Food for thought? Experimental Evidence on the Learning Impacts of a Large-Scale School Feeding Program

Elisabetta Aurino, Aulo Gelli, Clement Adamba, Isaac Osei-Akoto, Harold Alderman[1]

Abstract

There is limited experimental evidence on the effects of large-scale, government-led interventions on human capital in resource-constrained settings. We report results from a randomized trial of the government of Ghana's school feeding. After two years, the program led to moderate average increases in math and literacy standardized scores among pupils in treatment communities, and to larger achievement gains for girls and disadvantaged children and regions. Improvements in child schooling, cognition, and nutrition constituted suggestive impact mechanisms, especially for educationally-disadvantaged groups. The program combined

---

[1] Elisabetta Aurino is research fellow at the Department of Economics and Public Policy, Imperial College London (UK). Aulo Gelli is senior scientist at the International Food Policy Research Institute (US). Clement Adamba is research fellow at the School of Education and Leadership, University of Ghana (Ghana). Isaac Osei-Akoto is senior research fellow at the Institute of Statistical, Social, and Economic Research at the University of Ghana (Ghana). Harold Alderman is senior research fellow at the International Food Policy Research Institute (US). Corresponding author: Aurino. Email: e.aurino@imperial.ac.uk
Online Appendix can be found at http://jhr.uwpress.org/
The data used in this article can be obtained from Elisabetta Aurino or Aulo Gelli.
JEL classification: I210, I24, I250, O120
Disclosure statement: The authors have nothing to disclose (see attached statements).
Acknowledgments: The authors gratefully acknowledge the enabling environment provided by the government of Ghana, particularly the Ghana School Feeding Programme (GSFP), for supporting the trial. Specific individuals include GSFP national coordinator, Hon Siedu; and Paakuna Adamu and his team, including Mrs Kuma Mintah, Mrs Anima Wilson, Mrs Susan Torson, and Mr Kwame Nuako, as well as regional and monitoring coordinators and district desk officers, who also provided significant support. The authors also thank Prof Lesley Drake at the Partnership for Child Development (PCD) for her critical role in making the trial possible; they recognize the key role played by PCD staff and associates in London and Ghana, in particular, Don Bundy, Daniel Mumuni, and Josephine Kiamba, Getrude Ananse-Baiden, Lutuf Abdul-Rahman, Rosanna Agble, Abigail Bondzie, and Fred Amese. Also, they thank Felix Asante, Kwabena M. Bosompem, Gloria Folson, Anthoni Kusi, Daniel K. Arhinful, and Irene Ayi at the University of Ghana. We also wish to thank Jere Behrman, Jacob De Hoop, Chung Truong Hoang, Pedro Rosa Dias, Jan Ross, Franco Sassi, Abhijeet Singh, and participants at seminars at 2017 IPA Conference on Education in Ghana, 2018 Italian Conference of Public Economics, UNICEF Innocenti Research Centre; CSAE 2019 and IZA Workshop on Gender and Family Economics 2019 for their helpful comments.

equitable human capital accumulation with social protection, contributing to the "learning for all" sustainable development agenda.

I. Introduction

Average learning levels for primary school pupils in low- and middle-income countries (LMICs) are dismal: for instance, only 40 percent of students in Sub-Saharan Africa (SSA) master basic literacy and numeracy at the end of primary school (World Bank 2018). Further, large disparities in achievements are present, with children from lower socioeconomic status or rural households, and, sometimes girls, lagging behind the average pupil. This "learning crisis" occurred despite unprecedented expansion in primary school access and completion: in SSA, for example, 78 percent of children at primary school age were enrolled in 2014, up from 58 percent in 1999 (World Bank 2017). Consistent with the principle of "quality education for all" underscored by the Sustainable Development Goal 4 (SDG4), raising average learning achievements in an equitable way is a pressing global educational objective.

Currently, there is very limited rigorous evidence focusing on the effectiveness and distributional impacts of large-scale, government-led interventions on human capital, especially in SSA (Snilstveit et al. 2015). One such intervention is school feeding, which ranks amongst the world's most common forms of social protection (Alderman, Gentilini, and Yemtsov 2018). Every day, about 368 million children receive some form of school feeding globally, for an estimated investment of $70 billion a year (WFP 2013). In SSA, since the early 2000s, many governments have invested in school feeding as a multisectoral strategy involving education, health, and agriculture, with funding mostly stemming from Ministries of Education (Alderman and Bundy 2012; Drake et al. 2017). At an average cost of US$54 and US$82 per child per year in low- and middle-income countries, respectively, and often with limited poverty targeting, the share of the educational budgets devoted to school feeding is often considerable (Aulo Gelli and Daryanani 2013).

This paper experimentally addresses whether large-scale, government-led school feeding programs can contribute to equitable learning goals in resource-constrained settings. While school feeding has a robust track record in increasing school participation (Kristjansson et al. 2015; Drake et al. 2017), experimental evidence on its effectiveness on learning is more limited, and with mixed findings (see Appendix 1 and Snilstveit et al. [2015] for a meta-analysis).

Importantly, existing experiments have evaluated small-scale programs implemented as part of international food assistance, usually by the World Food Programme (WFP) or other international non-governmental organizations (NGOs), or by researchers. As an illustration, of the 16 randomized control trials (RCTs) of school feeding in LMICs published since 1980 onwards that we were able to find, all assessed programs implemented by either NGOs or researchers. Additionally, only three were implemented in more than one district in the country (Table A.1, Appendix 1). As noted by Vivalt (2019), smaller-size programs or programs implemented by academics and international organizations tend to report larger effect sizes than government-led ones reaching large populations daily. The latter may suffer from additional challenges as compared to smaller-scale interventions, including: market equilibrium effects and spillovers (Acemoglu 2010; Filmer et al. 2018); endogenous political economy reactions (Bold et al. 2018); heterogeneity by site or in organizational effectiveness (Allcott and Mullainathan 2012; Vivalt 2015); and scale-related implementation issues, including poor monitoring, limited administrative capacity or bureaucrat incentives for the proper functioning of the program (Deaton 2010; Muralidharan and Niehaus 2017; Berry et al. 2018). Further, existing experiments were conducted during limited time periods between baseline and follow-up, and programs often employed complex or unsustainable supply chain logistics (e.g., menus including perishable and/or higher-cost foods). Given these issues, the generalizability of existing evidence stemming from smaller-scale, internally-valid trials may fail to translate to "real-world" programs reaching millions of children daily through sizable government budgetary expenditures.

Further, in contexts characterized by widespread food insecurity, school feeding programs, through targeting the transfer directly at the child conditional on school attendance, may be more effective in raising learning through lowering educational costs and tackling child short-term hunger than alternative social protection measures targeting households, such as cash transfers or generalized food assistance. The most vulnerable groups of learners, such as girls and children from economically-disadvantaged households and geographical areas, may benefit disproportionally more from the transfer than less disadvantaged pupils. This is because the transfer may induce steeper declines in the marginal opportunity costs of human capital investments for these groups, as compared to the average child (Akresh, Walque, and Kazianga 2013; Björkman-Nyqvist 2013). Yet, heterogeneity analysis focusing on educationally-vulnerable groups constitutes another under-investigated topic within the literature on educational interventions in LMICs (Evans and Yuan 2018; Bashir et al. 2018; World Bank 2018). In the case of school feeding, our review of evidence in Appendix 1 shows that there is a lack of systematic investigation of heterogeneity across gender and socio-economic status in existing experimental evidence.

We tackle these questions by evaluating the average and distributional effects of the Ghana School Feeding Programme (GSFP) on child learning achievements. The GSFP provides a free, hot-cooked daily meal to over two million pupils in government primary schools across[i]. In collaboration with the government, we conducted a randomized control trial designed around the re-targeting and scale-up of the GSFP to the most food insecure districts in all regions of Ghana. While the overall trial was aimed at assessing program impacts on education, nutrition, and agriculture[ii] (see Gelli et al., 2016), here we report on both treatment effects on pre-specified educational outcomes, including child math and literacy scores, and heterogeneity in treatment effects on per-protocol population subgroups.

Ghana's learning challenges are similar to the ones currently faced by many other LMICs. First, while the government's efforts to raise schooling in the 2000s resulted in primary enrolment rates that are among the highest in SSA, average learning levels remain disappointingly low: a 2017 study highlighted that more than 80 percent and 70 percent of Grade 2 and Grade 4 students, respectively, could not read a single familiar word or perform a two-digit subtraction (World Bank 2018). Second, wide inequalities in learning exist by gender, poverty, and place of residence (World Bank 2018). Further, Ghana is highly varied in terms of agroecology, ethnicity, socioeconomics, as well as political and administrative capacities. Uncovering the average effect of the program, in face of this diversity and potential heterogeneous program implementation and monitoring across regions, is therefore of interest for policy makers operating in settings characterized by high heterogeneity in administrative and socio-economic environments.

Following the methodology outlined in our protocol (A. Gelli et al. 2016), we document the following intent-to-treat (ITT) findings. After almost two academic years of implementation, exposure to school feeding led to average increases in math, literacy and a composite score of learning by about 0.15 standard deviations (s.d., hereafter). Turning to impact heterogeneity, we find strong variation in program impact in favor of the groups that are at higher risk of falling behind in terms of learning. Girls' math, literacy, and learning composite scores increased by more than 0.2s.d. in school feeding communities compared to controls. Treatment effects among children living in the northern regions, the country's most disadvantaged areas, and for children from households below the poverty line at baseline ranged between 0.25s.d. and 0.3s.d. across all scores. These findings are likely to correspond to lower bounds of potential effects, as program take-up was imperfect, and some implementation challenges were present. The latter mostly related to delays in financial disbursements to the caterers which are in charge of procuring food and cooking and serving the meals. The school feeding intervention also led to increases in grade attainment for the average child, while it promoted enrolment among children from poorest households and regions. In line with the results on learning, cognitive scores of attention span and short-term memory also improved moderately for the average child, while they expanded more markedly for educationally vulnerable groups. Nutritional outcomes also improved for girls and poorest children in treatment communities (Aulo Gelli et al. 2019).

To the best of our knowledge, this is the first large-scale RCT from a LMIC context, investigating the effects of a nationally-mandated, government-led program on educational attainments. Compared to existing trials on school feeding, the importance of this study lies in showing the social protection-*cum*-human capital accumulation of a government school feeding program. As discussed, the issue of implementation at scale is of critical importance as treatment effects tend to decrease with the size of the implementing organization (Muralidharan and Niehaus 2017; Vivalt 2019). Regularity and quality in the provision of school meals is in fact critical for the effectiveness of the program, as children and parents may respond to irregular or lower quality meal provision in multiple ways (e.g. going home for lunch and not returning to school afterwards, changing school, or not attending at all). So far, the effectiveness of full-scale government-led school feeding on learning in LMICs has been previously assessed only in the context of India, and with quasi-experimental methods. Specifically, Chakraborty and Jayaraman (2019) exploited staggered program implementation to identify positive effects of the local "midday-meal" scheme on math and literacy, finding moderate and positive average effects but no treatment effect heterogeneity by gender or household assets. We also add to the school feeding literature, and, more broadly, to the one on educational and social protection interventions in LMICs, by providing evidence on treatment effects heterogeneity. Our findings suggest that in contexts characterized by wide educational inequalities such as Ghana, school feeding programs can contribute to "levelling the playing field" by raising learning outcomes, especially among children at the margin (Jukes, Drake, and Bundy 2008). More generally, this study highlights the importance of social protection on human capital. In LMICs, existing evidence has overwhelmingly focused on schooling (enrolment and attendance), rather than on learning achievements (Baird et al. 2014).

