N	870	685	685	558	508

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in Panels A-E the RD coefficients are estimated by fitting a local linear regression using a triangular kernel to the right and the left of the cutoff. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 4.3: Average treatment effect for the population close to the cutoff of nutrition information on cognition and education for different bandwidth and functional forms

	(1) Block design	(2) Digit span forward	(3) Digit span backward	(4) Progressive matrices	(5) Day and night	(6) Cognitive index	(7) Math	(8) Reading	(9) Attendance
A CCT Bandwidth	-0.480**	0.137	-0.246	-0.149	-0.096	-0.310*	-0.104	-0.054	0.056
Standard error	0.232	0.234	0.187	0.185	0.183	0.183	0.163	0.195	0.036
P-value	0.041	0.559	0.190	0.425	0.599	0.093	0.527	0.784	0.129
Bandwidth	0.5	0.4	0.7	0.5	0.7	0.4	0.7	0.6	0.6
N	514	395	691	514	691	395	691	602	563
B Alternative bandwidth									
Bandwidth 0.3	-0.477	0.381	-0.361	-0.210	-0.087	-0.229	-0.294	-0.033	0.078*
Standard error	0.301	0.276	0.289	0.234	0.287	0.222	0.270	0.307	0.046
P-value	0.117	0.172	0.214	0.373	0.763	0.305	0.278	0.915	0.092
N	279	279	279	279	279	279	279	279	260
Bandwidth 0.5	-0.480**	0.015	-0.262	-0.149	-0.121	-0.300*	-0.107	0.017	0.058
Standard error	0.232	0.193	0.216	0.185	0.222	0.156	0.200	0.221	0.039
P-value	0.041	0.938	0.227	0.425	0.587	0.056	0.594	0.940	0.143
N	514	514	514	514	514	514	514	514	482
Bandwidth 1.0	-0.260	0.028	-0.212	-0.043	0.025	-0.140	-0.108	-0.120	0.044
Standard error	0.165	0.127	0.157	0.151	0.148	0.116	0.132	0.133	0.028
P-value	0.119	0.824	0.181	0.776	0.867	0.231	0.414	0.367	0.116
N	955	955	955	955	955	955	955	955	899
Bandwidth 1.5	-0.272**	0.095	-0.145	-0.018	0.043	-0.092	-0.106	-0.106	0.033
Standard error	0.133	0.113	0.132	0.135	0.135	0.103	0.112	0.109	0.024
P-value	0.043	0.401	0.274	0.896	0.753	0.377	0.343	0.332	0.182
N	1,290	1,290	1,290	1,290	1,290	1,290	1,289	1,290	1,216
Bandwidth 2.0	-0.294**	0.092	-0.124	-0.032	0.030	-0.101	-0.104	-0.069	0.024
Standard error	0.118	0.101	0.118	0.126	0.124	0.095	0.101	0.098	0.022
P-value	0.015	0.367	0.295	0.799	0.810	0.290	0.306	0.483	0.277
N	1,488	1,488	1,488	1,488	1,488	1,488	1,487	1,487	1,405
Bandwidth 2.5	-0.290***	0.064	-0.103	-0.029	0.026	-0.101	-0.096	-0.040	0.019

Standard error	0.109	0.091	0.111	0.118	0.113	0.088	0.094	0.094	0.020
P-value	0.009	0.482	0.359	0.808	0.816	0.252	0.312	0.670	0.340
N	1,584	1,584	1,584	1,584	1,584	1,584	1,583	1,583	1,493
C Rectangular kernel	-0.401*	-0.029	-0.244	-0.075	-0.048	-0.361**	-0.145	-0.163	0.053
Standard error	0.208	0.219	0.184	0.183	0.166	0.173	0.142	0.171	0.034
P-value	0.056	0.895	0.188	0.682	0.771	0.039	0.310	0.343	0.121
Bandwidth	0.5	0.4	0.7	0.5	0.7	0.4	0.7	0.6	0.6
N	514	395	691	514	691	395	691	602	563
D With controls	-0.436*	0.016	-0.264	-0.122	-0.122	-0.357**	-0.029	0.025	0.066
Standard error	0.241	0.228	0.212	0.194	0.200	0.173	0.168	0.198	0.034
P-value	0.073	0.943	0.216	0.532	0.544	0.042	0.862	0.899	0.054
Bandwidth	0.5	0.4	0.7	0.5	0.7	0.4	0.7	0.6	0.6
N	476	364	636	476	636	364	636	555	522
E Local polynomial									
2nd order	-0.509**	0.024	-0.170	-0.075	-0.059	-0.266*	-0.104	-0.115	0.049
Standard error	0.231	0.172	0.179	0.171	0.164	0.138	0.146	0.184	0.033
P-value	0.030	0.892	0.345	0.662	0.720	0.056	0.477	0.536	0.144
Bandwidth	0.6	0.7	0.9	0.8	1. 0	0.7	1. 0	0.7	0.8
N	602	691	864	775	955	691	955	691	727
F Global polynomial regressions									
Polynomial 1st order	-0.268***	-0.005	-0.083	-0.032	-0.030	-0.126	-0.081	-0.011	0.011
Standard error	0.096	0.074	0.100	0.105	0.096	0.080	0.086	0.093	0.017
P-value	0.006	0.952	0.410	0.762	0.755	0.116	0.347	0.902	0.514
N	1,627	1,627	1,627	1,627	1,627	1,627	1,626	1,626	1,535
Polynomial 2nd order	-0.322**	0.115	-0.052	-0.071	0.069	-0.080	-0.094	-0.053	0.021
Standard error	0.126	0.102	0.127	0.140	0.140	0.103	0.108	0.109	0.023
P-value	0.012	0.261	0.684	0.612	0.625	0.437	0.384	0.628	0.366
N	1,627	1,627	1,627	1,627	1,627	1,627	1,626	1,626	1,535
Polynomial 3rd	-0.321**	0.125	-0.182	0.013	0.045	-0.101	-0.144	-0.207	0.044

order									
Standard error	0.158	0.142	0.155	0.159	0.160	0.122	0.138	0.139	0.031
P-value	0.044	0.379	0.245	0.937	0.780	0.410	0.302	0.138	0.158
N	1,627	1,627	1,627	1,627	1,627	1,627	1,626	1,626	1,535
G Donut (excluding Hb values 10.9 and 11)	-0.193	-0.468**	-0.230	0.072	0.186	-0.234	-0.056	-0.153	0.027
Standard error	0.266	0.221	0.210	0.251	0.214	0.184	0.168	0.229	0.040
P-value	0.470	0.037	0.275	0.776	0.388	0.208	0.739	0.506	0.493
Bandwidth	0.6	0.7	0.9	0.8	1. 0	0.7	1. 0	0.7	0.8
N	558	647	820	731	911	647	911	647	688

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in Panels A-E the RD coefficients are estimated by fitting a local linear regression using a triangular kernel to the right and the left of the cutoff. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

For the cognitive measures, a few point estimates are statistically significant but most are not (columns 1-6, table 4.3). Based on the estimates using the CCT bandwidth (Panel A), the revealing of a child's anemia status and the provision of information on better feeding practices, on average, decreased the change in the *block design* test score by 0.480 standard deviations, compared to the mean in 2014 (P-value: 0.041, SE: 0.232). Since the cognitive index is a composite index of all five cognitive tests the statistically significant and qualitatively large point estimate for the *block design* tests is also reflected in a decrease of the cognitive index by 0.310 standard deviations (significant at the 10% level, P-value: 0.093, SE: 0.183). The coefficient for *block design* remains statistically significant for most specifications and the cognitive index is statistically significant only for some of the other specifications (bandwidth of 0.5, rectangular kernel, inclusion of control variables and local polynomial of 2nd order). However, coefficients show the unexpected, opposite sign. Since estimates for the different cognitive tests are not consistent, i.e. only one cognitive test shows robust statistically significant estimates (*block design*) and the direction of the coefficients for the different cognitive tests are also not uniform, we cannot draw a general conclusion regarding cognitive ability. If anything, we find a undesirable negative effect of the nutrition information intervention. Finding an effect on cognition but not on feeding practices and hemoglobin would also be counterintuitive, as cognitive outcomes could only be affected through a change in feeding practices and an increase in hemoglobin values. For the education outcomes (columns 7-9, table 4.3), none of the coefficients are statistically significant and in some specifications, they have the unexpected, opposite sign.

Overall, even though our treatment combines nutrition information with revealing the nutritional disorder of a child, we do not find a positive treatment effect that is robust across specifications and consistent across indicators. If anything there is weak evidence for counterintuitive and undesirable treatment effect for the hemoglobin and one cognition tests.

4.5 Interpretation

One potential reason for this finding might be that positive effects were only short-term and that they had diminished after two years.

Another explanation for this finding might be unavailability and accessibility of iron-rich food items. Standard economic theory would predict that in perfect markets and with fully informed individuals a pareto-efficient allocation of goods and services would be

reached (Mankiw and Taylor, 2011), i.e. the information intervention would enable parents to make the rational decision to invest in child nutrition and thereby promote their child's health, cognitive ability and productivity. Markets are, however, rarely perfect. Especially in rural areas of developing countries where financial and commodity markets are often missing entirely, are imperfect or transaction costs to access markets are high (e.g. Ray 2007). Hence, even though parents in our sample might have wanted to adopt their nutrition behavior in reaction to the nutrition information, missing markets might have prevented them from doing so. In the introduction, we provided motivation for the nutrition information intervention by pointing out that the availability of commodities or technologies is not the only requirement for their adoption (supply side interventions), but that also the demand for these technologies must be created by nutrition information. The opposite is obviously equally true: For the information intervention to materialize in a change in nutrition behavior, the availability and the (financial) accessibility of the technology (in our case, the availability of iron-rich food products) are also required. The WHO (2001) states that improvements in nutrition build on these three pillars: the availability of micronutrient-rich food, the financial accessibility to those food items as well as a change in feeding practices (i.e. utilization). However, the tested intervention only addresses the latter pillar. In the follow up survey in 2016, we included two questions that give an idea about the availability and financial accessibility of iron-rich food items. Answers from these questions are presented in table 4.7. Table 4.7 shows that both availability and accessibility might have constrained parents from providing a more diverse diet for their children, even though they would have been willing to do so. In the follow-up survey, nearly 86% of parents that were given the nutrition information reported that they were often or sometimes not able to afford feeding their child a balanced meal and 77% reported that they were often or sometimes not able to feed their child a balanced meal, because only a limited variety of food was available in their surroundings.

Table 4.7: Availability and Accessibility of a diverse diet for treated households

	(1) Often	(2) Sometimes	(3) Never
Accessibility			
In the last 12 months, how often were you NOT able to afford to feed your child balanced meal?	48%	39%	12%
Availability			
In the last 12 months, how often were you NOT able feed your child a balanced meal because only a limited variety of food was available in your surrounding?	35%	43%	21%

We would like to point out that these numbers are from a subjective measure and hence reliability might be limited. However, we interpret table 4.7 as suggestive evidence that unavailability and inaccessibility, i.e. that imperfect markets have indeed hampered a potentially intended change in feeding practices. The unavailability and inaccessibility of iron-rich food items might also explain the potential presents of a counterintuitive and undesirable treatment effect. Maybe parents were not able to purchase the products the medical personnel recommended to them, but rather changed their feeding practices in an undesirable way.

Table 4.7 provides an idea of a possible reason why information alone does not change child nutrition and cognitive ability, and supports the interconnection between information, availability and accessibility for nutrition and health products. Since table 4.7 can only provide suggestive evidence, we do not want to argue that this is the full story. Empirical evidence supports the notion that even when availability and accessibility is ensured, people do not necessarily respond to health information or at least not as much as one would expect. E.g. children who were educated about the adverse effects of worm infections and means to prevent infection did not adapt their health behavior (i.e. wear shoes and adopt more hygienic behavior), even though adoption should have been cheap and feasible (Kremer and Miguel, 2007). Fernandes et al. (2014) conduct a meta-analysis on financial education and find that financial information explains very little of the savings and borrowing behavior, despite the existence of financial services. In the context of nutrition, Banerjee et al. (2015) found that making iron-fortified iodized salt (DFS) available in Indian villages and informing households with a flyer of the product's availability and its benefits, did not encourage take-up. In contrast, an education movie had a positive impact on adoption of the fortified salt compared to the control group; however, the overall take-up was only about 10%. Assuming that constraints in availability and accessibility are not or only partly the problem, it might be that individuals are making a fully rational decision by not adopting feeding practices in the way it was promoted by the medical staff in our intervention, because the medical staff is not their only source of information. Childs et al. (1997) argue that parents are confronted with many different views, information and ideas on how a child should be fed. They get advice from their extended family, their community, advertisements and the media. The nutrition information given in the scope of this study must therefore be seen as a complex set of information that parents encounter in their decision making process.

Childs et al. (1997), as well as Banerjee & Duflo (2012), also argue that beliefs, social norms, habits and culture play a crucial role when people make decision about food intake. In case new information contrasts with deep-seated beliefs and habits they will hardly be taken-up. Emotions, beliefs, habits and culture as underlying drivers of decisions, only entered economics recently when sociology and psychology was applied to economic problems. Insights from psychology, also point to certain mental tendencies that prevent human beings from converting their intentions into actions, including limited attention[71] and present bias (Laibson 1997, O'Donoghue & Rabin 1999, Banerjee & Mullainathan 2010, Dupas 2011, Karlan et al. 2006). The present bias, i.e. the fact that individuals give stronger weight to rewards and discomforts that are in the present, might also play a crucial role in explaining why revealing that a child is moderately anemic is less effective compared to revealing the presence of other diseases (HIV, malaria). While the degree of suffering from a nutritional disorder might not be perceived as very severe, mostly because the cause and symptoms are more subtle and salient, the costs of changing nutritional habits are high. Moreover, the benefits of a nutritional change need time to materialize and lie very much in the future and are hence, according to the present bias theory, strongly discounted. The weak evidence for an counterintuitive and undesirable treatment effect might also be explained by mental tendencies. It might have been the case that learning that one's child is anemic discouraged parents to invest more in this child and rather invested in their sibling. Unfortunately we cannot test for this hypothesis.

4.6 Robustness checks

4.6.1 Power

As RDD estimates are generally considered more reliable within a small bandwidth around the cutoff point, a lot of observations are normally excluded from the analysis. Furthermore, the correlation between the RDD forcing variable and the treatment status reduces the power of the estimates. This is why large sample sizes are generally required for estimating treatment effect using RDDs (Schochet 2009).

[71] Limited attention refers to the idea that given the wealth of information a human being is exposed to, one can hardly consider all of the dimensions and options (Mullainathan & Sharif, 2013).

We calculate the minimal detectable effect (MDE) for different bandwidths taking the correlation between the treatment and the forcing variable into account (Table 4.4).[72] In the last column of table 4.4, we compare these MDE to those found in Luo et al. (2012) and the effect sizes found in the intervention evaluated in essay 2 (chapter 3). Recall that Luo et al. (2012) is the study that is closest to our evaluation and that in essay 2, we evaluated another nutrition intervention where we study the effect of using iron-fortified iodized salt in the Indian school feeding program with the same dataset. Hence those estimates can be used as benchmarks of reasonable effect sizes for similar nutrition interventions. At least for some of the outcomes, which we consider to be comparable (hemoglobin, math and reading), the MDE that can be found in this study are 2 to 3 times larger. With a bandwidth of 2.5 in this study, we would be able to detect a change in hemoglobin levels of 0.551 g/dl. For comparison, in the intervention from essay 2, we found an effect size of 0.138 g/dl (significant at the 10% level) and Luo et al. (2012) found an effect size of 0.275 g/dl (not statistically significant) in one of their information interventions. In contrast, in a multivitamin supplementation intervention that was used by Luo et al. (2012) as a benchmark, effect sizes ranged from 0.202 to 0.416 g/dl. Estimates for math and reading from the intervention of essay 2, were in the range of 0.2 standard deviations (for 90% school attendance) and estimates that we would be able to detect in this study would be in the range of 0.6 standard deviations. These numbers indicate that insufficient power might indeed be a reason for the non-detection of an effect of the intervention on some of the outcomes. Or framed differently we are only able to detect relatively large effect sizes with the given dataset, which is a drawback of this study.

[72] The MDE is calculated using the following formula:
$MDE = (t_{(1-\kappa)}+t_{\alpha})/(1/P(1-P)J)^{1/2} * (\rho + (1- \rho)/n\ \sigma)^{1/2} * RDDE$
where $RDDE = 1/(1-r^2)$ and r is the correlation between treatment status and the continuous assignment variable, i.e. RDDE is the RDD design effect. For a normal distribution and a position of the cutoff at 25% of the distribution – the conditions that apply to the data set used in this analysis - Schochet (2009) calculate a RDDE of 2.17 for a linear functional form. The remaining part of the formula is the standard formula for calculating MDE in RCTs. We hence multiply the MDE that could be detected in a RCT setting by the factor 2.17. We assume a power of $\kappa = 80\%$ and a significance level of $\alpha = 5\%$. Standard deviation σ, number of clusters J, fraction of treated individuals P, n the average number of individual in each cluster and ρ the intra-cluster correlation is taken from the dataset itself for observations within the respective bandwidth and for the respective outcome. We do not know the take-up of the nutrition information and we assume 100% take-up by parents of children with a hemoglobin value $\leq$ 10.9 g/dl and 0% take-up of the nutrition information by parents of children with a hemoglobin value > 10.9 g/dl.

Table 4.4: Minimal detectable effects for different bandwidth

	Bandwidth						Effect size in other nutrition interventions
	0.3	**0.5**	**1.0**	**1.5**	**2.0**	**2.5**	
Dietary diversity	1.776	1.361	0.97	0.906	0.876	0.843	
Frequency of meat consumption	1.409	0.991	0.667	0.613	0.605	0.579	
Frequency of green vegetable consumption	2.081	1.66	1.318	1.186	1.109	1.051	
Hemoglobin	0.871	0.724	0.601	0.545	0.563	0.551	0.136*[1] 0.151[2] 0.275[3] 0.202**[4] 0.416**[4]
Number anemia symptoms	1.733	1.328	1.073	0.974	0.951	0.916	-0.137[1] -0.157[2]
Block design	1.202	0.922	0.698	0.699	0.667	0.639	0.012[1] 0.045[2]
Digit span forwards	1.246	0.99	0.655	0.599	0.585	0.555	-0.105[1] -0.135[2]
Digit span backwards	1.278	0.996	0.788	0.715	0.691	0.668	0.009[1] -0.23[2]
Progressive matrices	1.123	1.211	0.849	0.76	0.713	0.692	0.070[1] 0.112[2]
Day and night	1.32	1.134	0.862	0.754	0.723	0.683	0.116[1] 0.210[2]
Cognitive index	0.919	0.91	0.69	0.634	0.601	0.58	0.028[1] 0.058[2]
Math	1.273	1.054	0.79	0.729	0.701	0.663	0.112 0.197*[2]
Reading	1.253	1.039	0.711	0.664	0.635	0.603	0.129[1] 0.182*[2]
Attendance	0.265	0.19	0.156	0.147	0.143	0.136	-0.005[1]

1 Effect size from the evaluation of the intervention from essay 2 for the Difference-in-Difference estimates.
2 Effect size from the evaluation of the intervention from essay 2 for the Difference-in-Difference estimates at 90% school attendance.
3 Effect size in Luo et al. (2012), Information experiment 2.
4 Effect size in Luo et al (2012), Experiments 1 and 2, multivitamin supplement treatment arm.
*, **, *** denote significance at the 10%, 5% and 1% level, respectively.

To increase power, we run regression (1) without taking differences between the values in 2014 and 2016 for the outcome variables, i.e. instead of ΔY we include Y, the value of the outcome variable in 2016. The smallest MDE that can be found without taking differences in outcomes are presented in table 4.A.3 in the appendix. Without taking differences, the MDE only get slightly smaller. Estimates without taking differences in

outcomes are very similar to the estimates from the previous specification (regression output not shown).

We cannot rule out the possibility that we lack power to detect an effect that is smaller than numbers reported in table 4.4 and in table 4.A.3 in the appendix. The detection of even smaller effect sizes would, however, become less relevant from an economic perspective.

4.6.2 Irregularities in the Conveyance of the Nutrition Information

There are two sources of potential irregularities in the conveyance of the nutrition information, which might provoke an attenuation bias, i.e. bias estimates towards zero.[73] The first potential source is that the information did not reach the parents or did not reach the person that is responsible for child feeding. We consider this risk to be rather low since the hemoglobin testing drew a lot of attention in the village and most of the times many people gathered together during the testing. Thus, if the information was not taken up directly by the parents it was very likely taken up by a neighbor, a sibling or grandparent and possibly shared with the parents later. However, we assume that there might be stronger treatment effects for the conveyance of the nutrition information directly to the mother, i.e. when it was ensured that the information reached the mother. We came up with this assumption because some empirical evidence shows that decisions on child health and nutrition are mostly made by mothers and grandmother (Thomas 2011, Thomas 1993). From another household survey that one of the authors conducted in another district of Bihar, we also know that mothers mostly make decisions about what to cook and what children eat in our study area.[74] Our dataset allows us to distinguish between observations where mothers were present during the hemoglobin testing from those where mothers were absent, because whenever mothers were present maternal anthropometrics were taken. We therefore run regression (1) for the subgroup of children for which we collected maternal anthropometrics. Results of this regression are shown in table 4.5 for the feeding practice and anemia outcomes and in table 4.6 for the cognitive and education outcomes. The picture looks very much the same as in the regressions before. There is no statistically significant effect for any of the specifications on

[73] Before we were essentially assuming a situation of perfect compliance: Everybody who was earmarked to receive the nutrition information also received it, and every parent to whom the nutrition information was not earmarked, did not receive it. Now we are discussing the implications of this assumption and we relax the assumption to some extent.

[74] In the other household survey, it was found that in 68% of the households, females between 18-49 years of age decided what is cooked in the household (which would be mostly the mothers in our survey), and in 22% of households, grandmothers decide.

feeding practices and the number of anemia symptoms. For the smaller bandwidths, hemoglobin levels appear to be statistically significant; however, the estimates are not robust at the larger bandwidths and different functional forms. Furthermore, coefficients have a negative sign. For the cognitive outcomes, some coefficients are statistically significant; however, except for the *block design* estimates, they do not seem to be robust across the different specifications and the signs of the coefficients are also not consistent across cognitive tests. Estimates for the education outcomes are insignificant across the different specifications and mostly have the expected sign. Taking this evidence together, same conclusion holds for the subgroup of children where mothers were present during the hemoglobin testing: We find no robust impact of revealing a child's anemia status and providing nutrition information to parents of anemic children.