This paper is organized as follows. The next section presents the background and the study design. Then, Section III illustrates the data and identification strategy. Sections IV and V, respectively, present the ITT estimates and potential mechanisms for impact. Section VI presents some robustness checks, while Section VII concludes, including a concise discussion of costs.

II. Background and Study Design

A. Educational setting and the GSFP

Despite the rapid economic growth of Ghana in the past decades, food insecurity and poverty are widespread, particularly in rural areas. During the 2000s, the country prioritized school participation through various initiatives, including the GSFP. For example, it made basic education compulsory between 5 and 15 years. These efforts resulted in a substantial expansion of basic education, with primary enrolment increasing from 61 percent in 1999 to 87 percent in 2016 (World Bank 2017). Despite these impressive achievements, an estimated 300,000 to 800,000 children are still out of primary school, mostly from households below the poverty line and from the country's northern regions (UNDP Ghana 2015). Moreover, Ghana's success in expanding schooling has not been matched by corresponding improvements in learning, which remain overwhelmingly low as compared to international standards (Ministry of Education/RTI International 2014). Wide inequalities in achievements exist by gender, poverty, and place of residence (northern *viz*. southern regions) (World Bank 2018).

The government of Ghana initiated the GSFP in 2005 with a 4-year program budget of over US$200 million (GSFP 2006). Funding for the program is now integrated to the government annual budget. GSFP coordination and implementation are undertaken by a National Secretariat, with program oversight provided by the Ministry of Gender, Children and Social Protection. The program is decentralized; private caterers are awarded contracts by the GSFP to procure, prepare, and serve food to pupils in the targeted schools. Cash transfers (and, recently, electronic payments) are made from the District Assemblies to caterers based on 54 Ghana pesewas per child per day (circa US$0.33) every two weeks. Each caterer is responsible for procuring food from the market on a competitive basis, preparing school meals and distributing food to pupils. Supervision at the school level is undertaken by the School Implementing Committees. Delayed reimbursements to caterers are common, with delays as long as half a year or even a whole year (SEND-Ghana 2013). Delayed payments to caterers often result in caterers reducing the quantity or quality of food provided, or adjusting the school feeding menus, thus likely influencing program quality, and potentially, effectiveness (Ghana Institute of Management and Public Administration 2011).

B. Evaluation design

The trial was designed around the scale-up of the GSFP based on a retargeting exercise conducted in 2012. The government's decision to retarget the GSFP followed a report that highlighted that the program overwhelmingly benefited non-poor households, with only 21% of benefits accruing to poor families (World Bank 2012). The retargeting was guided by the development of poverty and food insecurity rankings to assess potential priority districts across the whole of Ghana. Rankings were used to generate district-level indices on the share of national poverty and food insecurity, through which 58 priority districts (out of the country's 170 at the time of this exercise) were identified for the scale-up of GSFP (see Gelli et al., [2016], for details).

The trial focused on assessing program effects on children (education, health, cognition) and district-level agriculture. In order to affect these different sets of outcomes, program components were delivered at different administrative levels: the school feeding service (which was hypothesized to affect mostly child education and health) was designed to be delivered at the school-level, while the agriculture-related activities were delivered at the district-level, also affecting communities that would not offer school feeding. A complex study design was therefore required, which was achieved through a multiple-level randomization (Figure 1). Due to the relatively small number of clusters in the study, to ensure balance in the comparison groups, the random allocation was undertaken through restricted randomization by modelling selection using a set of school- and village-level variables obtained through the Education Management Information System annual school census data from 2011-2012 (for details, see Gelli et al. [2016]). This step utilized a list from the GSFP secretariat of schools not currently covered by the GSFP. Following Hayes and Moulton (2017), the variables in the restricted randomization were selected on availability and their potential influence on the main study outcomes. The variables in the model included school enrollment, gender ratio, classroom numbers and infrastructure conditions, accessibility, and NGO support. A program was developed to identify the best combination of pairs of schools in each district, by regressing selection status on the school- and village-level covariates over 2,000 random allocations and selecting the combination of two groups of schools that minimised the R-squared statistic for the predicted selection. The allocation into school feeding and control groups was then randomized. A similar restricted randomization procedure was developed to allocate the school feeding group into two sub-groups (GSFP and HGSF) based on additional variables that characterized the agricultural environment, including agroecological zone, maize productivity, employment, alongside the education related variables. We note that the schools were selected from separate communities. The distance between communities is geographically wide enough to minimise cross-community school enrolment, as each two villages are at least 6km apart.

Using a household census at baseline, approximately twenty-five households with children in the 5-15 target-age group were then randomly selected for interview from each community. Control communities were scheduled to receive school feeding at the end of the trial. The multi-level design compared the child-level outcomes (e.g., education, health) between children belonging to school feeding (including both GSFP and HGSF modalities) and control communities, and the agriculture impacts of the HGSF pilot relative to the regular GSFP at district-level[iii].

[FIGURE 1 ABOUT HERE]

III. Data and Sample Description

A. Timeline and Sample

A baseline survey was undertaken in 116 communities between June and September 2013. Due to an error in the lists received by the GSFP, twenty-five schools in the study population, including approximately 18 percent of children in the target age group (5-15 years), had already been receiving school meals at baseline and were removed from the study population. We excluded these schools (13 controls, 12 treatment) from the follow-up. Analysis of child and household characteristics show that the excluded communities were more likely to be rural and located in the north of Ghana, and households to be slightly worse off in terms of some socio-demographic characteristics. Children from excluded communities had lower learning achievements, although all these differences were not large (Online Appendix 1). Two additional communities from the same district in Northern Ghana were excluded from the endline survey, due to logistical problems related to local insecurity.

Implementation in most treatment communities started in the academic year 2014/15, due to bureaucratic delays (see Section III.C). The follow-up survey was conducted in February-March 2016. Given that the academic year in Ghana usually runs from August to May, the program was evaluated after roughly two academic years of implementation.

Both rounds of surveys included detailed modules on household demographics, farm and other assets, expenditures, farming and other economic activities, child anthropometry, and child self-reported[iv] education indicators for all target-age children in the household, including enrolment, attendance and grade attainment, and educational achievement tests. Of the 4,269 target-age children sampled in 2013, 836 were in the last year of primary school or had already completed primary school. As such, they were not eligible to receive the GSFP intervention when implementation began and were threfore excluded from the sample. After three years, we successfully reinterviewed 92 percent of children of target-age and eligible to receive school feeding, leading to a longitudinal sample of 3,170 children. Data on schools and caterers were also collected.

B. Balance of Baseline Covariates and Attrition at Endline

Table 1 presents descriptive statistics of characteristics of the baseline sample by treatment arm. The average child was about 8.5 years old, with children from the school feeding arm on average a month older than control. Almost all children were enrolled in school at baseline, and a tenth of them attended private schools. The average child had completed less than two years of schooling, and about 11 percent had repeated a grade. Along with the descriptive statistics, we present a balance test to assess whether the randomization was successful in achieving balance of baseline covariates. The only difference between the two groups that was statistically significant at 10 percent level was age of household heads in the school feeding arm, which were about one-year-and-a-half older than control communities. These findings, together with the relatively small size of the differences, suggests that the randomization was overall successful in achieving balance.

[TABLE 1 ABOUT HERE]

Table 2 presents analysis of attrition at the child level. We do not observe any inbalance in the probability of remaining in the longitudinal sample based on offer of school feeding[v]. Column 2 presents analysis of whether baseline test scores were associated with children with higher baseline test scores being more likely to be resurveyed, which did not appear to be the case. Finally, column 3 investigates whether the treatment was associated with some child characteristics in predicting likelihood of remaining in the sample. We did so by interacting treatment assignment with the background characteristics we use for heterogeneity analysis. This time, we find a joint significance of all regressors at the 5 percent level. Also, the interaction between treatment and children from poor households was moderately significant, as children from poor households in treatment areas were slightly more likely to be re-interviewed at endline (93 percent of baseline children were followed up in treatment communities, viz. 91 percent in control areas, translating in, respectively, for a total of 22 additional children lost in control areas compared to treatment). Also, boys and children from northern regions were slighlty more likely to be re-interviewed. To evaluate further the possible effects of potential attrition bias on validity of the impact estimates, the table in Online Appendix 2 presents the balance of baseline and endline characteristics across treatment groups for the full longitudinal sample, as well as for the longitudinal sample stratified by gender, household poverty and northern region. Across a wide range of baseline child and household backgrounds, there were no differences between school feeding and control arms in key characteristics at both baseline and endline for the longitudinal sample. The only exception is age in months for children from poor households, whereby children in school feeding areas from poor households were older at both baseline and endline than children in control areas. We address this issue by employing age-standardising test scores, as highlighted in section III.D. Thus, even if there was some concern of differential attrition by treatment in the case of children from poorest households, balance was generally maintained, particularly in light of the relatively low levels of attrition overall, which lessens concerns of a change in the sampling frame by treatment assignment due to attrition.

[TABLE 2 ABOUT HERE]

C. Program Uptake and Implementation

61 percent of eligible children at baseline in treatment areas reported receiving school meals in the previous week at endline, which we refer to as overall uptake rate. The uptake rate was 83 percent for those in public primary education, indicating that most children that were still in basic government education (where the program is served) did in fact receive school meals. . On the other hand, fewer than 2 percent of children in control areas were found to have received school feeding at endline, ruling out the possibility of significant crossover, which

would have hampered the experimental design. We also checked whether the introduction of the program led to children in treatment communities to switch from private to public schools to receive the program, but we did not find evidence of such instance (available upon request).

As the indicator of program uptake was self-reported by the child (or the caregiver in case of young children), we cross-checked it with mean uptake at the community level, to assess whether responses from children living in the same communities were consistent. For 80% of the communities, mean uptake was more than 70% (with half of them having an average uptake exceeding 90%) (results available upon request). Only 4 communities (all located in the south of Ghana) had average uptake below a quarter of all eligible children, which may be a signal of poor implementation

Eighty percent of children that reported receiving school feeding in the treatment arm at endline ate at school during all days in the previous week, suggesting a fairly regular service provision. Twenty-three percent of children in the treatment group reported they were more likely to eat less food at home on days they eat at school, indicating some substitution between meals. However, only four percent reported to bring their food from the school meal to share at home. Online Appendix 3 presents correlates of child endline program uptake (independent of primary enrolment status) among children in treatment communities. Children aged 5-11 years at baseline were two times more likely to receive school feeding compared to adolescents (12-15 years at baseline), coherently with expectations of older children having progressed to secondary school or being out of school. There was no gender variation in the odds of uptake, while household poverty at baseline and northern regions were predictive of about 2 times higher chances of reporting school meals receipt. Baseline math and literacy scores were positively associated with lower odds of school feeding. This finding may be due to faster progression to secondary school for pupils that had higher achievements at baseline. The type of program to which a child was assigned (e.g, standard GSFP *viz.* HGSF pilot) was not predictive of uptake, which reassures about potential concerns of implementation variation between the two school feeding modalities.