Table 4.5: Average treatment effect for the population close to the cutoff of nutrition information on feeding practices and hemoglobin for the subgroup of hemoglobin tests where mothers were present for different bandwidth and functional forms

	(1) Dietary diversity score	(2) Frequency of meat consumption	(3) Frequency of green veg consumption	(4) Hemoglobin	(5) N anemia symptoms
A CCT Bandwidth	0.279	-0.072	0.010	-0.388*	0.365
Standard error	0.292	0.206	0.275	0.225	0.394
P-value	0.342	0.726	0.971	0.088	0.357
Bandwidth	0.6	0.7	0.5	0.4	0.4
N	560	647	484	350	322
B Alternative bandwidth					
Bandwidth 0.3	0.510	-0.109	-0.029	-0.439*	0.697
Standard error	0.439	0.330	0.379	0.252	0.478
P-value	0.249	0.743	0.940	0.084	0.148
N	270	270	270	249	229
Bandwidth 0.5	0.355	-0.098	0.010	-0.287	0.285
Standard error	0.319	0.247	0.275	0.206	0.330
P-value	0.269	0.691	0.971	0.167	0.390
N	484	484	484	457	422
Bandwidth 1.0	0.175	-0.114	-0.158	-0.093	0.189
Standard error	0.216	0.168	0.195	0.147	0.198
P-value	0.420	0.500	0.420	0.528	0.344
N	907	907	907	856	784
Bandwidth 1.5	0.084	-0.125	-0.102	-0.062	0.069
Standard error	0.180	0.144	0.175	0.123	0.166
P-value	0.644	0.386	0.562	0.616	0.677
N	1,226	1,226	1,226	1,159	1,065
Bandwidth 2.0	0.009	-0.104	-0.045	-0.049	0.002
Standard error	0.161	0.129	0.160	0.111	0.153
P-value	0.954	0.423	0.779	0.660	0.988
N	1,428	1,428	1,428	1,347	1,243
Bandwidth 2.5	-0.022	-0.093	-0.033	-0.016	0.019
Standard error	0.145	0.120	0.151	0.104	0.147
P-value	0.880	0.439	0.825	0.881	0.898
N	1,522	1,522	1,522	1,438	1,327
C Rectangular kernel	0.154	-0.072	-0.049	-0.359	0.133
Standard error	0.267	0.187	0.253	0.222	0.365
P-value	0.565	0.702	0.847	0.109	0.716
Bandwidth	0.6	0.7	0.5	0.4	0.4
N	560	647	484	350	322
D With controls	0.131	-0.038	-0.043	-0.346	0.392
Standard error	0.309	0.210	0.274	0.245	0.334
P-value	0.674	0.857	0.875	0.161	0.244
Bandwidth	0.6	0.7	0.5	0.4	0.4
N	525	607	452	328	302
E Local polynomial					
2nd order	0.295	-0.037	-0.094	-0.314	0.267
Standard error	0.274	0.235	0.277	0.226	0.328
P-value	0.284	0.874	0.736	0.167	0.418

Bandwidth	0.8	0.7	0.6	0.5	0.6
N	813	647	560	457	487
F Global polynomial regressions					
Polynomial 1st order	-0.154	-0.107	-0.015	0.019	0.063
Standard error	0.122	0.105	0.138	0.095	0.146
P-value	0.209	0.310	0.912	0.846	0.666
N	1,570	1,570	1,570	1,482	1,368
Polynomial 2nd order	-0.015	-0.132	-0.084	0.004	0.059
Standard error	0.181	0.144	0.182	0.117	0.165
P-value	0.932	0.359	0.646	0.971	0.723
N	1,570	1,570	1,570	1,482	1,368
Polynomial 3rd order	0.160	-0.104	-0.100	-0.103	-0.019
Standard error	0.223	0.174	0.227	0.146	0.207
P-value	0.475	0.552	0.659	0.480	0.928
N	1,570	1,570	1,570	1,482	1,368
G Donut (excluding Hb values 10.9 and 11)	0.026	-0.252	0.011	-0.006	-0.142
Standard error	0.343	0.254	0.409	0.343	0.435
P-value	0.940	0.324	0.979	0.986	0.745
Bandwidth	0.8	0.7	0.6	0.5	0.6
N	727	561	474	375	408

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in Panels A-E, the RD coefficients are estimated by fitting a local linear regression using a triangular kernel to the right and left of the cutoff. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 4.6: Average treatment effect for the population close to the cutoff of nutrition information on cognition and education for the subgroup of hemoglobin tests where mothers were present for different bandwidth and functional forms

	(1) Block design	(2) Digit span forward	(3) Digit span backward	(4) Progressive matrices	(5) Day and night	(6) Cognitive index	(7) Math	(8) Reading	(9) Attendance
A CCT Bandwidth	-0.440*	0.067	-0.266	-0.147	-0.115	-0.318*	-0.159	-0.045	0.056
Standard error	0.243	0.241	0.225	0.196	0.206	0.188	0.146	0.222	0.039
P-value	0.073	0.780	0.240	0.454	0.576	0.093	0.279	0.839	0.154
Bandwidth	0.5	0.4	0.6	0.5	0.6	0.4	0.8	0.5	0.6
N	460	354	535	460	535	354	688	460	504
B Alternative bandwidth									
Bandwidth 0.3	-0.391	0.347	-0.298	-0.128	-0.087	-0.172	-0.262	-0.044	0.076
Standard error	0.309	0.280	0.313	0.248	0.278	0.222	0.269	0.304	0.048
P-value	0.210	0.218	0.344	0.606	0.753	0.441	0.334	0.886	0.121
N	251	251	251	251	251	251	251	251	236
Bandwidth 0.5	-0.440*	-0.052	-0.267	-0.147	-0.141	-0.314*	-0.153	-0.045	0.058
Standard error	0.243	0.203	0.239	0.196	0.221	0.164	0.193	0.222	0.042
P-value	0.073	0.799	0.267	0.454	0.525	0.058	0.429	0.839	0.165
N	460	460	460	460	460	460	460	460	435
Bandwidth 1.0	-0.256	-0.004	-0.246	-0.052	0.047	-0.154	-0.152	-0.177	0.043
Standard error	0.171	0.141	0.175	0.165	0.155	0.127	0.126	0.133	0.029
P-value	0.137	0.979	0.162	0.751	0.762	0.229	0.229	0.184	0.137
N	852	852	852	852	852	852	852	852	805
Bandwidth 1.5	-0.255*	0.085	-0.173	-0.035	0.064	-0.097	-0.145	-0.161	0.034
Standard error	0.141	0.124	0.145	0.145	0.143	0.114	0.108	0.108	0.025
P-value	0.073	0.496	0.237	0.808	0.653	0.398	0.180	0.138	0.170
N	1,152	1,152	1,152	1,152	1,152	1,152	1,151	1,152	1,090
Bandwidth 2.0	-0.265**	0.077	-0.142	-0.041	0.046	-0.100	-0.147	-0.115	0.027
Standard error	0.128	0.110	0.130	0.134	0.128	0.103	0.097	0.098	0.022
P-value	0.041	0.485	0.275	0.764	0.721	0.336	0.132	0.241	0.233
N	1,330	1,330	1,330	1,330	1,330	1,330	1,329	1,329	1,263

Bandwidth 2.5	-0.260**	0.050	-0.116	-0.034	0.033	-0.100	-0.139	-0.078	0.021
Standard error	0.118	0.098	0.123	0.124	0.115	0.095	0.088	0.095	0.021
P-value	0.030	0.609	0.350	0.784	0.775	0.293	0.115	0.411	0.311
N	1,420	1,420	1,420	1,420	1,420	1,420	1,419	1,419	1,347
C Rectangular kernel	-0.355	-0.133	-0.272	-0.099	-0.073	-0.419	-0.150	-0.106	0.053
Standard error	0.217	0.226	0.220	0.196	0.201	0.183	0.136	0.209	0.037
P-value	0.105	0.559	0.218	0.617	0.718	0.024	0.272	0.615	0.149
Bandwidth	0.5	0.4	0.6	0.5	0.6	0.4	0.8	0.5	0.6
N	460	354	535	460	535	354	688	460	504
D With controls	-0.444	-0.104	-0.292	-0.183	-0.164	-0.430	-0.065	-0.028	0.076*
Standard error	0.260	0.243	0.240	0.187	0.220	0.179	0.151	0.213	0.037
P-value	0.091	0.669	0.227	0.330	0.459	0.018	0.668	0.894	0.040
Bandwidth	0.5	0.4	0.6	0.5	0.6	0.4	0.8	0.5	0.6
N	429	329	499	429	499	329	637	429	470
E Local polynomial									
2nd order	-0.451*	-0. 9	-0.180	-0.035	-0.125	-0.264	-0.146	-0.094	0.043
Standard error	0.246	0.178	0.220	0.183	0.195	0.159	0.154	0.205	0.035
P-value	0.069	0.958	0.417	0.850	0.525	0.100	0.343	0.648	0.212
Bandwidth	0.6	0.7	0.8	0.8	0.8	0.6	0.9	0.6	0.9
N	535	613	688	688	688	535	768	535	727
F Global polynomial regressions									
Polynomial 1st order	-0.236**	-0.009	-0.094	-0.030	-0.032	-0.122	-0.123	-0.049	0.010
Standard error	0.103	0.078	0.113	0.106	0.099	0.084	0.079	0.094	0.019
P-value	0.023	0.911	0.406	0.776	0.746	0.152	0.122	0.606	0.580
N	1,460	1,460	1,460	1,460	1,460	1,460	1,459	1,459	1,386
Polynomial 2nd order	-0.282**	0.090	-0.064	-0.075	0.089	-0.074	-0.134	-0.097	0.025
Standard error	0.135	0.109	0.142	0.147	0.144	0.110	0.101	0.109	0.024
P-value	0.038	0.415	0.652	0.612	0.537	0.502	0.191	0.377	0.308
N	1,460	1,460	1,460	1,460	1,460	1,460	1,459	1,459	1,386

Polynomial 3rd order	-0.308*	0.100	-0.206	-0.005	0.060	-0.112	-0.189	-0.264	0.046
Standard error	0.168	0.151	0.167	0.174	0.171	0.133	0.139	0.141	0.031
P-value	0.070	0.507	0.222	0.975	0.727	0.403	0.176	0.064	0.143
N	1,460	1,460	1,460	1,460	1,460	1,460	1,459	1,459	1,386
G Donut (excluding Hb values 10.9 and 11)	-0.247	-0.667**	-0.328	-0.041	0.148	-0.467	-0.129	-0.175	0.027
Standard error	0.282	0.260	0.243	0.276	0.258	0.234	0.179	0.285	0.040
P-value	0.383	0.012	0.180	0.884	0.567	0.048	0.471	0.541	0.500
Bandwidth	0.6	0.7	0.8	0.8	0.8	0.6	0.9	0.6	0.9
N	454	532	607	607	607	454	687	454	652

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in Panels A-E, the RD coefficients are estimated by fitting a local linear regression using a triangular kernel to the right and the left of the cutoff. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Finally, as previously stated, we cannot rule out the possibility that our medical staff sometimes forgot to inform parents or that the parents of the control children took-up the treatment and subsequently changed feeding practices, when the nutrition information was addressed to parents of anemic children (in the evaluation framework those parents would be called crossovers). Even when we assume that the irregularities in the conveyance of the treatment were not systematic but occurred randomly, this circumstance might cause an attenuation bias and might hence be the reason why we do not find any significant robust treatment effect. A similar concern, with the same implication, might be that the information was not conveyed very precisely, particularly at the cutoff, i.e. even though being instructed differently it might have been the case that the medical personnel did not take the cutoff too seriously and sometimes advised parents of children with a hemoglobin of 11 g/dl to change their feeding practice or did not advice parents of children with a hemoglobin of 10.9 g/dl. To investigate this concern, we perform a regression where we exclude observations directly at the cutoff (10.9 and 11 g/dl). Results of the *donut* regression are shown in the last panel (panel G) of table 4.2 and 4.3 for the main results and in tables 4.5 and 4.6 for the subgroup of hemoglobin tests where the nutrition information was directly conveyed to the mother. Results are in line with the previous estimates such that we do not think that this type of imprecision was an issue. Since it might be the case that the quality and carefulness of the medical staff differed, we test if estimates change when excluding each of the six medical staff once. Again, the results do not change much. There are no statistically significant effects on feeding practices and anemia and the few estimates for cognitive outcomes that show up to be statistically significant, do not show a uniform picture (regressions not shown).

4.6.3 Attenuation Bias

As the data used for this evaluation comes from the intervention described in essay 2, which tested the effects of using iron-fortified iodized salt in the Indian school-feeding program, there might be another source of attenuation bias. In general, the circumstance that another nutrition intervention was evaluated with the same dataset does not bias the results since the intervention evaluated in essay 2 was randomized at the school level and hence children that were treated by the iron-fortified salt intervention were equally distributed across the different hemoglobin values found in 2014. Table 4.1 shows the share of children that belonged to the treatment group from essay 2, just above and below the threshold (first variable in the covariates section). To the left of the cutoff, 52% of the children were treated by the

intervention described in essay 2 and to the right of the cutoff, this is true for 55% of the children, showing that belonging to the treatment group from essay 2 is quite balanced above and below the cutoff. Nevertheless, we included the information if the child was treated by the intervention evaluated in essay 2 in the set of control variables for one of our specifications. The inclusion of covariates did not make much of a difference in estimation results (tables 4.2, 4.3, 4.5, 4.6 panel D). Since the intervention evaluated in essay 2 looks at similar outcomes (hemoglobin, cognition and education) and at least the hemoglobin levels and also the math and reading scores were weakly positively affected by the intervention from essay 2, there might be the risk that a saturation effect had occurred. The saturation might again bias our estimates towards zero. It might have been the case that the intervention evaluated in essay 2 had already increased the hemoglobin values and other outcomes in such a way that a change in feeding practices could not add anything or very little. In case our results are indeed biased toward zero, we should still find a statistically significant change in feeding practices, which we do not. As a robustness check to the potential presence of a saturation effect, we run regression (1) exclusively for the control group of the intervention evaluated in essay 2 (due to the above stated problem of limited power we had used the whole sample form RCT – essay 2 – in the analysis before). Doing this, we lose about half of our observations and hence a lot of power, which is the reason why we did not restrict our main analysis to the control group of the intervention from essay 2. Results for this subsample are shown in table 4.A.1 in the appendix for the feeding practice and anemia outcomes and in table 4.A.2 in the appendix for the cognition and education outcomes. The pattern of the estimates remains similar to the full sample; however, the coefficients for hemoglobin are not statistically significant for any of the specifications, whereas the coefficients for the *backward digit-span* and the *day and night tests* turn out to be statistically significant. Again, the estimates for none of the outcomes are robust across specifications and they often have the unexpected sign. If these estimates are not underpowered, they indicate that an attenuation bias due to saturation is not leading to an underestimation of the true treatment effect, but rather that there is no effect from the information intervention.

4.7 Conclusion

In this article, we report results from a RDD analysis, studying the impact of revealing the anemia status of school-aged children and advising parents to feed their child iron-rich food items in rural India. As outcome variables, we looked at feeding practices, hemoglobin

levels, cognitive ability and education outcomes. At least for the medium term (two years after the intervention), we did not find any robust impact on any of the measured outcomes. Our results rather point to a counterintuitive and undesirable treatment effect for some of the outcome variables (hemoglobin and one cognition test). The non-detection of any treatment effect for many of the outcomes is in line with the findings from Luo et al. (2012), who also did not find any effect of a very similar intervention in China. It is, however, in contrast to other health information treatments that also included the revealing of individual health status and where individuals – at least partly – reacted to the treatment (Thornton 2008, Cohen et al. 2015). A draw back of the study is that we would not be able to detect them if the effect sizes are not sufficiently large because of a lack of power.

Given that too little power is not a problem in our analysis, potential explanations for the non-detection of a treatment effect might be that positive effects were only short-term and that they had diminished after two years. It might further be the case that unavailability and inaccessibility of iron-rich food items prevented parents from improving feeding practices or led to an undesirable change in feedings practices. We also point to certain mental tendencies that prevent parents from putting their intentions into actions, including present bias, limited attention and a discouragement effect.

Unfortunately, we are not able to disentangle the true underlying mechanism why information, even when combined with revealing the existence of a nutritional disorder of a child, did not show any significant treatment effect. However, given what we already know about what limits and what motivates human behavior from a neoclassical perspective, as well as from insights from sociology and psychology, it becomes clear that any policy or intervention that aims at changing nutritional behavior needs to consider a complex array of different factors[75]. Certainly

> *"…we should recognize – indeed assume – that information alone will not do the trick. This is just how things are, for the poor, as for us…"* (Banerjee & Duflo 2012, p. 70).

This seems to be true even when the presence of a nutritional disorder is revealed. Understanding the reasons why some health and nutrition information work and others do not and under which circumstances they work is a core interest in making information an effective tool for development.

[75] Some innovative interventions have been developed and successfully tested in the field. They range from commitment opportunities, nudging (i.e. small incentives), reminders to enforced mandatory policies (Dupas, 2011b).

4.A Appendix

Text 4.A.1 Control Variables

We use a set of control variables for the three different outcome categories. For the feeding practice and anemia outcomes we include socioeconomic characteristics (rural or urban, block, a wealth index, the father's and mother's years of schooling, caste, religion, the number of household members and if the child's father lives in the household), nutritional factors (an indicator for household food security, the number of meals the child eats every day, if the child took iron supplements in the last year, the average intake of calories and iron from the school meal, an indicator for maternal health knowledge), access to healthcare (dummy for institutional delivery of the child and if any household member is covered by a health insurance), morbidity indicators (if the child suffered from diarrhea in the last 30 days and if the household possesses an improved sanitation facility) as well as one biological factor (sex of the child). For the cognitive outcomes, we include the same covariates. In addition, we use indicators for psychosocial stimuli (a dummy if the mother helps the child with its homework, the time the mother spends on giving physical care to the child and if parents participate in parent-teacher meetings at school, a dummy if the father lives in the household) and a dummy for the test administrator was included. Further indicators for quality of schooling (total school enrollment, the student-teacher ratio, the number of children that attended second grade at the baseline and the fourth grade at the endline on the day of the interview and the distance to the school) were included as controls for the education outcomes. Since it might be the case that the quality of information conveyed differs between the six medical people that conducted the medical tests, we further control for who conducted the blood test for anemia.

Table 4.A.1: Average treatment effect for the population close to the cutoff of nutrition information on feeding practices and hemoglobin for different bandwidth and functional excluding treatment group of the intervention from essay 2

	(1) Dietary diversity score	(2) Frequency of meat consumption	(3) Frequency of green veg consumption	(4) Hemoglobin	(5) N anemia symptoms
A CCT Bandwidth	0.444	-0.127	-0.040	-0.298	-0.286
Standard error	0.562	0.546	0.841	0.258	0.551
P-value	0.434	0.818	0.963	0.253	0.607
Bandwidth	0.4	0.4	0.3	0.6	0.4
N	172	172	122	266	155
B Alternative bandwidth					
Bandwidth 0.3	0.425	-0.236	-0.040	-0.670	-0.189
Standard error	0.645	0.619	0.841	0.411	0.686
P-value	0.514	0.705	0.963	0.110	0.785

N	122	122	122	122	106
Bandwidth 0.5	0.413	-0.098	-0.038	-0.340	-0.218
Standard error	0.520	0.469	0.667	0.285	0.469
P-value	0.431	0.835	0.955	0.238	0.644
N	225	225	225	227	204
Bandwidth 1.0	0.176	-0.034	-0.118	-0.190	-0.054
Standard error	0.349	0.285	0.443	0.210	0.273
P-value	0.616	0.904	0.791	0.371	0.845
N	421	421	421	420	379
Bandwidth 1.5	0.117	-0.067	-0.022	-0.154	0.010
Standard error	0.285	0.231	0.346	0.190	0.224
P-value	0.684	0.772	0.948	0.420	0.965
N	583	583	583	575	522
Bandwidth 2.0	0.103	-0.075	0.097	-0.109	0.017
Standard error	0.256	0.206	0.292	0.168	0.209
P-value	0.690	0.716	0.742	0.520	0.935
N	692	692	692	682	623
Bandwidth 2.5	0.090	-0.087	0.093	-0.072	0.069
Standard error	0.226	0.190	0.257	0.154	0.200
P-value	0.694	0.648	0.718	0.643	0.732
N	743	743	743	735	672
C Rectangular kernel	0.482	-0.089	-0.189	-0.238	-0.323
Standard error	0.569	0.515	0.778	0.242	0.521
P-value	0.402	0.863	0.810	0.329	0.539
Bandwidth	0.4	0.4	0.3	0.6	0.4
N	172	172	122	266	155
D With controls	0.451	-0.068	0.208	-0.237	0.079
Standard error	0.698	0.515	0.808	0.220	0.532
P-value	0.522	0.895	0.798	0.287	0.883
Bandwidth	0.4	0.4	0.3	0.6	0.4
N	154	154	109	241	139
E Local polynomial					
2nd order	0.195	0.008	-0.121	-0.201	-0.335
Standard error	0.434	0.368	0.698	0.226	0.482
P-value	0.655	0.982	0.863	0.378	0.490
Bandwidth	0.7	0.8	0.5	0.9	0.5
N	305	342	225	382	204
F Global polynomial regressions					
Polynomial 1st order	-0.095	-0.152	0.096	-0.075	0.207
Standard error	0.185	0.176	0.196	0.133	0.190
P-value	0.609	0.390	0.625	0.573	0.279
N	767	767	767	758	693
Polynomial 2nd order	0.129	-0.131	-0.004	-0.043	0.250
Standard error	0.291	0.216	0.332	0.174	0.246
P-value	0.659	0.547	0.991	0.807	0.314
N	767	767	767	758	693
Polynomial 3rd order	0.169	0.106	-0.049	-0.152	-0.206
Standard error	0.345	0.287	0.468	0.225	0.288
P-value	0.626	0.713	0.917	0.503	0.477
N	767	767	767	758	693

G Donut (excluding Hb value 10.9 and 11)	-0.014	0.054	-0.562	0.073	-0.455
Standard error	0.526	0.306	0.765	0.236	0.449
P-value	0.979	0.861	0.466	0.757	0.316
Bandwidth	0.7	0.8	0.5	0.9	0.5
N	290	327	210	367	190

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in panels A-E, the RD coefficients are estimated by fitting a local linear regression separately using a triangular kernel. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5%, 1% level respective.

Table 4.A.2: Average treatment effect for the population close to the cutoff of nutrition information on cognition and education for different bandwidth and functional forms excluding treatment group of the intervention from essay 2

	(1) Block design	(2) Digit span forward	(3) Digit span backward	(4) Progressive matrices	(5) Day and night	(6) Cognitive index	(7) Math	(8) Reading	(9) Attendance
A CCT Bandwidth	-0.202	0.047	-0.541**	0.069	0.474*	-0.204	0.287	0.422	0.055
Standard error	0.418	0.443	0.258	0.314	0.242	0.289	0.307	0.262	0.041
P-value	0.632	0.916	0.041	0.826	0.056	0.485	0.354	0.113	0.186
Bandwidth	0.5	0.3	0.7	0.5	0.6	0.4	0.5	0.5	0.7
N	243	131	328	243	284	187	243	243	307
B Alternative bandwidth									
Bandwidth 0.3	-0.161	0.047	-0.694	0.077	0.632	-0.037	0.231	0.411	0.073
Standard error	0.602	0.443	0.444	0.455	0.392	0.369	0.465	0.392	0.087
P-value	0.790	0.916	0.126	0.867	0.114	0.921	0.622	0.301	0.406
N	131	131	131	131	131	131	131	131	122
Bandwidth 0.5	-0.202	-0.363	-0.607*	0.069	0.510*	-0.178	0.287	0.422	0.069
Standard error	0.418	0.298	0.324	0.314	0.278	0.233	0.307	0.262	0.054
P-value	0.632	0.229	0.067	0.826	0.073	0.447	0.354	0.113	0.202
N	243	243	243	243	243	243	243	243	229
Bandwidth 1.0	-0.189	-0.145	-0.378*	0.117	0.467**	-0.041	0.161	0.194	0.042
Standard error	0.263	0.193	0.207	0.239	0.191	0.165	0.185	0.156	0.036
P-value	0.474	0.457	0.073	0.627	0.018	0.803	0.390	0.220	0.245
N	448	448	448	448	448	448	448	448	423
Bandwidth 1.5	-0.404**	-0.059	-0.203	0.133	0.263	-0.088	-0.052	0.067	0.030
Standard error	0.187	0.149	0.165	0.188	0.176	0.135	0.135	0.119	0.028
P-value	0.035	0.695	0.224	0.481	0.142	0.516	0.701	0.578	0.292
N	724	724	724	724	724	724	723	724	682
Bandwidth 2.0	-0.432**	-0.083	-0.193	0.108	0.203	-0.126	-0.104	0.036	0.023
Standard error	0.174	0.131	0.154	0.178	0.161	0.123	0.120	0.116	0.026
P-value	0.016	0.527	0.214	0.549	0.213	0.314	0.388	0.761	0.378
N	777	777	777	777	777	777	776	777	731
Bandwidth 2.5	-0.319	-0.036	-0.223	0.139	0.336*	-0.038	0.039	0.120	0.036