Data at the school level show that for some schools, the program started as originally planned in the first semester of 2013, but for the majority of schools (n=30), the program started in the early months of 2014. Only one school started in February 2015. There was no indication of discontinuation of the program, however only 37% of schools reported having a copy of the district GSFP menu, potentially signalling varying adherence to the nutritional guidelines set by the GSFP secretariat (results available upon request). No regional differences were evident. Nearly 85 percent of caterers of treatment schools indicated that often payments were insufficient to cover operational costs, which led them to recur to credit to avoid changing the content and size of meals (83 percent), cutting on portion sizes (9 percent), or adopting a mix of other strategies to reduce costs (e.g. reduce personnel).

We do not have access to administrative data on program implementation. Nonetheless, to further understand whether the financial challenges experienced by the caterers translated into poor-quality meals, we analyzed data from the caterers' weekly meal logs, which provided the ingredients used for the meal served during the survey day and the following day. The most frequent meal served was the combination of a starchy food (e.g. rice, yam, gari, etc.) with some type of legumes (46% of meals), followed by a stew or a soup combining starchy foods and animal-source proteins (mostly dry fish, chicken or meat) (37% of meals), and a starch with vegetables, mostly okra or tomato (9% of meals). All these meals are consistent with the GSFP menu. Only in one school in the Brong Ahafo region the caterer reported serving no meal in both days, while in only three separate instances (but only for one of the two meals surveyed) the caterer reported to have served only a starchy food. Figure 2 presents meal content by region, showing some degree of regional variation in meal nutritional quality. Although we do not have data on quantities served per child, these descriptive findings suggest that at least the implementation guidelines regarding food diversity seemed to have been followed in most cases. Further, focus groups with children, caregivers, teachers and caterers did not highlight particular irregularities in service provision (Fernandes et al. 2017).

[FIGURE 2 HERE]

D. Measures of Child Learning

Given the wide age range included in the target sample, learning assessments evaluated a basic set of skills in literacy and math. Each section of the test began with basic domain-specific questions that progressively increased in difficulty in order to cover different ability levels. The math assessment included questions on recognition of single or double-digit numbers, basic arithmetics, fractions, and basic problems (e.g., how many minutes/hours in 120 minutes), while the literacy test assessed letter recognition, reading short words and sentences, and three final questions on completing a sentence with the correct item among four possible choices. The same 15-item math and literacy tests were administered across rounds. Tests were administered at home to ensure that even children out of school were tested, enhancing internal validity. Parents or schools did not have any knowledge of the test, nor of its specific date and timing.

Test scores were standardized by child age in months for each survey round, with the control group having mean 0 and standard deviation 1, in order to deal with the wide age groups assessed as part of the evaluation. In line with the literature (e.g., Banerjee et al. [2007] and Das et al. [2013]), this was achieved first by removing interviewer effects from the raw scores through a OLS regression on interviewer dummies[vi]. The residuals from these regressions were non-parametrically estimated to obtain age-conditional means and standard deviations. We also generated a composite indicator of learning to address potential issues related to multiple testing, which should enhance statistical power to detect effects that go in the same direction (Kling, Liebman, and Katz 2006). We computed this index as an average from the normalized test scores, and then standardized again to the control group within each round[vii]. In this way, estimated ITT effects can be interpreted as the effect size relative to the control group (Banerjee et al. 2015).

Table 3 presents descriptive statistics of raw and age-standardized tests scores in the two learning domains by intervention arm for the longitudinal sample. School feeding children had larger scores in both rounds, with the difference from control being more pronounced at endline. However, none of the differences prior to the beginning of the intervention appeared to be statistically distinguishable from zero[viii]. The analysis of the raw scores highlights the low achievement levels in each outcome and survey round: at baseline, on average, children were not even able to respond to two out of 15 questions in the math and literacy tests. This proportion increased slightly three years later. However, raw endline scores were still very low, with the average pupil only being able to respond to about 4 out of 15 correct questions for math and literacy, confirming Ghana's learning challenges. Consistent with these average low levels of achievements, there were no ceiling effects by age at endline due to the test design: for instance, children between five and 10 years responded correctly to three questions for both math and literacy, while children aged between 11 and 15 years were able to correctly answer to 5 questions. The analysis of age-standardized test scores at endline highlight the progress of children in the school feeding arm across all competencies.

[TABLE 3 ABOUT HERE]

Figure 3 presents the non-parametric distributions of raw (Panel A) and age-standardized (Panel B) scores in math and literacy by treatment arm at both rounds. Floor effects were present, particularly in the baseline data, highlighting the tests were challenging, particularly for the younger children. A basic reading assessment in Ghana reported similar floor effects, whereby 42 percent and 20 percent of Grade 3 and Grade 6 students, respectively, did not respond correctly to any of the test's six questions (Balwanz and Darvas 2013). Second, there was an improvement in mean achievement in both competences between baseline and endline, although achievements were widely dispersed across the sample. This may reflect alleviation of the floor effects by the endline, as well as widening of educational inequalities in the transition from primary to higher levels of education, by which the most vulnerable children tend to enter the labor market while the others progress to secondary school (De Groot et al. 2015). Third, the distribution of age-standardized achievements of the school feeding group appeared to be above the one of control at endline across the mid- to upper-end of the distribution of both math and literacy.

[FIGURE 3 HERE]

Online Appendix 5 present raw scores at both rounds by child gender, household poverty status and by residence (south vs. north Ghana). At both rounds, there were no significant and large differences between girls and boys, while gaps between nonpoor and poor children were evident across all outcomes. The greatest disparities in baseline raw achievements, however, were between northern and southern regions, underscoring important geographic inequalities in educational quality between north and south Ghana. Children from the southern regions had, on average, responded to about one additional question than northern peers across both competences. This gap substantially reduced or closed at endline. Figure 4 presents empirical distributions of age-standardized test scores by gender (Panel A), poverty (Panel B) and place of residence (Panel C). While at baseline the distribution of achievements tended to overlap between treatment and control group, highlighting balance of outcomes between treatment and control by those factors prior to the start of the program, the nonparametric distributions for the school feeding group often tended to shift toward the right at endline, particularly across the mid- to upper-ends of the distribution, indicating larger gains in learning and cognition for school feeding children after two academic years of program exposure, as compared to control.

[FIGURE 4 HERE]

Autocorrelations of test scores between baseline and follow-up were low (math: ρ =0.23; literacy: ρ =0.31, all significant at <0.01)[ix] . This finding may be partially explained by some degree of measurement error, and partly by the three-year lag between the assessments. We checked with a different longitudinal dataset (the Young Lives study from Ethiopia, India, Peru and Vietnam) whether low autocorrelation among test scores in different waves is common. We found that autocorrelation in vocabulary scores between 5 and 8 years in this sample was also low and roughly comparable to the one related to our literacy scores ($\rho = 0.38$).

E. Identification

We assessed program impact through an ITT approach by comparing test scores between eligible children that were in communities randomly assigned to the school feeding program or the control. The ITT parameter represents the average effect of offering school feeding to children that were eligible to the program at baseline in treatment communities, regardless of whether they actually had school lunches at endline. Given the imperfect program uptake documented in Section 3.3, ITT parameters may likely represent a lower bound for potential program effects. Yet, they are of high policy relevance as program offering can only partially influence uptake.

In the analysis plan we outlined two potential strategies to estimate the ITT parameters, depending on outcomes of interest: ANCOVA and difference-in-differences (DiD). The former improves statistical power by conditioning the endline outcome on the assignment to treatment and the baseline value of the outcome. Following McKenzie (2012) and Frison and Pocock (1992), this is our preferred estimator due to its greater efficiency (defined as retaining unbiasdness with lower variance) in estimating average treatment effects with experimental data, as compared to DiD or a post-estimator approach. Gains in efficiency are more marked when outcomes have low autocorrelation, as in our case. In econometric terms, we estimate the following model:

$$y_{it,j} \; = \alpha_{it,j} + \beta_1 SF_{it,j} + \; \beta_2 y_{i(t-1),j} + \theta_r + \; \varepsilon_{it,j} \quad (1),$$

Where y_{it}, j and $y_{i(t-1),j}$ represent, respectively, the endline and baseline test scores (when available)[x] for child *i* residing in community *j*; $SF_{it,j}$is a dichotomous variable for a child residing in a community randomly assigned to school feeding and thus uncorrelated with $y_{i(t-1),j}$; and θ_r is a vector of region dummies to capture region-specific unobservable characteristics or potential regional variation in quality of implementation. Standard errors were clustered at the community level, which is the unit of randomization for school feeding. β_1, the coefficient related to school feeding, provides the estimate of the treatment effects. Although we analyze treatment effects on pre-specified outcomes, and we also generated a composite index of learning, we further address multiple hyphotesis testing by adjusting p-values through the Romano-Wolf step-down method (Romano and Wolf 2005, 2016) using 2,000 iterations and clustering by community.

IV. Impact of School Feeding on Learning

Table 4, Panel A presents ITT estimates for the full sample. The randomized offer of school feeding led to moderately significant increases across all test scores (of about 0.15 s.d.), after adjusting for multiple hyphotesis testing. We then investigate heterogeneity in program effects. Table 4, Panels B, C, and D, respectively, report treatment effects in models that stratify for child gender, household poverty, and geographical regions, so that we can evaluate total program effects for policy-relevant sub-populations[xi]. School feeding led to sizeable and statistically significant learning gains across all competencies for girls, children from households below the poverty line, and from northern Ghana. In the case of girls, math and literacy scores increased by 0.24 s.d. and 0.2 s.d., respectively. By contrast, the program had a much smaller and not significant effect for boys. For children from households below the poverty line at baseline (Panel C), gains in math and in the composite scores amounted to a third of a standard deviation, while the increases in literacy accounted to 0.23s.d. . Similarly, children from northern regions had moderately significant increases in math and literacy accounting to a quarter of a standard deviation each. As for boys, gains among children from better-off households or regions were smaller and never significant. For completeness, we provide DiD coefficients in Online Appendix 6. While the treatment effects arising from both ANCOVA and DiD are in most cases similar, as anticipated, the former estimator proved being more efficient compared to DiD.

[TABLE 4 ABOUT HERE]

We also investigated variation in treatment effects by age in Online Appendix 7. The latter shows that the effect of school feeding was mostly similar, with the exception of math, between children of different age groups at baseline. However, in the younger cohort (children that were aged 6-11 years at baseline), effects were more precisely estimated, probably due to larger sample sizes. Finally, although it was not part of the analysis plan, we investigated variation by treatment modality in Online Appendix 8. No substantial differences in treatment effects on educational achievements between the GSFP and HGSF were detectable.