Standard error	0.207	0.173	0.181	0.203	0.189	0.147	0.153	0.131	0.032
P-value	0.129	0.835	0.224	0.498	0.082	0.799	0.799	0.362	0.266
N	614	614	614	614	614	614	613	614	577
C Rectangular kernel									
	-0.115	-0.124	-0.456*	0.166	0.417*	-0.298	0.210	0.332	0.046
Standard error	0.344	0.444	0.236	0.284	0.223	0.273	0.294	0.252	0.039
P-value	0.739	0.781	0.059	0.561	0.067	0.281	0.479	0.193	0.252
Bandwidth	0.5	0.3	0.7	0.5	0.6	0.4	0.5	0.5	0.7
N	243	131	328	243	284	187	243	243	307
D With controls	-0.318	-0.210	-0.495*	0.014	0.442	-0.119	0.421	0.681**	0.046
Standard error	0.445	0.495	0.294	0.280	0.291	0.346	0.398	0.268	0.048
P-value	0.478	0.674	0.099	0.962	0.136	0.733	0.295	0.014	0.347
Bandwidth	0.5	0.3	0.7	0.5	0.6	0.4	0.5	0.5	0.7
N	217	114	293	217	253	165	217	217	274
E Local polynomial									
2nd order	-0.298	-0.239	-0.658*	0.129	0.406*	-0.153	0.280	0.236	0.049
Standard error	0.403	0.294	0.340	0.279	0.224	0.202	0.265	0.217	0.044
P-value	0.463	0.420	0.059	0.644	0.076	0.454	0.295	0.283	0.280
Bandwidth	0.6	0.5	0.5	0.7	0.8	0.7	0.7	0.7	0.7
N	284	243	243	328	369	328	328	328	307
F Global polynomial regressions									
Polynomial 1st order	-0.403**	-0.138	-0.098	0.072	0.132	-0.135	-0.155	0.018	0.011
Standard error	0.158	0.098	0.136	0.163	0.141	0.112	0.103	0.125	0.024
P-value	0.014	0.166	0.473	0.662	0.351	0.236	0.136	0.885	0.628
N	801	801	801	801	801	801	800	801	755
Polynomial 2nd order	-0.459**	-0.009	-0.127	0.080	0.269	-0.080	-0.075	0.013	0.032
Standard error	0.198	0.136	0.199	0.199	0.189	0.142	0.146	0.132	0.030
P-value	0.024	0.948	0.525	0.690	0.161	0.576	0.611	0.921	0.280
N	801	801	801	801	801	801	800	801	755
Polynomial 3rd order	-0.403*	-0.086	-0.294	0.158	0.344	-0.092	0.101	0.062	0.050
Standard error	0.226	0.211	0.225	0.233	0.225	0.169	0.188	0.181	0.042
P-value	0.081	0.685	0.197	0.502	0.133	0.589	0.592	0.732	0.238
N	801	801	801	801	801	801	800	801	755

G Donut (excluding Hb value 10.9 and 11)	-0.007	-1.292***	-0.415	0.194	0.384	-0.146	0.192	0.186	0.034
Standard error	0.342	0.392	0.439	0.436	0.313	0.217	0.302	0.262	0.049
P-value	0.985	0.002	0.349	0.658	0.225	0.502	0.529	0.481	0.497
Bandwidth	0.6	0.5	0.5	0.7	0.8	0.7	0.7	0.7	0.7
N	269	228	228	313	354	313	313	313	294

N: Number of observations. Each cell represents a different regression. Unless otherwise indicated in Panels A-E the RD coefficients are estimated by fitting a local linear regression separately using a triangular kernel. All specifications allow for different slopes to the left and the right of the cutoff and standard errors are clustered at the school level. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Table 4.A.3: Minimal detectable effects for different bandwidth (no difference in outcome variable)

	Bandwidth						Effect size in other nutrition interventions
	0.3	**0.5**	**1.0**	**1.5**	**2.0**	**2.5**	
Dietary diversity	1.308	1.028	0.822	0.76	0.741	0.712	
Frequency of meat consumption	0.842	0.653	0.522	0.488	0.477	0.459	
Frequency of green vegetable consumption	1.627	1.37	0.997	0.888	0.848	0.809	
Hemoglobin	0.87	0.732	0.547	0.51	0.513	0.495	0.136*[1] 0.151[2] 0.275[3] 0.202**[4] 0.416**[4]
Number anemia symptoms	1.208	1.145	0.817	0.76	0.744	0.692	-0.137[1] -0.157[2]
Block design	1.254	0.829	0.65	0.62	0.581	0.552	0.012[1] 0.045[2]
Digit span forwards	1.205	0.894	0.592	0.574	0.541	0.518	-0.105[1] -0.135[2]
Digit span backwards	1.429	1.057	0.831	0.728	0.682	0.66	0.009[1] -0.23[2]
Progressive matrices	0.779	0.582	0.499	0.462	0.423	0.41	0.070[1] 0.112[2]
Day and night	1.457	1.121	0.861	0.749	0.71	0.676	0.116[1] 0.210[2]
Cognitive index	0.919	0.91	0.69	0.634	0.601	0.58	0.028[1] 0.058[2]
Math	1.273	1.054	0.79	0.729	0.701	0.663	0.112 0.197*[2]
Reading	1.253	1.039	0.711	0.664	0.635	0.603	0.129[1] 0.182*[2]
Attendance	0.265	0.19	0.156	0.147	0.143	0.136	-0.005[1]

1 Effect size from the evaluation of the intervention from essay 2 for the Difference-in-Difference estimates.
2 Effect size from the evaluation of the intervention from essay 2 for the Difference-in-Difference estimates at 90% school attendance.
3 Effect size in Luo et al. (2012), Information experiment 2.
4 Effect size in Luo et al (2012), Experiments 1 and 2, multivitamin supplement treatment arm.
*, **, *** denote significance at the 10%, 5% and 1% level, respectively.

5. References

Alexander, J. M. (2008) *Capabilities and Social Justice: The Political Philosophy of Amartya Sen and Martha Nussbaum*. Ashgate Publishing, Ltd.

Alkire, S. and Foster, J. (2007) 'Counting and Multidimensional Poverty', *OPHI Working Papers*, 7, pp. 77–89. doi: 10.2499/9780896296602BK.

Allen, L. *et al.* (2006) 'Guidelines on Food Fortification With Micronutrients', *Who, Fao Un*, p. 341. doi: 10.1242/jeb.02490.

Altman, D. G. (1985) 'Comparability of Randomized Groups', *Journal for the Royal Statistical Society, Series D (The Statistician)*, 34(1), pp. 125–36.

Andang'o, P. E. A. *et al.* (2007) 'Efficacy of iron-fortified whole maize flour on iron status of schoolchildren in Kenya: a randomised controlled trial.', *The Lancet*, 369(9575), pp. 1799–806. doi: 10.1016/S0140-6736(07)60817-4.

Anderson, P. (2002) 'Assessment and Development of Executive Function (EF) During Childhood', *Child Neuropsychology*, 8(2), pp. 71–82. doi: 10.1076/chin.8.2.71.8724.

Anderson, V. (2001) 'Assessing executive functions in children: biological, psychological, and developmental considerations', *Pediatric Rehabilitation*, 4(3), pp. 119–36. doi: 10.1080/13638490110091347.

Angrist, J. D. and Pischke, J.-S. (2008) 'Mostly harmless econometrics: An empiricist's companion', *An empiricist's companion*, (March), p. 392. doi: 10.1057/be.2009.37.

Ardila, A. *et al.* (2005) 'The Influence of the Parents' Educational Level on the Development of Executive Functions', *Developmental Neuropsychology*, 28(1), pp. 539–60. doi: 10.1207/s15326942dn2801_5.

ASER (2014) *Aser tools & testing*. Available at: http://www.asercentre.org.

Ashraf, N., Field, E. and Lee, J. (2014) 'Household Bargaining and Excess Fertility: An Experimental Study in Zambia', *American Economic Review*, 104(7), pp. 2210 – 2237.

Azizi, F. *et al.* (1995) 'Physical, neuromotor and intellectual impairment in non-cretinous schoolchildren with iodine deficiency', *International Journal for Vitamin and Nutrition Research*, 65(3), pp. 199–205. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8830000.

Baird, S. *et al.* (2003) 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities', *Econometrica*, 14(0), pp. 31–35. doi: 10.1111/j.1468-0262.2004.00481.x.

Banerjee, A. *et al.* (2015) 'The miracle of microfinance? Evidence from a randomized evaluation', *American Economic Journal: Applied Economics*, 7(1), pp. 22–53. doi: 10.1257/app.20130533.

Banerjee, A., Barnhardt, S. and Duflo, E. (2014) 'Nutrition, Iron Deficiency Anemia, and the Demand for Iron Fortified Salt: Evidence from an Experiment in Rural Bihar', in Wise, D. (ed.) *Discoveries in the Economics of Aging*. National Bureau of Economic Research, pp. 343–84.

Banerjee, A., Barnhardt, S. and Duflo, E. (2016) *Can Iron-Fortified Salt Control Anemia? Evidence from Two Experiments in Rural Bihar*. Cambridge, MA. doi: 10.3386/w22121.

Banerjee, A., Duflo, E. and Barnhardt, S. (2015) 'Movies, Margins and Marketing: Encouraging the Adoption of Iron-Fortified Salt', *NBER Working Paper Series*, p. 26. doi:

10.3386/w21616.

Banerjee, A., Duflo, E. and Glennerster, R. (2011) ‘Is Decentralized Iron Fortification a Feasible Option to Fight Anemia Among the Poorest?’, in Wise, D. A. (ed.) *Explorations in the Economics of Aging*. Chicago: University of Chicago Press, pp. 317–44.

Banerjee, A. and Mullainathan, S. (2010) ‘The Shape of Temptation: Implications for the Economic Lives of the Poor’, *National Bureau of Economic Research Working Paper Series*, 0(1990), p. 15973+. doi: 10.2139/ssrn.1598547.

Banerjee, A. V. *et al.* (2010) ‘Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives.’, *BMJ (Clinical research ed.)*, 340(9655), p. c2220. doi: 10.1136/bmj.c2220.

Banerjee, A. V. and Duflo, E. (2012) ‘Poor Economics’, *Poor economics*, p. 303. doi: 10.1007/s13398-014-0173-7.2.

Bartolotta, T. and Shulman, B. (2010) ‘Child development.’, *Current problems in pediatrics*, 23(2), pp. 44–49. doi: 10.1016/0045-9380(93)90002-T.

Bautista, A. *et al.* (1982) ‘The effects of oral iodized oil on intelligence, thyroid status, and somatic growth in school-age children from an area of endemic goiter’, *The American journal of clinical nutrition*, 35(1), pp. 127–34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6278919.

Beard, J. (2003) ‘Iron deficiency alters brain development and functioning’, *The Journal of nutrition*, 133(5 Suppl 1), p. 1468S–72S. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12730445.

Beard, J. L. (2008) ‘Why iron deficiency is important in infant development.’, *The Journal of nutrition*, 138(12), pp. 2534–2536. doi: 138/12/2534 [pii].

de Benoist, B. *et al.* (2008) ‘Iodine Deficiency in 2007: Global Progress since 2003’, *Food and Nutrition Bulletin*, 29(3), pp. 195–202. doi: 10.1177/156482650802900305.

Berk, L. E. (2012) *Child development*. Pearson.

Bhutta, Z. A. *et al.* (2013) ‘Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost?’, *The Lancet*, 382(9890), pp. 452–77. doi: 10.1016/S0140-6736(13)60996-4.

Black, R. E. *et al.* (2016) ‘Reproductive, maternal, newborn, and child health: Key messages from Disease Control Priorities 3rd Edition’, *The Lancet*. doi: 10.1016/S0140-6736(16)00738-8.

Blake, H. *et al.* (2002) ‘An evaluation of screening measures for cognitive impairment after stroke’, *Age and ageing*, 31(6), pp. 451–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12446291.

Bobonis, G. J., Miguel, E. and Puri-Sharma, C. (2006) ‘Anemia and School Participation’, *The Journal of Human Resources*, 41(4), pp. 692–721.

Bortolini, G. A. and Vitolo, M. R. (2011) ‘The impact of systematic dietary counseling during the first year of life on prevalence rates of anemia and iron deficiency at 12-16 months’, *Jornal de Pediatria*, 88(1), pp. 33–9. doi: 10.2223/JPED.2156.

Bourguignon, F. and Chakravarty, S. (2003) ‘The measurement of multidimensional poverty’, *Journal of Economic inequality*, 1(1), pp. 25–49. doi: 10.1007/978-0-387-79253-8.

Calonico, S. *et al.* (2016) ‘Regression discon- tinuity designs using covariates.’, *Working paper. University of Michigan.*

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014) 'Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs', *Econometrica*, 82(6), pp. 2295–2326. doi: 10.3982/ECTA11757.

Carlson, S. M. (2005) 'Developmentally sensitive measures of executive function in preschool children', *Developmental neuropsychology*, 28(2), pp. 595–616. doi: 10.1207/s15326942dn2802_3.

Carpenter, P. A., Just, M. A. and Shell, P. (1990) 'What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test', *Psychological review*, 97(3), pp. 404–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2381998.

Cattaneo, M. D., Titiunik, R. and Vazquez-Bare, G. (2016) 'Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality'.

Cattaneo, M. D. and Vazquez-Bare, G. (2016) 'The Choice of Neighborhood in Regression Discontinuity Designs', *Observational Studies*, 2, pp. 134–146.

Ceda, G. P. *et al.* (1992) 'Regulation of insulin-like growth factor-binding protein expression by thyroid hormone in rat GH3 pituitary tumor cells', *Endocrinology*, 130(3), pp. 1483–9. doi: 10.1210/endo.130.3.1371451.

Chaouki, M. L. and Benmiloud, M. (1994) 'Prevention of iodine deficiency disorders by oral administration of lipiodol during pregnancy', *European Journal of Endocrinology*, 130(6), pp. 547–51. doi: 10.1530/eje.0.1300547.

Chen, C. and Stevenson, H. W. (1988) 'Cross-linguistic differences in digit span of preschool children', *Journal of Experimental Child Psychology*, 46(1), pp. 150–8. doi: 10.1016/0022-0965(88)90027-6.

Childs, F. *et al.* (1997) 'Dietary education and iron deficiency anaemia in the inner city', *Archives of Disease in Childhood*, 76(2), pp. 144–47. doi: 10.1136/adc.76.2.144.

Ciolli, A. (2008) 'Mandatory school vaccinations: The role of tort law', *Yale Journal of Biology and Medicine*, pp. 129–137.

Cohen, J., Dupas, P. and Schaner, S. (2015) 'Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial', *American Economic Review*, 105(2), pp. 609–645. doi: 10.1257/aer.20130267.

Collier, P. and Dollar, D. (2004) 'Development effectiveness: What have we learnt?', *Economic Journal*, 114(496). doi: 10.1111/j.1468-0297.2004.00221.x.

Considine, G. and Zappalà, G. (2002) 'Factors influencing the Educational Performance of Students from Disadvantaged Backgrounds', in Eardley, T. and Bradbury, B. (eds) *Competing Visions: Refereed Proceedings of the National Social Policy Conference 2001, SPRC Report 1/02*. Sydney: Social Policy Research Centre, University of New South Wales, pp. 91–107.

Dasgupta, P. and Ray, D. (1986) 'Inequality as a Determinant of Malnutrition and Unemployment: Theory', *The Economic Journal*, 96(384), pp. 1011–1034. doi: 10.2307/2233329.

Dean Karlan, Margaret McConnell, Sendhil Mullainathan, and J. Z. (2013) 'Getting To the Top of Mind: How Reminders Increase Saving', *Journal of Chemical Information and Modeling*, 53(9), pp. 1689–1699. doi: 10.1017/CBO9781107415324.004.

DiNardo, J. (2008) 'Natural Experiments and Quasi-Natural Experiments', in *New Palgrave Dictionary of Economics*, pp. 1–16. doi: 10.1057/b.9780631218234.2008.X.

Drèze, J. and Sen, A. (2013) *An Uncertain Glory India and Its Contradictions*, *Journal of*

Chemical Information and Modeling. doi: 10.1017/CBO9781107415324.004.

Duflo, E., Glennerster, R. and Kremer, M. (2007) 'Chapter 61 Using Randomization in Development Economics Research: A Toolkit', *Handbook of Development Economics*, pp. 3895–3962. doi: 10.1016/S1573-4471(07)04061-2.

Dunn, J. T. (1992) 'Iodine-deficiency: The next target for elimination?', *New England Journal of Medicine*, 326(4), pp. 267–268.

Dupas, P. (2011a) 'Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya', *American Economic Journal: Applied Economics*, 3(1), pp. 1–34. doi: 10.1257/app.3.1.1.

Dupas, P. (2011b) 'Health Behavior in Developing Countries', *Annual Review of Economics*, 3(1), pp. 425–449. doi: 10.1146/annurev-economics-111809-125029.

Dupas, P. and Robinson, J. (2013) 'Why don't the poor save more? Evidence from health savings experiments', *American Economic Review*, pp. 1138–1171. doi: 10.1257/aer.103.4.1138.

Easterly, W. (2006) 'The White Man's Burden', *Lancet*, p. 2060. doi: 10.1016/S0140-6736(06)68925-3.

Ezzat, S. *et al.* (1991) 'Growth Hormone Regulation in Primary Fetal and Neonatal Rat Pituitary Cell Cultures: The Role of Thyroid Hormone', *Endocrinology*, 128(2), pp. 937–43.

FAO (2007) *Nutrition and Consumer Protection Division. Guidelines for measuring household and individual dietary diversity*. Rome. Available at: http://www.fao.org/3/a-i1983e.pdf.

FAO, IFAD and WFP (2014) *The State of Food Insecurity in the World 2014. Strengthening the enabling environment for food security and nutrition.* Rome. FAO.

FAO and WHO (2002) *Human Vitamin and Mineral Requirements. Report of a joint FAO/WHO expert consultation, Bangkok, Thailand*. Rome. Available at: http://www.fao.org/3/a-y2809e.pdf.

Farebrother, J. *et al.* (2015) 'Systematic review of the effects of iodised salt and iodine supplements on prenatal and postnatal growth: study protocol', *BMJ Open*, 5(4), pp. e007238–e007238. doi: 10.1136/bmjopen-2014-007238.

Fenske, N. *et al.* (2013) 'Understanding Child Stunting in India: A Comprehensive Analysis of Socio-Economic, Nutritional and Environmental Determinants Using Additive Quantile Regression', *PLoS ONE*. Edited by Z. A. Bhutta, 8(11). doi: 10.1371/journal.pone.0078692.

Fernandes, D., Lynch Jr, J. G., & Netemeyer, R. G. (2014) 'Financial literacy, financial education, and downstream financial behaviors', .*Management Science*, 60(8), pp. 1861–1883. doi: 10.1287.

Ferry, R. J., Cerri, R. W. and Cohen, P. (1999) 'Insulin-like growth factor binding proteins: new proteins, new functions', *Hormone research*, 51(2), pp. 53–67. doi: 23315.

Fewtrell, M. S. *et al.* (2008) 'How much loss to follow-up is acceptable in long-term randomised trials and prospective studies?', *Archives of disease in childhood*, 93(6), pp. 458–461. doi: 10.1136/adc.2007.127316.

Food and Agriculture Organization (FAO) (2006) 'The double burden of malnutrition. Case studies from six developing countries.', *FAO Food and Nutrition Paper*, 84, pp. 1–334. Available at: http://www.fao.org/docrep/009/a0442e/a0442e03.htm#bm03.

Fox, R. (2003) 'Food and eating: an anthropological perspective', *Social issues research*

center, pp. 1–22. Available at: http://sirc.org/publik/foxfood.pdf.

Gerstadt, C. L., Hong, Y. J. and Diamond, A. (1994) 'The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test', *Cognition*, 53(2), pp. 129–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7805351.

Giovagnoli, A. R. (2001) 'Relation of sorting impairment to hippocampal damage in temporal lobe epilepsy', *Neuropsychologia*, 39(2), pp. 140–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11163372.

Glewwe, P., Kremer, M. and Moulin, S. (2009) 'Many children left behind? Textbooks and test scores in Kenya', *American Economic Journal: Applied Economics*, 1(1), pp. 112–135. doi: 10.1257/app.1.1.112.

Global Database on Child Growth and Malnutrition (no date) *Child growth standards and cut-off points*. Available at: http://www.who.int/nutgrowthdb/about/introduction/en/index5.html.

Global Panel (2015) 'BIOFORTIFICATION : An Agricultural Investment for Nutrition', *Policy Brief*, (1), pp. 1–8.

Goswmai, S. and Das, K. K. (2015) 'Socio-economic and demographic determinants of childhood anemia', *Jornal de Pediatria*, 91(5), pp. 471–7. doi: 10.1016/j.jped.2014.09.009.

Graham, R. D. *et al.* (2007) 'Nutritious Subsistence Food Systems', *Advances in Agronomy*, pp. 1–74. doi: 10.1016/S0065-2113(04)92001-9.

von Grebmer, K. *et al.* (2014) '2014 Global Hunger Index: The Challenge of Hidden Hunger', *October*, 12, pp. 1–6. doi: 10.2499/9780896299269GHI2010.

Gugerty, M. K. (2007) 'You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya', *Economic Development and Cultural Change*, 55(2), pp. 251–282. doi: 10.1086/508716.

Haas, J. D. *et al.* (2014) 'Double-fortified salt is efficacious in improving indicators of iron deficiency in female Indian tea pickers.', *The Journal of nutrition*, 144(6), pp. 957–64. doi: 10.3945/jn.113.183228.

Hahn, J., Todd, P. and Klaauw, W. (2001) 'Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design', *Econometrica*, 69(1), pp. 201–9. doi: 10.1111/1468-0262.00183.

Hale, J. B., Hoeppner, J.-A. B. and Fiorello, C. A. (2002) 'Analyzing digit span components for assessment of attention processes', *Journal of Psychoeducational Assessment*, 20(2), pp. 128–43.

Halterman, J. S. *et al.* (2001) 'Iron deficiency and cognitive achievement among school-aged children and adolescents in the United States', *Pediatrics*, 107(6), pp. 1381–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11389261.

Horton, S. and Ross, J. (2003) 'The economics of iron deficiency', *Food Policy*, 28(1), pp. 51–75. doi: 10.1016/S0306-9192(02)00070-2.

Hudspeth, W. J. and Pribram, K. H. (1990) 'Stages of brain and cognitive maturation.', *Journal of Educational Psychology*, 82(4), pp. 881–4. doi: 10.1037/0022-0663.82.4.881.

ICCIDD Global Network (2014) *Iodine Deficiency. Where is the problem?* Available at: http://www.iccidd.org/p142000269.html.

ICF Macro (2011) *DHS Final Reports*. Calverton.

Imai, K., King, G. and Stuart, E. A. (2008) 'Misunderstandings between experimentalists and observationalists about causal inference', *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(2), pp. 481–502. doi: 10.1111/j.1467-985X.2007.00527.x.

Imbens, G. W. and Kalyanaraman, K. (2012) 'Optimal Bandwidth Choice for the Regression Discontinuity Estimator', *The Review of Economic Studies*, 79(3), pp. 933–59. doi: 10.1093/restud/rdr043.

Imbens, G. W. and Lemieux, T. (2008) 'Regression discontinuity designs: A guide to practice', *Journal of Econometrics*, 142, pp. 615–35.