V. Mechanisms

While the RCT was designed to investigate educational outcomes in terms of learning, we offer a supportive exploration of possible mechanisms. Improved schooling, nutritional status, and cognitive capacities constitute potential channels through which school feeding can affect learning. First, school meals may promote enrolment, attendance and grade attainment by subsidizing educational costs through the provision of a free meal conditional on attendance. Second, by addressing hunger and micronutrient deficiencies, school feeding can positively affect children's learning via reduced morbidity-related absenteeism, better nutritional status, and increased cognitive skills in the classroom, including increased attention and memory(Kristjansson et al. 2015; Afridi, Barooah, and Somanathan 2019). Further, it may be plausible that teachers can be more motivated by interacting with more attentive and responsive pupils (Afridi, Barooah, and Somanathan 2013; Glewwe and Kremer 2006). The potential health impacts of school feeding may be offset by substitution between meals, or changes in the intrahousehold distribution of food, as this could be diverted away from the child receiving the free meal, though evidence of this effect is inconclusive (Jacoby 2002; Ahmed 2004; Chakraborty and Jayaraman 2019; Harounan Kazianga, de Walque, and Alderman 2014). Also, high heterogeneity in the health pathway may be present, with effects mostly concentrated among malnourished children (Krämer, Kumar, and Vollmer 2018; Powell et al. 1998). The remainder of this section investigates the role of these intermediate outcomes as potential pathways for program effects. Descriptive statistics are presented in Online Appendix 9.

A. Changes in Schooling

Table 5 presents ITT estimates of school feeding on the following indicators: school enrolment in any educational level; school attendance (conditional on enrolment), as measured by the number of days the child attended school out of a five-day week; and current grade attended by the child. All of these variables were measured in the household survey with questions directed to the child or her caregiver (for young children) in both survey rounds. Although these outcomes are included in the study protocol as key schooling outcomes potentially affected by the intervention, we also compute Romano-Wolf adjusted p-values. Panel A reports estimates of school feeding for the full sample, while Panel B, C, and D report ITT effects by gender, household poverty, and geographical areas, respectively[xii]. Increases in school enrolment emerge as an important plausible channel for impact, but only for children from poorest households and geographical areas. This finding is expected in contexts such as Ghana, where basic enrolment rates are already high and only poorest children are excluded from basic education. Treatment effects for attendance and grade attainment were positive across all groups, but only significant in the case of boys and non-poor children for grade attainment.

[TABLE 5 HERE]

5.2 Changes in cognition

Table 6 presents treatment effects on two indicators of child cognitive development: the standardized progressive matrices (SPM) and the digit span tests. The former is an adaptation of the Raven's progressive matrices test, a commonly-used measure of nonverbal fluid intelligence and problem-solving ability. For each question of the SPM test, the child was given a set of images, and was asked to choose the image that would complete the picture. The digit span test focuses on assessing working memory, whereby the child was presented sequences of numbers of increasing lengths, and was asked to recall the sequences as prompted (forwards) and reversing the number order (backwards). The same 12-item tests were administered across both rounds. Similarly to the learning outcomes, we generated a composite measure of child cognitive development. Also, although these indicators are also listed in the protocol as potentially affected by the intervention, we adjust p-values with the Romano-Wolf method.

School feeding had a positive effect on cognitive skills of the average child, with an increase of 0.12 s.d. in both the digit span and SPM scores, and of 0.14 s.d. in the composite score. Also, consistently with the results on learning, school feeding especially improved the cognitive development of disadvantaged learners groups. Specifically, the offer of school feeding led to an increase of 0.19 s.d., 0.27 s.d. and 0.25 s.d. in the digit span scores of treatment girls, children from poor households and northern Ghana, respectively, as compared to peers in control groups. School feeding also led to increases in the SPM score of more than 0.2 s.d. among children from most disadvantaged households and regions. The improvement in the composite cognitive score following the offer of school feeding accounted to 0.18 s.d. for girls, and slighly less than 0.3 s.d. for children from poor households and northern Ghana. There were also improvements in cognitive development among boys in the treatment arm: specifically, their SPM score improved by about 0.15 s.d., and their 0.12 s.d. in the composite index.

[TABLE 6 HERE]

5.3 Changes in nutritional status

A separate paper presents impact results on nutritional status (Aulo Gelli et al. 2019). However, given the potential relevance of this channel for learning, we report results. School feeding had no effect on the height-for-age z-scores (HAZ), a marker of chronic nutritional status, and on BMI-for-age z-scores (BAZ), an indicator of concurrent nutritional status, for the whole sample. However, the program had significant effects on HAZ of girls (effect size: 0.12 s.d.) and for young children in households living below the poverty line (effect size: 0.22 s.d.).

VI. Robustness checks

We investigate whether our results might have been driven by the fact that children may perform better in learning assessments after having eaten breakfast or the school lunch (Figlio and Winicki 2005). Although we did not record the time in which the tests were undertaken, we check if there are differences in whether children from the treatment and control groups had breakfast before schools (which would be relevant for those children that took the test in the morning), in the total number of meals consumed by the children, and in the overall diversity of the diet (a proxy for macro- and micro-nutrient dietary quality). The latter is measured as the number of food groups consumed by the child in the previous day (Ruel 2003), so that we can also investigate issues of quality of the diet beyond frequency of meal consumption. We do not find differences in any of these indicators by arm for the full sample or for the sub-groups (Online Appendix 10).

VII. Concluding Remarks

Most governments globally invest in providing food at school as a social protection strategy to enhance children's education and health, yet experimental evidence from LMICs is limited. Understanding whether large-scale, government-led school feeding is effective in raising human capital, and whether it enhances achievements for marginalized learner groups, is therefore a critical policy question and evidence gap in order to prioritize competing intervention options available to resource-constrained LMIC governments. We report treatment effects on learning from a trial assessing a program reaching daily two million children in Ghana. After two academic years of implementation, the offer of school feeding in randomized communities led to positive learning gains for the average pupil. The magnitude of the effects (about 0.15 s.d.) was comparable to estimates from a meta-analysis of school feeding programs implemented by NGOs or researchers in smaller-scale trials (Snilstveit et al. 2015), which is remarkable, as usually the latter tend to report larger treatment effects than larger programs (Vivalt 2019). Beyond average effects, the program had larger impacts among the groups that are more vulnerable to poor learning outcomes in Ghana. For girls, children from poor households and children residing in the country's northern regions, school feeding led to dramatic improvements in learning – ranging between 0.2 s.d. to 0.3 s.d.. All estimates are likely to represent lower bounds due to imperfect uptake, due to eligible children at baseline progressing to secondary (where the program is not served), being out of school or being enrolled in private education. Therefore, treatment-on-the-treated effects (which we did not estimate) would point to even larger effects.

To the best of our knowledge, this is the first study providing experimental evidence on a government's school feeding program implemented across all regions of a LMIC. A potential concern about the external validity of our results to the broader population of Ghana relates to our sampling frame. While the retargeting exercise allowed us to randomize the scale-up of the intervention, our sample draws from the country's poorest communities, which the government and donors decided to prioritize for the retargeting. However, there is a high degree of consistence between the poverty and extreme poverty headcounts (based on the national consumption expenditure poverty line) in our sample (23% and 7%, respectively) and the 2012-13 national headcounts (24% and 8%)(Ghana Statistical Service 2018). Although our purpose was not to assess the retargeting of the program, this finding highlights that the program is still not finely targeted: rather than being strongly pro-poor, our sample exhibits poverty levels comparable to nationally representative data. This, then, supports the view that the results have external validity. Importantly, even though the issues related to limited poverty targeting remain, our findings highlight that the relatively poorest in our sample benefit more from the program, which has implication in terms of potential increases in program impact if the GFSP was even more strategically targeted to more disadvantaged areas or households (World Bank 2012).

We note that we cannot experimentally distinguish between the relative contribution of changes in schooling, cognition and nutrition following the intervention. School feeding had a composition effect on educational participation by attracting children from poorest households and regions, which are at higher risk of exclusion from basic education in Ghana, to school. This result is consistent with other evidence from SSA, which showed that reductions in the cost of education are effective in raising enrolment, particularly for most educationally-vulnerable population segments (Aurino et al. 2018; Björkman-Nyqvist 2013; H. Kazianga, de Walque, and Alderman 2012). Further, the associated improvements in nutrition and child cognition, both inputs in the production function of human capital (Alderman et al. 1996), have likely enhanced the effects of school feeding on learning for those groups. On the other hand, our evidence is suggestive that for children from more advantaged backgrounds, which very likely were already in school, school feeding favored grade attainments for boys and non-poor children, and had led to some moderate improvements in cognition for boys.

The provision of a full cost-benefit/cost-effectiveness analysis of the GSFP is beyond the remit of this paper, especially given that the program is meant to affect a wide set of outcomes (education, health, social protection, agriculture), which should be assessed jointly to provide an accurate measure of cost-effectiveness (Aulo Gelli et al. 2014). Also, even if we would decide to restrict the focus on educational achievements, this exercise may be still unsatisfactory, as the life course and intergenerational effects of gains from increased human capital are not yet fully known. For instance, Bütikofer and co-authors have estimated that access to a school breakfast program run in the 1930s in Norway had positive long-term and intergenerational effects on education and earnings[xiii] (Bütikofer, Mølland, and Salvanes 2018). While we leave these important issues for future research, back-of-the-envelope calculations based on the government of Ghana's transfer to caterers and an average of 200 school-days per year suggest that the program costed about US$66 per child per year in 2015/16. While this is a very rough estimation as it does not include full implementation costs (e.g., other costs at the school-level that are not included in the government budget for school feeding), this figure falls within the range of the average cost per child of school meals in LMICs reported in Gelli and Darayani (2013) (see Introduction). Taking inflation into account,[xiv] the GSFP thus compares well with other programs in LMICs in terms of costs. Also, Gelli and Darayani's estimations of program costs were based on WFP operating costs. As the WFP is the largest school feeding implementer in the world and operates through a centralized model that allows economies of scale, its cost estimates likely provide a lower bound for government programs. This is especially relevant for countries seeking food procurement within national boundaries using "home-grown" approaches, such as Ghana, in order to stimulate internal agricultural production and rural poverty reduction, at the potential cost of raising programmatic budgets through the purchase of higher-cost, locally-grown staples.

Overall, our findings highlight the role of government-led, large-scale school feeding programs as a social protection tool with positive and equitable impacts on human capital accumulation, particularly for marginalized groups of learners. Program impacts are especially remarkable when contextualized to the normal implementation challenges related to large-scale programs run in LMICs. The implementation challenges we highlighted add to the generalizability of our findings to "real-world" interventions, which may face additional financial, implementation, and monitoring constraints than small-scale trials. Increasing average learning levels by narrowing the gaps in the distribution of achievements is critical for sustainable economic and social development. Therefore, government-led, large-scale school feeding programs remain an important educational and social protection tool for attaining the 2030 "learning for all" agenda.