Jalan, J. and Somanathan, E. (2008) 'The Importance of Being Informed: Experimental Evidence on Demand for Environmental Quality', *Journal of Development Economics*, 87(1), pp. 14–28.

Jones, J. I. and Clemmons, D. R. (1995) 'Insulin-like growth factors and their binding proteins: biological actions', *Endocrine reviews*, 16(1), pp. 3–34. doi: 10.1210/edrv-16-1-3.

Jooste, P. L. and Strydom, E. (2010) 'Methods for determination of iodine in urine and salt', *Best practice & research. Clinical endocrinology & metabolism*, 24(1), pp. 77–88. doi: 10.1016/j.beem.2009.08.006.

Karlan, D. *et al.* (2006) 'Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines', *The Quarterly Journal of Economics*, 121(2), pp. 635–672. doi: 10.1162/qjec.2006.121.2.635.

Karlan, D., Ratan, A. L. and Zinman, J. (2014) 'Savings by and for the Poor: A research review and agenda', *Review of Income and Wealth*, 60(1), pp. 36–78. doi: 10.1111/roiw.12101.

Kennedy, G. *et al.* (2010) 'Measurement of Dietary Diversity for monitoring the impact of food based approaches 1 , 2 Gina Kennedy, Maylis Razes, Terri Ballard and Marie Claude Dop', *International Symposium on Food and Nutrition Security: Food-based approaches for improving diets and raising levels of nutrition*.

Kennedy, G. L. *et al.* (2007) 'Dietary diversity score is a useful indicator of micronutrient intake in non-breast-feeding Filipino children.', *The Journal of nutrition*, 137(2), pp. 472–7. doi: 137/2/472 [pii].

Kennedy, G., Nantel, G. and Shetty, P. (2003) 'The scourge of "hidden hunger": global dimensions of micronutrient deficiencies', *Food Nutrition and Agriculture*, 32, pp. 8–16. Available at: ftp://ftp.fao.org/docrep../fao/005/y8346m/y8346m01.pdf.

Kent, G. (2005) *Freedom from What: The Human Right to Adequate Food*. Washington D.C.: Georgetown University Press.

Khandker, S. R., Koolwal, G. B. and Samad, H. a. (2010) *Handbook on Impact Evaluation: Quantitative Methods and Practices*, *Learning*. doi: 10.1596/978-0-8213-8028-4.

Kiess, L. (no date) *Nutrition in the Post-2015 Context*. Available at: https://sustainabledevelopment.un.org/content/documents/1936nutrition.pdf.

Kremer, M. and Miguel, E. (2007) 'The Illusion of Sustainability', *The Quarterly Journal of Economics*, 122(3), pp. 1007–1065. doi: 10.1162/qjec.122.3.1007.

Kretchmer, N., Beard, J. L. and Carlson, S. (1996) 'The role of nutrition in the development of normal cognition', *American Journal of Clinical Nutrition*, 63(6).

Laibson, D. (1997) 'Golden eggs and hyperbolic discounting', *Quarterly Journal of Economics*, 112(2), pp. 443–477. doi: 10.1162/003355397555253.

Lal, R. B., Srivastava, V. K. and Chandra, R. (1996) 'A study of spectrum of iodine deficiency disorders in rural area of Uttar Pradesh', *Indian Journal of Public Health*, 40(1), pp. 10–2. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9090894.

Larsen, B., Seitz, A. and Larsen, R. (2000) 'Recovery of cognitive function after remifentanil-propofol anesthesia: a comparison with desflurane and sevoflurane anesthesia', *Anesthesia and Analgesia*, 90(1), pp. 168–174. doi: 10.1097/00132586-200010000-00018.

Lee, D. S. and Lemieux, T. (2010) 'Regression Discontinuity Designs in Economics', *Journal of Economic Literature*, 48(2), pp. 281–355. doi: 10.1257/jel.48.2.281.

Leeds, L. *et al.* (2001) 'A comparison of the new executive functioning domains of the CAMCOG-R with existing tests of executive function in elderly stroke survivors', *Age and ageing*, 30(3), pp. 251–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11443027.

Lezak, M. D. (1995) *Neuropsychological assessment*. 3rd edn. New York: Oxford University Press.

Lozoff, B. *et al.* (1998) 'Behavior of infants with iron-deficiency anemia.', *Child development*, 69(1), pp. 24–36. doi: 10.2307/1132067.

Lozoff, B. (2007) 'Iron deficiency and child development', *Food Nutr Bull*, 28(4), pp. S560–71.

Lozoff, B., Jimenez, E. and Smith, J. B. (2006) 'Double burden of iron deficiency in infancy and low socioeconomic status: a longitudinal analysis of cognitive test scores to age 19 years', *Archives of pediatrics & adolescent medicine*, 160(11), pp. 1108–13. doi: 10.1001/archpedi.160.11.1108.

Ludwig, J. and Miller, D. L. (2007) 'Does Head Start improve children's life chances? Evidence from a regression discontinuity design', *The Quarterly Journal of Economics*, 122(1), pp. 159–208. doi: 10.1162/qjec.122.1.159.

Luo, R. *et al.* (2012) 'The limits of health and nutrition education: Evidence from three randomized-controlled trials in rural China', *CESifo Economic Studies*, 58(2), pp. 385–404. doi: 10.1093/cesifo/ifs023.

Ma, G. *et al.* (2007) 'Iron and zinc deficiencies in China : What is feasible and cost-effective strategy ?', *Public Health Nutrition*, 11(6), pp. 632–638. doi: 10.1017/S1368980007001085.

Ma, G. (2015) 'Food, eating behavior, and culture in Chinese society', *Journal of Ethnic Foods*, 2(4), pp. 195–199. doi: 10.1016/j.jef.2015.11.004.

Mahomed, K. and Gülmezoglu, A. M. (2006) 'Maternal iodine supplements in areas of deficiency', in Mahomed, K. (ed.) *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/14651858.CD000135.pub2.

Malins, A. (1969) *Malins Intelligence Scale for Children; 1969., Indian Psychological Corporation*. Lucknow.

Maluccio, J. A. *et al.* (2009) 'The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults', *The Economic Journal*, 119(537), pp. 734–63. doi: 10.1111/j.1468-0297.2009.02220.x.

Mankiw, N. G. and Taylor, M. P. (2011) *Principles of Economics, Second Edition, Book*. doi: 10.1017/CBO9780511511455.

Mason, J. B. *et al.* (2002) 'Iodine fortification is related to increased weight-for-age and birthweight in children in Asia', *Food and Nutrition Bulletin*, 23(3), pp. 292–308.

McLean, E. *et al.* (2009) 'Worldwide prevalence of anaemia, WHO Vitamin and Mineral

Nutrition Information System, 1993-2005', *Public health nutrition*, 12(4), pp. 444–54. doi: 10.1017/S1368980008002401.

Medical Research Council (2004) *MRC Ethics Guide: Medical Research involving Children*. London. Available at: https://www.mrc.ac.uk/documents/pdf/medical-research-involving-children/.

Mei, Z. and Grummer-Strawn, L. M. (2007) 'Standard deviation of anthropometric Z-scores as a data quality assessment tool using the 2006 WHO growth standards: a cross country analysis', *Bulletin of the World Health Organization*, 85(6), pp. 441–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17639241.

Melse-Boonstra, A. *et al.* (2012) 'Iodine supplementation in pregnancy and its effect on child cognition', *Journal of Trace Elements in Medicine and Biology*, 26(2-3), pp. 134–6. doi: 10.1016/j.jtemb.2012.03.005.

Moreno-Reyes, R. *et al.* (2003) 'Selenium and iodine supplementation of rural Tibetan children affected by Kashin-Beck osteoarthropathy', *The American journal of clinical nutrition*, 78(1), pp. 137–44. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12816783.

Moretti, D. *et al.* (2006) 'Extruded rice fortified with micronized ground ferric pyrophosphate reduces iron deficiency in Indian schoolchildren: a double-blind randomized controlled trial', *The American journal of clinical nutrition*, 84(4), pp. 822–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17023709.

Mudur, G. S. (2013) *Doubts on fortified midday-meal salt - safe but no clear proof it increases haemoglobin: Scientists. The Telegraph*. Available at: https://www.telegraphindia.com/1131210/jsp/nation/story_17663877.jsp#.WMAAPn9u2mo.

Mullainathan, S. and Sharif, E. (2013) *Scarcity: Why Having Too Little Means So Much*. Picador.

Murray, C. J. L. and Lopez, A. D. (2013) 'Measuring the Global Burden of Disease', *New England Journal of Medicine*, 369(5), pp. 448–457. doi: 10.1056/NEJMra1201534.

Nairl, S. *et al.* (2013) 'Impact of double fortified salt on iron and iodine deficient school children (6 to 12 years) of rural Vadodara', *Journal of Public Health and Epidemiology*, 5(9), pp. 370–3.

Neumann, C. G. and Harrison, G. G. (1994) 'Onset and evolution of stunting in infants and children. Examples from the Human Nutrition Collaborative Research Support Program. Kenya and Egypt studies', *European journal of clinical nutrition*, 48 Suppl 1, pp. 90–102. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8005095.

NFHS (2015a) *National Family Health Survey 4. District Fact Sheet Jehanabad Bihar*. Mumbai.

NFHS (2015b) *National Family Health Survey 4. India Fact Sheet*. Mumbai.

Nussbaum, M. C. (2011) 'Creating Capabilities', *The Human Development Approach*, (2011), pp. 17–45. doi: 10.2307/j.ctt2jbt31.

O'Donoghue, T. and Rabin, M. (1999) 'Doing it now or later', *American Economic Review*, 89(1), pp. 103–124. doi: 10.2307/116981.

Orsini, A. *et al.* (1987) 'Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children', *Italian journal of neurological sciences*, 8(6), pp. 539–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3429213.

Osório, M. M. (2002) 'Determinant factors of anemia in children', *Jornal de Pediatria*, 78(4),

pp. 269–78.

Pandav, C. S. *et al.* (2000) 'Validation of spot-testing kits to determine iodine content in salt', *Bulletin of the World Health Organization*, 78(8), pp. 975–80.

Pasricha, S.-R. *et al.* (2010) 'Determinants of anemia among young children in rural India.', *Pediatrics*, 126(1), pp. e140–e149. doi: 10.1542/peds.2009-3108.

Paxson, C. H. and Schady, N. (2007) 'Cognitive Development Among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting', *The Journal of Human Resources*, 42(1), pp. 49–84. doi: 10.2139/ssrn.753548.

Pratham (2015) *ASER 2014 - Annual Status of Education Report: National Findings*. New Delhi.

Qian, M. *et al.* (2005) 'The effects of iodine on intelligence in children: a meta-analysis of studies conducted in China', *Asia Pacific journal of clinical nutrition*, 14(1), pp. 32–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15734706.

Radhika, M. *et al.* (2011) 'Micronized ferric pyrophosphate supplied through extruded rice kernels improves body iron stores in children: a double-blind, randomized, placebo-controlled midday meal feeding trial in Indian schoolchildren', *American Journal of Clinical Nutrition*, 94(5), pp. 1202–10. doi: 10.3945/ajcn.110.007179.

Ramakrishnan, U. *et al.* (2004) 'Multimicronutrient interventions but not vitamin a or iron interventions alone improve child growth: results of 3 meta-analyses.', *The Journal of nutrition*, 134(August), pp. 2592–2602. doi: 134/10/2592 [pii].

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

Rawls, J. (1971) *A Theory Of Justice*, *Saudi Med J*. doi: 10.1073/pnas.0703993104.

Ray, D. (1999) *Development Economics*. Princeton University Press.

Ray, D. (2007) 'Development Economics', *New Palgrave Dictionary of Economics*, (March), pp. 1–31. doi: 10.1057/9780230226203.0385.

Ren, Q. *et al.* (2002) 'Effect of environmental supplementation of iodine on infant mortality and growth in children in Xinjiang, China', *Zhonghua liu xing bing xue za zhi*, 23(3), pp. 198–202. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12411089.