References

Acemoglu, Daron. 2010. "Theory, General Equilibrium, and Political Economy in Development Economics." *Journal of Economic Perspectives* 24 (3): 17–32. https://doi.org/10.1257/jep.24.3.17.

Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2013. "School Meals and Classroom Effort: Evidence from India." *International Growth Centre.* Vol. 25. http://www.theigc.org/wp-content/uploads/2014/09/Afridi-Et-Al-2013-Working-Paper1.pdf.

———. 2019. "Hunger and Performance in the Classroom."

Ahmed, Akhter U. 2004. "Impact of Feeding Children in School: Evidence from Bangladesh." *International Food Policy Research Institute* 2033: 20006–1002. http://www.lcgbangladesh.org/FSN/reports/IFPRI Final Report_School Feeding in Bangladesh.pdf.

Akresh, Richard, D De Walque, and Harounan Kazianga. 2013. "Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality." *World Bank Policy Research*. Policy Research Working Papers. The World Bank. https://doi.org/10.1037/e547172013-001.

Alderman, Harold, Jere R. Behrman, David R. Ross, and Richard Sabot. 1996. "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labour Market." *Oxford Bulletin of Economics and Statistics*. https://doi.org/10.1111/j.1468-0084.1996.mp58001003.x.

Alderman, Harold, and Donald Bundy. 2012. "School Feeding Programs and Development: Are We Framing the Question Correctly?" *World Bank Research Observer* 27 (2): 204–21. https://doi.org/10.1093/wbro/lkr005.

Alderman, Harold, Ugo Gentilini, and Ruslan Yemtsov. 2018. *The 1.5 Billion People Question: Food, Vouchers, or Cash Transfers?* Edited by Harold Alderman, Ugo Gentilini, and Ruslan Yemtsov. Washington, D. C.: The World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/27907/9781464810879.pdf.

Alderman, Harold, Daniel O. Gilligan, and Kim Lehrer. 2012. "The Impact of Food for Education Programs on School Participation in Northern Uganda." *Economic Development and Cultural Change* 61 (1): 187–218. https://doi.org/10.1086/666949.

Allcott, Hunt, and Sendhil Mullainathan. 2012. "External Validity and Partner Selection Bias." *NBER Working Paper 18373*, 1–44. https://doi.org/10.1093/qje/qjv015.Advance.

Andang'o, Pauline EA, Saskia JM Osendarp, Rosemary Ayah, Clive E. West, David L. Mwaniki, Corine A. De Wolf, Rob Kraaijenhagen, Frans J. Kok, and Hans Verhoef. 2007. "Efficacy of Iron-Fortified Whole Maize Flour on Iron Status of Schoolchildren in Kenya: A Randomised Controlled Trial." *Lancet*. https://doi.org/10.1016/S0140-6736(07)60817-4.

Ash, Deborah M., Simon R. Tatala, Edward A. Frongillo, Godwin D. Ndossi, and Michael C. Latham. 2003. "Randomized Efficacy Trial of a Micronutrient-Fortified Beverage in Primary School Children in Tanzania." *The American Journal of Clinical Nutrition*.

Aurino, Elisabetta, Jean-Pierre Tranchant, Amadou Sekou Diallo, and Aulo Gelli. 2018. "School

Feeding or General Food Distribution? Quasi-Experimental Evidence on the Educational Impacts of Emergency Food Assistance during Conflict in Mali." Innocenti Working Paper. Florence. https://www.unicef-irc.org/publications/pdf/Aurino et al for WEB.pdf.

Baird, Sarah, Francisco H.G. Ferreira, Berk Özler, and Michael Woolcock. 2014. "Conditional, Unconditional and Everything in between: A Systematic Review of the Effects of Cash Transfer Programmes on Schooling Outcomes." *Journal of Development Effectiveness* 6 (1): 1–43. https://doi.org/10.1080/19439342.2014.890362.

Balwanz, David, and Peter Darvas. 2013. *Basic Education beyond the Millennium Development Goals in Ghana: How Equity in Service Delivery Affects Educational and Learning Outcomes*. The World Bank. https://doi.org/10.1596/978-1-4648-0098-6.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236). https://doi.org/10.1126/science.1260799.

Banerjee, Abhijit V., Shawn Cole, Esther; Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics* 122 (3): 1235–64. https://doi.org/10.1162/qjec.122.3.1235.

Bashir, Sajitha, Marlaine Lockheed, Elizabeth Ninan, and Jee-Peng Tan. 2018. *Facing Forward: Schooling for Learning in Africa*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1260-6.

Berry, James, S. Metha, P. Mukherjee, H. Ruebeck, and G. Kartini Shastry. 2018. "Inputs, Monitoring and Crowd-Out in India School-Based Health Interventions." S-89206-INC-2. London.

Björkman-Nyqvist, Martina. 2013. "Income Shocks and Gender Gaps in Education: Evidence from Uganda." *Journal of Development Economics* 105 (November): 237–53. https://doi.org/10.1016/j.jdeveco.2013.07.013.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2018. "Experimental Evidence on Scaling up Education Reforms in Kenya." *Journal of Public Economics* 168 (December): 1–20. https://doi.org/10.1016/J.JPUBECO.2018.08.007.

Bütikofer, Aline, Eirin Mølland, and Kjell G. Salvanes. 2018. "Childhood Nutrition and Labor Market Outcomes: Evidence from a School Breakfast Program." *Journal of Public Economics* 168 (December): 62–80. https://doi.org/10.1016/j.jpubeco.2018.08.008.

Chakraborty, Tanika, and Rajshri Jayaraman. 2019. "School Feeding and Learning Achievement: Evidence from India's Midday Meal Program." *Journal of Development Economics*, April. https://doi.org/10.1016/J.JDEVECO.2018.10.011.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Score." *American Economic Journal: Applied Economics* 5 (2): 29–57. https://doi.org/10.1257/app.5.2.29.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55. https://doi.org/10.1257/jel.48.2.424.

Drake, Lesley, Meena Fernandes, Elisabetta Aurino, Josephine Kiamba, Boitshepo Giyose, Carmen Burbano, Harold Alderman, Lu Mai, Arlene Mitchell, and Aulo Gelli. 2017. "School Feeding Programs in Middle Childhood and Adolescence." In *Disease Control Priorities 3*, edited by D. Bundy, N. De Silva, S. Horton, D. Jamison, and G.C. Patton. Washington, D.C.: The World Bank. http://dcp-3.org/sites/default/files/chapters/DCP3

CAHD_Ch 12.pdf.

Duflo, Esther;, Pascaline Dupas, and Michael Kremer. 2017. “The Impact of Free Secondary Education: Experimental Evidence from Ghana.” https://www.poverty-action.org/sites/default/files/publications/DDK_GhanaScholarships.pdf.

Evans, D. K., and F. Yuan. 2018. “Commentary.” In *Learning at the Bottom of the Pyramid.*, edited by Daniel A. Wagner, Sharon Wolf, and Robert F. Boruch, 232–34. Paris: UNESCO.

Fernandes, Meenakshi, Gloria Folson, Elisabetta Aurino, and Aulo Gelli. 2017. “A Free Lunch or a Walk Back Home? The School Food Environment and Dietary Behaviours among Children and Adolescents in Ghana.” *Food Security*, October 1, 2017. https://doi.org/10.1007/s12571-017-0712-0.

Figlio, David N., and Joshua Winicki. 2005. “Food for Thought: The Effects of School Accountability Plans on School Nutrition.” *Journal of Public Economics* 89 (2–3): 381–94. https://doi.org/10.1016/J.JPUBECO.2003.10.007.

Filmer, Deon P., Jed Friedman, Eeshani Kandpal, and Junko Onishi. 2018. “General Equilibrium Effects of Targeted Cash Transfers: Nutrition Impacts on Non-Beneficiary Children.” 8377. *Policy Research Working Paper*. Vol. 8377. Washington, D.C. http://documents.worldbank.org/curated/en/989031522077749796/General-equilibrium-effects-of-targeted-cash-transfers-nutrition-impacts-on-non-beneficiary-children.

Frison, Lars, and Stuart J Pocock. 1992. “Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and Its Implications for Design.” *Statistics in Medicine* 11 (13): 1685–1704. https://doi.org/10.1002/sim.4780111304.

Gelli, A., E. Masset, G. Folson, A. Kusi, D.K. Arhinful, F. Asante, I. Ayi, et al. 2016. “Evaluation of Alternative School Feeding Models on Nutrition, Education, Agriculture and

Other Social Outcomes in Ghana: Rationale, Randomised Design and Baseline Data." *Trials* 17 (1). https://doi.org/10.1186/s13063-015-1116-0.

Gelli, Aulo, Elisabetta Aurino, Gloria Folson, Daniel Arhinful, Clement Adamba, Isaac Osei-Akoto, Edoardo Masset, et al. 2019. "A School Meals Program Implemented at Scale in Ghana Increases Height-for-Age during Midchildhood in Girls and in Children from Poor Households: A Cluster Randomized Trial." *The Journal of Nutrition*, May. https://doi.org/10.1093/jn/nxz079.

Gelli, Aulo, and Roshan Daryanani. 2013. "Are School Feeding Programs in Low-Income Settings Sustainable? Insights on the Costs of School Feeding Compared with Investments in Primary Education." *Food and Nutrition Bulletin* 34 (3): 310–17.

Gelli, Aulo, Francisco Espejo, Jing Shen, and Elizabeth Kristjansson. 2014. "Putting It All Together: Aggregating Impacts of School-Feeding Programmes on Education, Health and Nutrition: Two Proposed Methodologies." *WIDER Working Paper Series*. https://ideas.repec.org/p/unu/wpaper/wp2014-036.html.

Ghana Statistical Service. 2018. "Ghana Living Standards Survey (GLSS7): Poverty Trends in Ghana; 2005-2017." *Ghana Statistical Service*.

Glewwe, Paul, and Michael Kremer. 2006. "Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries." *Handbook of the Economics of Education* 2: 945–1017. https://doi.org/10.1016/S1574-0692(06)02016-2.

Graff Zivin, Joshua, Solomon M. Hsiang, and Matthew Neidell. 2018. "Temperature and Human Capital in the Short and Long Run." *Journal of the Association of Environmental and Resource Economists* 5 (1): 77–105. https://doi.org/10.1086/694177.

Grantham-McGregor, S M, S Chang, and S P Walker. 1998. "Evaluation of School Feeding

Programs: Some Jamaican Examples." *The American Journal of Clinical Nutrition* 67 (4): 785S-789S. http://www.ncbi.nlm.nih.gov/pubmed/9537629.

Groot, Richard De, Sudhanshu Handa, Michael Park, Robert Osei Darko, Isaac Osei-Akoto, Garima Bhalla, and Luigi Peter Ragno. 2015. "Heterogeneous Impacts of an Unconditional CashTransfer Programme on Schooling: Evidence from the Ghana LEAP Programme." Working Paper.

GSFP. 2006. "Ghana School Feeding Programme. Programme Document 2007-2010." Accra.

Hayes, Richard J., and Lawrence H. Moulton. 2017. *Cluster Randomised Trials*. 2nd ed. Chapman and Hall/CRC.

Hulett, Judie L., Robert E. Weiss, Nimrod O. Bwibo, Osman M. Galal, Natalie Drorbaugh, and Charlotte G. Neumann. 2014. "Animal Source Foods Have a Positive Impact on the Primary School Test Scores of Kenyan Schoolchildren in a Cluster-Randomised, Controlled Feeding Intervention Trial." *British Journal of Nutrition* 111 (05): 875–86. https://doi.org/10.1017/S0007114513003310.

Jacoby, Hanan G. 2002. "Is There an Intrahousehold ' Flypaper Effect '? Evidence from a School Feeding Programme." *Economic Journal* 112 (476): 196–221.

Jukes, M C H, L J Drake, and D A P Bundy. 2008. *School Health, Nutrition and Education for All: Levelling the Playing Field*. Wallingford: CABI Publishing. https://doi.org/http://dx.doi.org/10.1079/9781845933111.0011.

Kazianga, H., D. de Walque, and H. Alderman. 2012. "Educational and Child Labour Impacts of Two Food-for-Education Schemes: Evidence from a Randomised Trial in Rural Burkina Faso." *Journal of African Economies* 21 (5): 723–60. https://doi.org/10.1093/jae/ejs010.

Kazianga, Harounan, Damien de Walque, and Harlod Alderman. 2009. "Educational and Health

Impacts of Two School Feeding Schemes Evidence from a Randomized Trial in Rural Burkina Faso." 4976. World Bank Policy Research Working Paper.

Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2014. "School Feeding Programs, Intrahousehold Allocation and the Nutrition of Siblings: Evidence from a Randomized Trial in Rural Burkina Faso." *Journal of Development Economics* 106 (January): 15–34. https://doi.org/10.1016/j.jdeveco.2013.08.007.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2006. "Experimental Analysis of Neighborhood Effects." *Econometrica*, no. June: 1–34. https://doi.org/10.1111/j.1468-0262.2007.00733.x.

Krämer, Marion; Kumar, Santosh; Vollmer, Sebastian. 2018. "Improving Children Health and Cognition: Evidence from School-Based Nutrition Intervention in India." 203. GLO Discussion Paper. Maastricht.

Kristjansson, Elizabeth, Damian K Francis, Selma Liberato, Maria Benkhalti Jandu, Vivian Welch, Malek Batal, Trish Greenhalgh, et al. 2015. "Food Supplementation for Improving the Physical and Psychosocial Health of Socio-Economically Disadvantaged Children Aged Three Months to Five Years (Review)." *Cochrane Database of Systematic Reviews*, no. 3. https://doi.org/10.1002/14651858.CD009924.pub2.

Lawson, Ty M. 2012. "Impact of School Feeding Programs on Educational, Nutritional, and Agricultural Development Goals: A Systematic Review of Literature." Michigan State University. https://www.3ieimpact.org/evidence-hub/systematic-review-repository/impact-school-feeding-programs-educational-nutritional.

McKenzie, David. 2012. "Beyond Baseline and Follow-up: The Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210–21.

https://doi.org/10.1016/j.jdeveco.2012.01.002.

Ministry of Education & RTI International. 2014. "Ghana 2013 National Education Assessment - Technical Report | Global Reading Network." https://globalreadingnetwork.net/eddata/ghana-2013-national-education-assessment-technical-report.

Moretti, Diego, Michael B. Zimmermann, Sumithra Muthayya, Prashanth Thankachan, Tung Ching Lee, Anura V. Kurpad, and Richard F. Hurrell. 2006. "Extruded Rice Fortified with Micronized Ground Ferric Pyrophosphate Reduces Iron Deficiency in Indian Schoolchildren: A Double-Blind Randomized Controlled Trial1-3." *American Journal of Clinical Nutrition*. https://doi.org/10.1093/ajcn/84.4.822.

Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at Scale." *Journal of Economic Perspectives* 31 (4): 103–24. https://doi.org/10.1257/jep.31.4.103.

Muthayya, Sumithra, Tinku Thomas, Krishnamachari Srinivasan, Kirthi Rao, Anura V. Kurpad, Jan Willem van Klinken, Gail Owen, and Eveline A. de Bruin. 2007. "Consumption of a Mid-Morning Snack Improves Memory but Not Attention in School Children." *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2006.09.025.

Neumann, Charlotte G, Suzanne P Murphy, Connie Gewa, Monika Grillenberger, and Nimrod O Bwibo. 2007. "Meat Supplementation Improves Growth, Cognitive, and Behavioral Outcomes in Kenyan Children." *The Journal of Nutrition* 137 (4): 1119–23. http://www.ncbi.nlm.nih.gov/pubmed/17374691.

Osendarp, Saskia J.M., Katrine I. Baghurst, Janet Bryan, Eva Calvaresi, Donna Hughes, Mahdin Hussaini, Elvina Karyadi, et al. 2007. "Effect of a 12-Mo Micronutrient Intervention on Learning and Memory in Well-Nourished and Marginally Nourished School-Aged

Children: 2 Parallel, Randomized, Placebo-Controlled Studies in Australia and Indonesia." *American Journal of Clinical Nutrition*.

Powell, C A, S P Walker, S M Chang, and S M Grantham-McGregor. 1998a. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." *The American Journal of Clinical Nutrition* 68 (4): 873–79.

———. 1998b. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." *The American Journal of Clinical Nutrition* 68 (4): 873–79. http://www.ncbi.nlm.nih.gov/pubmed/9771865.

Radhika, Madhari S., Krishnapillai M. Nair, Rachakulla Hari Kumar, Mendu Vishnuvardhana Rao, Punjal Ravinder, Chitty Gal Reddy, and Ginnela N.V. Brahmam. 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*. https://doi.org/10.3945/ajcn.110.007179.

Romano, Joseph P., and Michael Wolf. 2005. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica*. https://doi.org/10.1111/j.1468-0262.2005.00615.x.

———. 2016. "Efficient Computation of Adjusted P-Values for Resampling-Based Stepdown Multiple Testing." *Statistics and Probability Letters*. https://doi.org/10.1016/j.spl.2016.02.012.

Ruel, M.T. 2003. "Operationalizing Dietary Diversity: A Review of Measurement Issues and Research Priorities." *Journal of Nutrition* 133: 3911S-3926S.

SEND-Ghana. 2013. "Ghana School Feeding Programme: Way Forward."

Snilstveit, Birte, Jennifer Stevenson, Daniel Phillips, Martina Vojtkova, Emma Gallagher, Tanja

Schmidt, Hannah Jobse, Maisie Geelen, Maria Grazia Pastorello, and John Eyers. 2015. "Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review." London.

Stuijvenberg, M E van, J D Kvalsvig, M Faber, M Kruger, D G Kenoyer, and a J Benadé. 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): 497–503.

UNDP Ghana. 2015. "2015 Ghana Millennium Development Goals Report." Accra. http://www.gh.undp.org/content/ghana/en/home/library/poverty/2015-ghana-millennium-development-goals-report.html.

Vermeersch, Christel, and Michael Kremer. 2005. "School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation." Policy Research Working Paper 3523. Washington, D. C. https://doi.org/10.2139/ssrn.667881.

Vivalt, Eva. 2015. "Heterogeneous Treatment Effects in Impact Evaluation." *American Economic Review* 105 (5): 467–70. https://doi.org/10.1257/aer.p20151015.

———. 2019. "How Much Can We Generalize from Impact Evaluations ?" *Forthcoming, Journal European Economic Association*.

WFP. 2013. *State of School Feeding Worldwide 2013*. Rome: World Food Programme.

Whaley, Shannon E., Marian Sigman, Charlotte Neumann, Nimrod Bwibo, Donald Guthrie, Robert E. Weiss, Susan Alber, and Suzanne P. Murphy. 2003. "The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children." *The Journal of Nutrition*. https://doi.org/10.1093/jn/133.11.3965s.

World Bank. 2012. *Improving the Targeting of Social Programs in Ghana*. Edited by Quentin

Wodon. *World Bank*. https://doi.org/10.1596/978-0-8213-9593-6.

———. 2017. “World DataBank: Education Statistics - All Indicators.” 2017. http://databank.worldbank.org/data/reports.aspx?source=education-statistics-~-all-indicators#.

———. 2018. “World Development Report 2018: Learning to Realize Education’s Promise.” Washington, D.C. https://doi.org/10.1596/978-1-4648-1096-1.

Zimmermann, Michael B., Christophe Zeder, Noureddine Chaouki, Amina Saad, Toni Torresani, and Richard F. Hurrell. 2003. “Dual Fortification of Salt with Iodine and Microencapsulated Iron: A Randomized, Double-Blind, Controlled Trial in Moroccan Schoolchildren.” *American Journal of Clinical Nutrition*. https://doi.org/10.1093/ajcn/77.2.425.

Figures

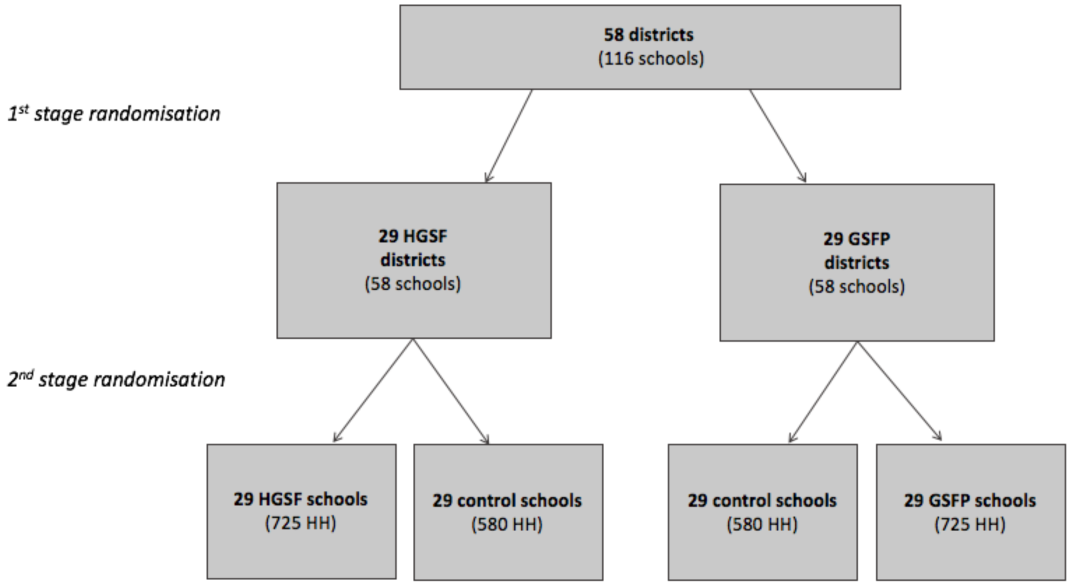


Figure 1.