Rippmann, C. E. *et al.* (1997) 'Hemocue, an accurate bedside method of hemoglobin measurement?', *Journal of clinical monitoring*, 13(6), pp. 373–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9495289.

Rivera, J. A. *et al.* (2001) 'Multiple micronutrient supplementation increases the growth of Mexican infants', *The American journal of clinical nutrition*, 74(5), pp. 657–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11684535.

Rogol, A. D., Clark, P. A. and Roemmich, J. N. (2000) 'Growth and pubertal development in children and adolescents: effects of diet and physical activity', *The American journal of clinical nutrition*, 72(2), pp. 521–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10919954.

Rubin, D. B. (1974) 'Estimating causal effects of treatments in randomized and nonrandomized studies.', *Journal of Educational Psychology*, 66(5), pp. 688–701. doi: 10.1037/h0037350.

Sachs, J. D. (2001) 'Macroeconomics and Health: Investing in Health for Economic Development: Report of the Commission on Macroeconomics and Health', *Nature Medicine*,

8(6), pp. 1–200. doi: ISBN 92 4 154550 X (NLM classification: WA 30).

Sachs, J. D. (2005) *The End of Poverty*, *The Penguin Press*. doi: 10.1353/lag.2006.0020.

Salimpoor, V. N. and Desrocher, M. (2006) 'Increasing the utility of EF assessment of executive function in children', *Developmental Disabilities Bulletin*, 34(1-2), pp. 15–42.

Salomon, J. A., Haagsma, J. A. and Davis, A. (2015) 'NDisability weights for the Global Burden of Disease 2013 studyo Title', *Lancet Global Health*, 3, pp. 712–23.

Samuels, M. H. *et al.* (1989) 'The effect of altered thyroid status on pituitary hormone messenger ribonucleic acid concentrations in the rat', *Endocrinology*, 124(5), pp. 2277–82. doi: 10.1210/endo-124-5-2277.

Santos, D. N. *et al.* (2008) 'Determinants of cognitive function in childhood: a cohort study in a middle income context.', *BMC public health*, 8, p. 202. doi: 10.1186/1471-2458-8-202.

Schochet, P. Z. (2009) 'Statistical Power for Regression Discontinuity Designs in Education Evaluations', *Journal of Educational and Behavioral Statistics*, 34(2), pp. 238–266. doi: 10.3102/1076998609332748.

Schultz, T. P. (2004) 'School subsidies for the poor: Evaluating the Mexican Progresa poverty program', *Journal of Development Economics*, 74(1), pp. 199–250. doi: 10.1016/j.jdeveco.2003.12.009.

Semba, R. D. and Delange, F. (2008) 'Iodine Deficiency Disorders', in Semba, R. D., Bloem, M. W., and Piot, P. (eds) *Nutrition and Health in Developing Countries*. Totowa: Humana Press, pp. 507–29.

Sen, A. (1999) 'Development as Freedom', *Oxford Press*, pp. 1–50. doi: 10.1215/0961754X-9-2-350.

Sivakumar, B. *et al.* (2001) 'Prospects of fortification of salt with iron and iodine', *British Journal of Nutrition*, 85(2), p. 167. doi: 10.1049/BJN2000310.

Stein, A. J. *et al.* (2005) 'Analyzing the Health Benefits of Biofortified Staple Crops by Means of the Disability-Adjusted Life Years Approach: a Handbook Focusing on Iron, Zinc and Vitamin A', *HarvestPlus Technical Monograph Series*, 4.

Stein, A. and Qaim, M. (2007) 'The Human and Economic Cost of Human Hunger.', *Food and Nutrition Bulletin*, 2, pp. 125–134.

Strauss, J. and Thomas, D. (1998) 'Health, nutrition, and economic development', *Journal of Economic Literature*, 36(2), pp. 766–817. doi: http://www.aeaweb.org/jel/index.php.

van Stuijvenberg, M. E. *et al.* (1999) 'Effect of iron-, iodine-, and beta-carotene-fortified biscuits on the micronutrient status of primary school children: a randomized controlled trial', *The American journal of clinical nutrition*, 69(3), pp. 497–503. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10075336.

van Stuijvenberg, M. E. (2005) 'Using the School Feeding System as a Vehicle for Micronutrient Fortification: Experience from South Africa', *Food and Nutrition Bulletin*, 26(2 suppl2), pp. S213–S219. doi: 10.1177/15648265050262S212.

Stuss, D. T. and Benson, F. D. (1987) 'The frontal lobes and control of cognition and memory', in Perecman, E. (ed.) *The frontal lobes revisited*. New York: IRBN Press, pp. 141–58.

Sunstein, C. and Thaler, R. (2009) *Nudge: Improving decisions about health, wealth, and happiness.* Penguin.

Thatcher, R. W. (1991) ‘Maturation of the human frontal lobes: Physiological evidence for staging’, *Developmental Neuropsychology*, 7(3), pp. 397–419. doi: 10.1080/87565649109540500.

The Ministry of Human Resource Development (2016a) *Mid Day Meal Scheme. About the Mid Day Meal Scheme*.

The Ministry of Human Resource Development (2016b) *Mid Day Meal Scheme. Meal provision*. Available at: http://mdm.nic.in.

The World Bank (2015) *World Development Report 2015: Mind, society, and behavior, International Bank for Reconstruction and Development / The World Bank*. doi: 10.1596/978-1-4648-0342-0.

Thistlethwaite, D. L. and Campbell, D. T. (1960) ‘Regression-discontinuity analysis: An alternative to the ex post facto experiment’, *Journal of Educational Psychology*, 51(6), pp. 309–317. doi: 10.1037/h0044319.

Thomas, D. (1993) ‘The Distribution of Income and Expenditure within the Household’, *Annales d’Économie et de Statistique*, (29), pp. 109–135. doi: 10.2307/20075898.

Thomas, D. (2011) ‘Intra-Household Resource Allocation: An Inferential Approach’, *Journal of Human Resources*, 25(4), pp. 635–664. doi: 10.2307/145670.

Thornton, R. (2005) ‘The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment’, *Response*, 98(5), p. 0.

Thornton, R. L. (2008) ‘The Demand for Learning HIV Status and the Impact on Sexual Behavior: Evidence from a Field Experiment’, *American Economic Review*, 98(5), pp. 1829–1863. doi: 10.1257/aer.98.5.1829.

Tomer, J. F. *et al.* (2013) ‘Choices, values, and frames’, *Psychological review*, 8(3), pp. 88–96. doi: 10.1037/0033-295X.106.3.529.

Torheim, L. E. *et al.* (2003) ‘Validation of food variety as an indicator of diet quality assessed with a food frequency questionnaire for Western Mali’, *European journal of clinical nutrition*, 57(10), pp. 1283–91. doi: 10.1038/sj.ejcn.1601686.

UNDP (2015) *Human Development Report 2015: Work for human development, Undp*. Available at: http://hdr.undp.org/sites/default/files/2015_human_development_report.pdf.

UNICEF (2008) *Sustainable elimination of iodine deficiency*. New York. Available at: https://www.unicef.org/publications/files/Sustainable_Elimination_of_Iodine_Deficiency.pdf.

UNICEF (2017) *Micronutrients and Hidden Hunger*. Available at: https://www.unicef.org/republicadominicana/english/survival_development_12473.htm.

United Nations (2017) *Sustainable Development Goals: 17 Goals to Transform our World*. Available at: http://www.un.org/sustainabledevelopment/sustainable-development-goals/.

USAID (2008) *Guide to DHS statistics. Demographic and Health Surveys Methodology*. Calverton. Available at: http://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf.

White, G., Ruther, M. and Kahn, J. (2015) *Educational Inequality in India: An Analysis of Gender Differences in Reading and Mathematics*.

WHO (1959) *Iron deficiency aneamia. Report of a study group*. Geneva. Available at: http://apps.who.int/iris/bitstream/10665/40447/1/WHO_TRS_182.pdf?ua=1&ua=1.

WHO (2001) *Iron deficiency anaemia: assessment, prevention and control*. Geneva. Available at: http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/WHO_NHD_01.3/en/index.html.

WHO (2007) *Assessment of iodine deficiency disorders and monitoring their elimination. Third edition*. Available at: http://apps.who.int/iris/bitstream/10665/43781/1/9789241595827_eng.pdf.

WHO (2008) *Worldwide prevalence of anaemia 1993–2005. WHO Global Database on Anaemia*. Geneva. Available at: http://apps.who.int/iris/bitstream/10665/43894/1/9789241596657_eng.pdf.

WHO (2011) *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity*. Geneva. Available at: http://www.who.int/vmnis/indicators/haemoglobin.pdf.

WHO (2017a) *Purchasing Power Parity 2005*.

WHO (2017b) *Q&A: Malnutrition and emergencies*. Available at: http://www.who.int/features/qa/malnutrition-emergencies/en/.

WHO Multicenter Growth Reference Study Group (2006) *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age*. Geneva. Available at: http://www.who.int/childgrowth/publications/en/.

Wilken, D. *et al.* (2012) 'What are the benefits of medical screening and surveillance?', *European Respiratory Review*, pp. 105–111. doi: 10.1183/09059180.00005011.

Wood, W. and Neal, D. T. (2009) 'The habitual consumer', *Journal of Consumer Psychology*, 19(4), pp. 579–592. doi: 10.1016/j.jcps.2009.08.003.

World Bank (2006) 'Repositioning Nutrition as Central to Development', *World*, 13(9), p. 272. doi: 10.1596-978-0-8213-6399-7.

World Bank (2017a) *GNI per capita, PPP (current international $)*.

World Bank (2017b) *Life expectancy at birth, total (years)*.

Zimmermann, M. B. *et al.* (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), pp. 425–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12540404.

Zimmermann, M. B. *et al.* (2007) 'Treatment of iodine deficiency in school-age children increases insulin-like growth factor (IGF)-I and IGF binding protein-3 concentrations and improves somatic growth', *The Journal of clinical endocrinology and metabolism*, 92(2), pp. 437–42. doi: 10.1210/jc.2006-1901.

Zimmermann, M. B. (2009) 'Iodine Deficiency', *Endocrine Reviews*, 30(4), pp. 376–408. doi: 10.1210/er.2009-0011.

Zimmermann, M. B. (2011) 'The role of iodine in human growth and development', *Seminars in cell & developmental biology*, 22(6), pp. 645–52. doi: 10.1016/j.semcdb.2011.07.009.

Zweifel, P., Breyer, F. and Kifmann, M. (2009) *Health economics*, *springer*. doi: 10.1007/978-3-540-68540-1.

6. Eidesstattliche Erklärung

Versicherung gemäß §14 Prüfungs- und Studienordnung für den Promotionsstudiengang Wirtschaftswissenschaften

1. Die Gelegenheit zum vorliegenden Promotionsvorhaben ist mir nicht kommerziell vermittelt worden. Insbesondere habe ich keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

2. Ich versichere, dass ich die eingereichte Dissertation “Nutrition and Child Development in Low- and Middle-Income Countries – Evaluation of Three Micronutrient Interventions” selbstständig und ohne unerlaubte Hilfsmittel verfasst habe; fremde Hilfe habe ich dazu weder unentgeltlich noch entgeltlich entgegengenommen und werde dies auch zukünftig so halten. Anderer als der von mir angegebenen Hilfsmittel und Schriften habe ich mich nicht bedient. Alle wörtlich oder sinngemäß den Schriften anderer Autoren entnommenen Stellen habe ich kenntlich gemacht.

3. Die Richtlinien zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen werden von mir beachtet.

4. Eine entsprechende Promotion wurde an keiner anderen Hochschule im In- oder Ausland beantragt; die eingereichte Dissertation oder Teile von ihr wurden nicht für ein anderes Promotionsvorhaben verwendet.

5. Des Weiteren ist mir bekannt, dass Unwahrhaftigkeiten hinsichtlich der vorstehenden Erklärung die Zulassung zur Promotion ausschließen bzw. später zum Verfahrensabbruch oder zur Rücknahme des erlangten Titels berechtigen.

_____

Date and Signature