Two-level randomization

Notes: this figure provides the study design: first, districts were randomly assigned to pilot (HGSF) and standard (GSFP) school feeding through a first-level randomization; second, within each district, two schools (and related households living within the school catchment areas, which were refer to as “communities”) were randomly assigned to school feeding or control. Note that due to the discovery of the GSFP already present in 25 communities at baseline, these were dropped from the original community sample. Two additional communities could not be resurveyed due to local violence at the time of the endline survey. An original third level of randomization was dropped soon after the baseline due to substantial delays in implementation and the limited number of schools still available after the removal of schools that we discovered had school feeding at baseline.

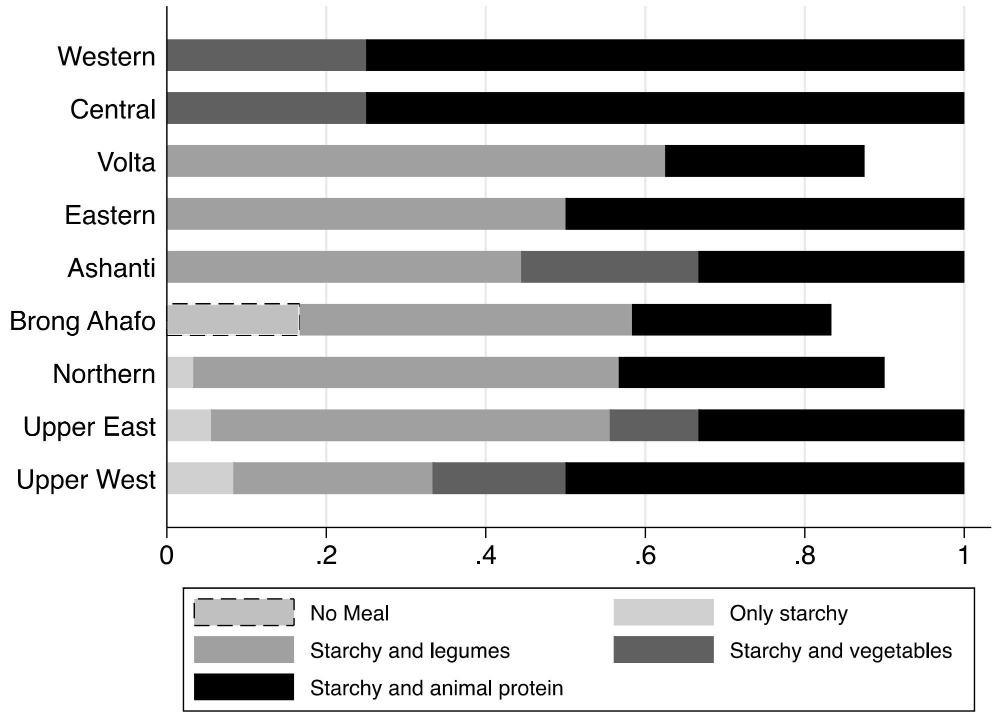
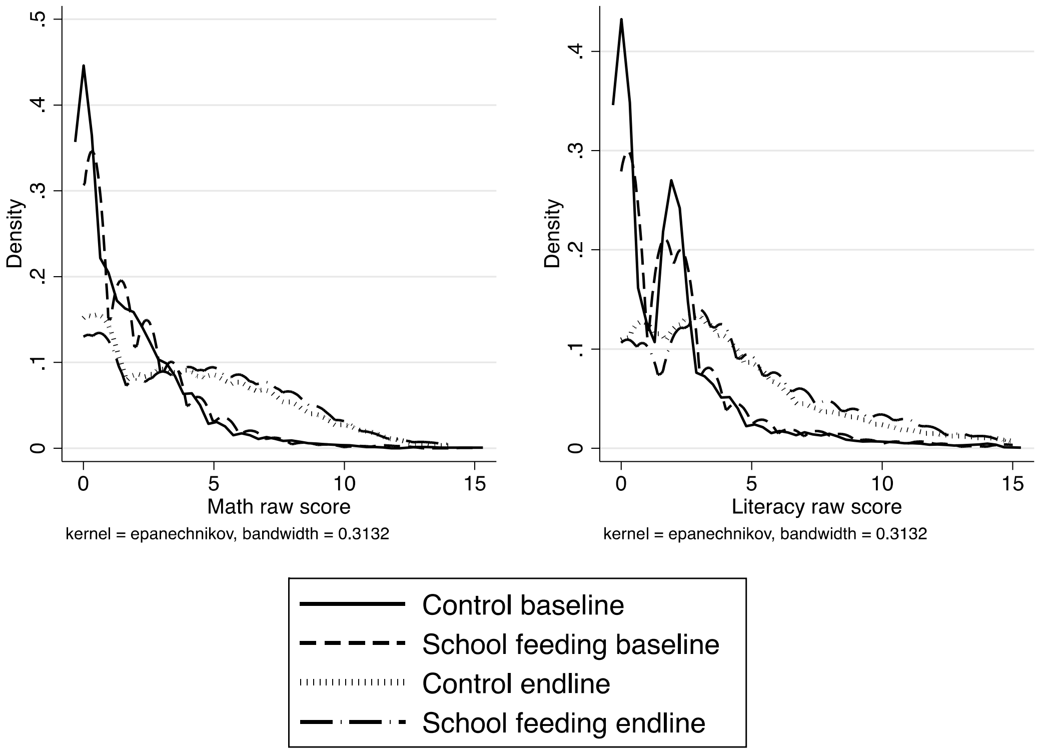


Figure 2.

Meal content by region

Notes: this figure presents the proportion of specific meal types served in the day of the interview and in the following day by region. Data on the meal served were taken by the caterer's weekly meal logs.

Panel A. Raw test scores



Panel B. Age-standardized scores

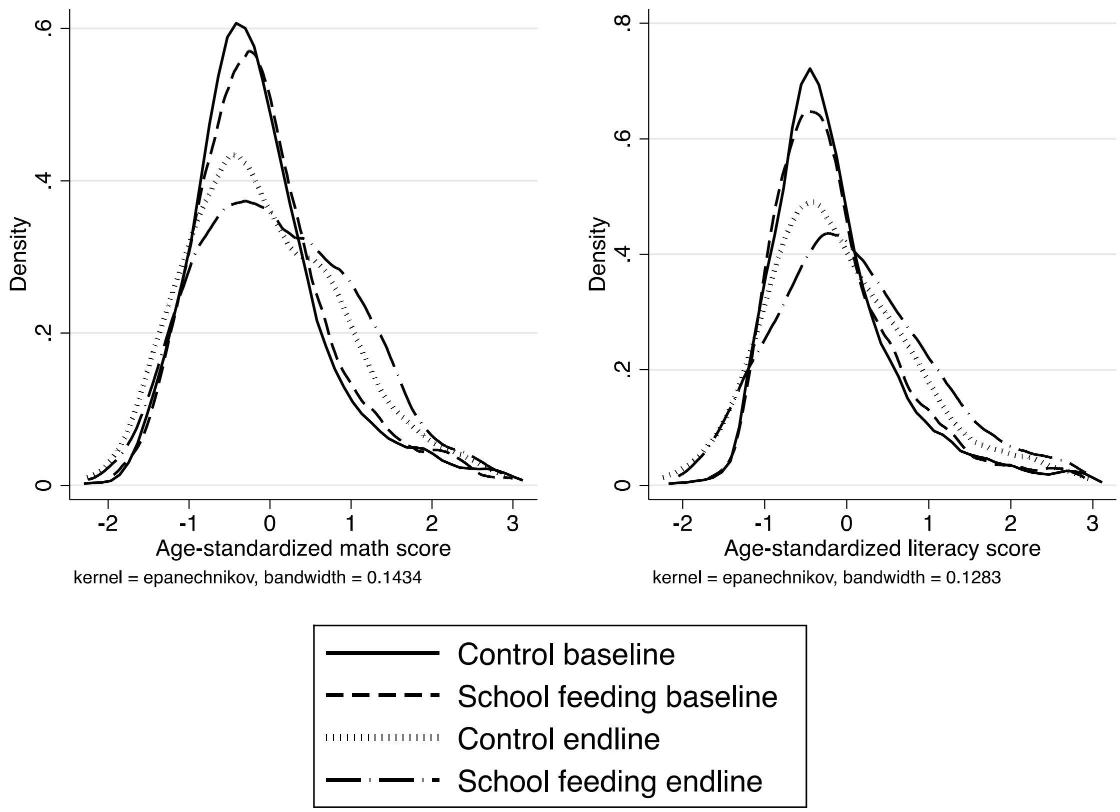
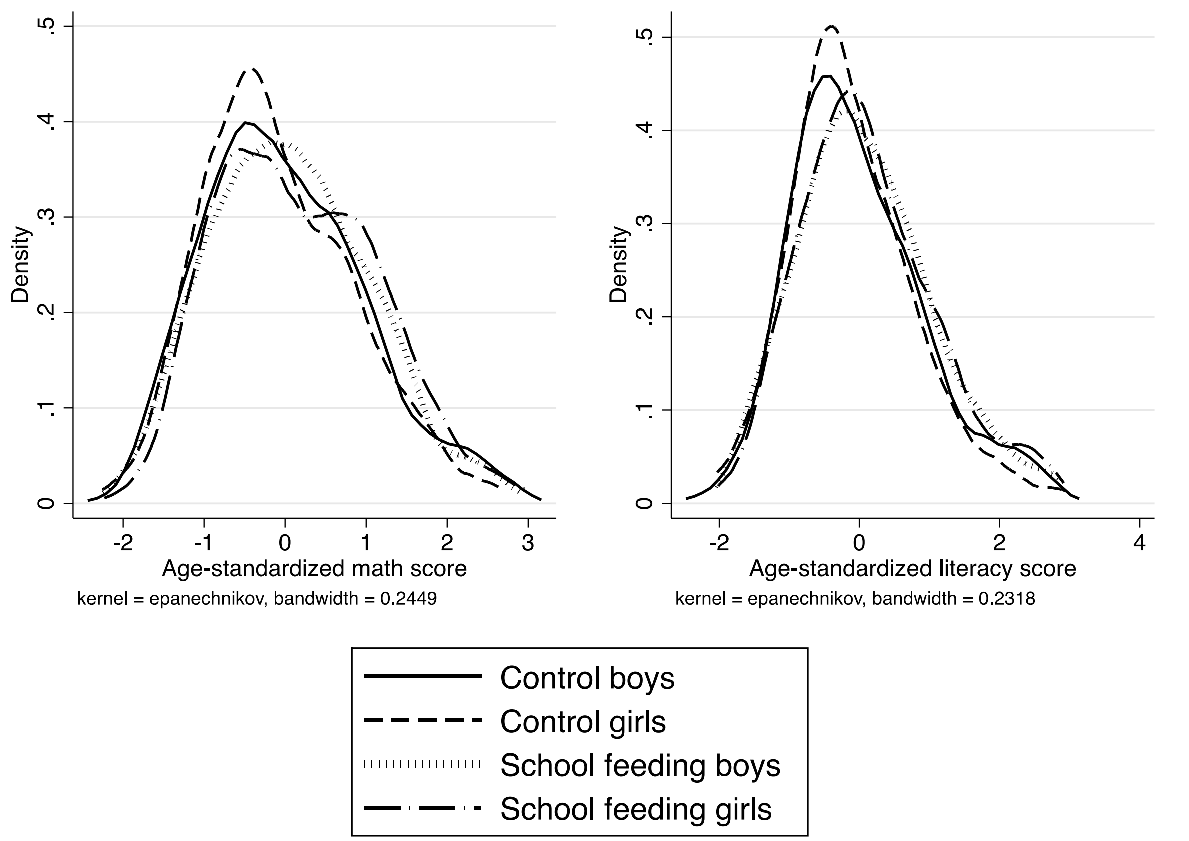


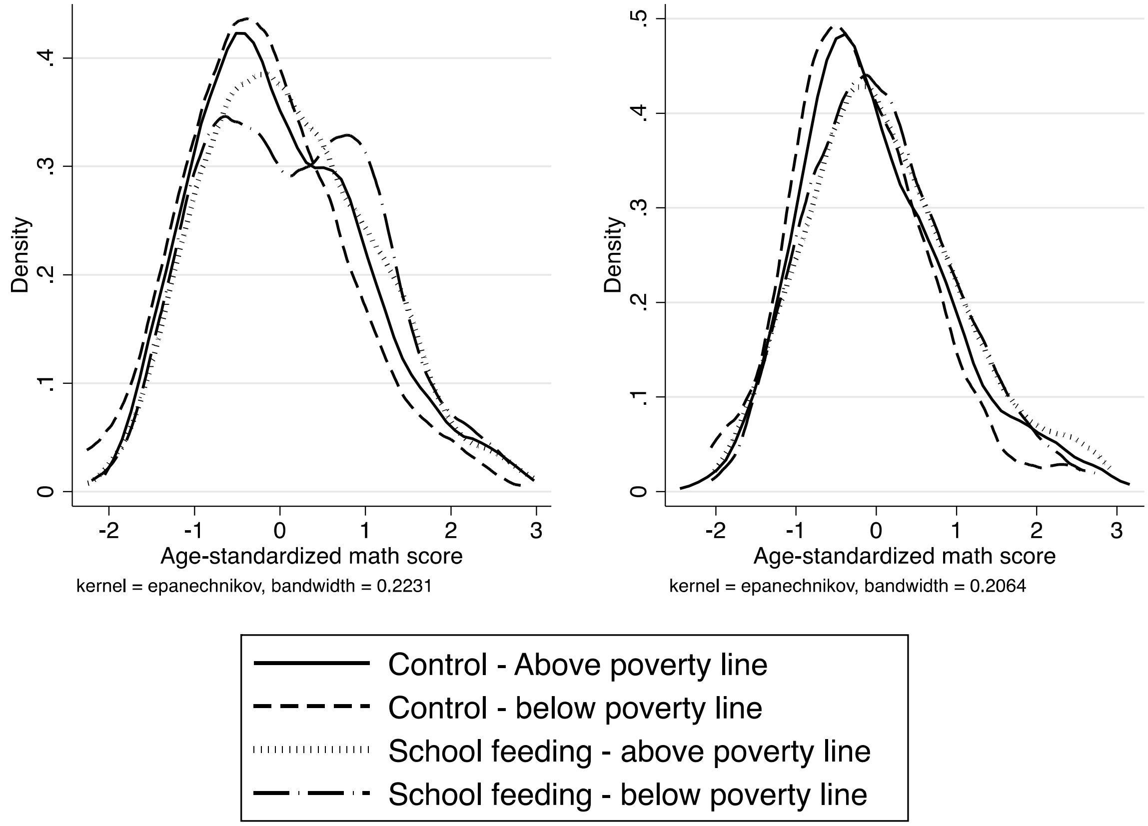
Figure 3.
Empirical distributions of test scores for maths (left-side panel) and literacy (right-side panel), by survey round and treatment arm

Notes: this figure presents, by treatment group and survey wave, the non-parametric distributions of math (left-side panel) and literacy (right-side panel) scores for the full longitudinal sample of children. Panel A and B show raw and age-standardized test scores, respectively. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel.

Panel A. Child Gender



Panel B. Household poverty



Panel C. Geographical region

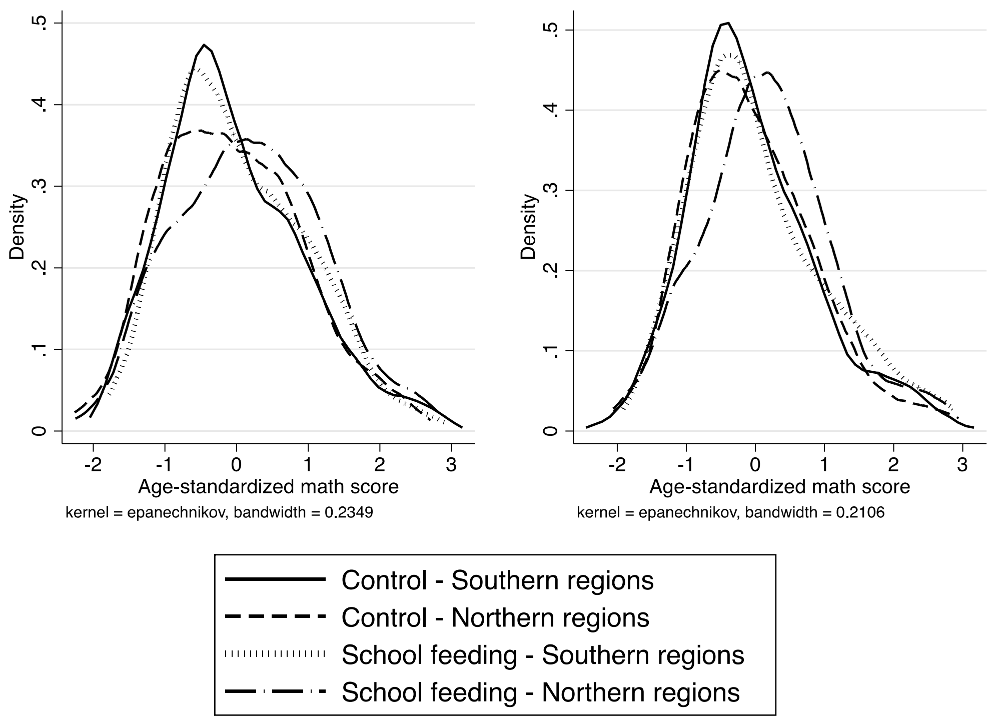


Figure 4.
Empirical endline distributions of age-standardised maths (left-side panel) and literacy (right-side panel) age-standardised test scores, by treatment arm and child gender, poverty and geographical region

Notes: these figures present endline age-standardized scores by treatment and gender (Panel A), household poverty status at baseline (Panel B) and geographical region (Panel C). Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel. Household poverty is a dichotomous indicator having the value of one if the household had baseline per capita consumption levels falling below the national consumption poverty line in 2013. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Tables

Table 1.

Descriptive statistics and balance of covariates at baseline, full baseline sample

	Control (N=1,612)	School feeding (N=1,821)	School feeding - control difference (SE)
Child age in months	102.73	103.92	1.185
	(30.77)	(31.45)	(1.485)
Child is male	0.54	0.51	-0.027
	(0.50)	(0.50)	(0.018)
Enrolled	0.99	0.98	-0.007
	(0.11)	(0.14)	(0.006)
Child has child fallen ill in last 7 days	0.10	0.09	-0.010
	(0.31)	(0.29)	(0.013)
Highest grade completed	1.67	1.76	0.088
	(1.53)	(1.48)	(0.083)
Child has repeated a grade	0.11	0.12	0.011
	(0.32)	(0.33)	(0.020)
Absent from school in past 7 days	0.12	0.17	0.050
	(0.66)	(0.76)	(0.041)
Private school	0.10	0.11	0.010
	(0.30)	(0.31)	(0.033)
Height-for-age Z-scores	-1.11	-1.05	0.062
	(1.35)	(1.29)	(0.088)
Number of children of target age	3.38	3.24	-0.142
	(1.69)	(1.71)	(0.184)
Number of children under 5 years	1.06	0.94	-0.117
	(0.94)	(0.96)	(0.092)
Household size	6.77	6.60	-0.178
	(2.72)	(2.67)	(0.313)
Head of the household is male	0.81	0.80	-0.004
	(0.39)	(0.40)	(0.040)
Head of the household’s age	44.06	45.52	1.477*
	(12.05)	(12.69)	(0.742)
Mother's age	37.45	38.58	1.128
	(10.83)	(10.95)	(0.740)
Mother's education in years	5.22	6.01	0.789
	(5.01)	(4.17)	(0.662)
Wealth index	13.21	13.38	0.174
	(11.45)	(11.77)	(1.541)
Sold agriculture produce in the past year	0.51	0.43	-0.074
	(0.50)	(0.50)	(0.055)
Per capita expenditure	2,085.17	2092.62	7.446
	(993.87)	(1,097.27)	(109.181)
Household owns livestock	0.68	0.66	-0.012
	(0.47)	(0.47)	(0.048)
Urban	0.06	0.06	0.001
	(0.24)	(0.24)	(0.039)
Northern regions	0.43	0.50	0.066
	(0.50)	(0.50)	(0.110)

Notes: * p<0.1. N= 3,433. This table presents descriptive statistics for the full baseline sample of eligible children at baseline, stratified by assignment to treatment. The sample refers to all children aged 5-15 interviewed at baseline, prior to attrition. Mean and standard deviation in parentheses. The school feeding-control difference column reports the school feeding coefficient of a basic OLS regression with each covariate as an outcome and standard errors clustered at the community level. For each variable, the estimated school feeding coefficient provides the difference between the school feeding and control groups and its standard errors.

Table 2.

Baseline correlates of children remaining in the longitudinal sample

	(1)	(2)	(3)
Treatment	-0.016	-0.016	-0.016
	(0.018)	(0.019)	(0.028)
Age-standardized maths score		-0.000	
		(0.010)	
Age-standardized literacy score		0.011	
		(0.010)	
Male			0.019*
			(0.012)
Male*Treatment			-0.013
			(0.018)
Northern regions			0.050***
			(0.017)
Northern regions* Treatment			-0.020
			(0.035)
Poor Household			-0.035
			(0.027)
Poor*Treatment			0.061*
			(0.036)
Constant	0.932***	0.933***	0.909***
	(0.010)	(0.010)	(0.014)
Observations	3,433	3,158	3,432
R-squared	0.001	0.002	0.009
Prob > F	0.388	0.400	0.012

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents probability of remaining in the longitudinal sample estimated through linear probability models, with standard errors clustered at the community level. N= 3,433 children of target-age prior to attrition. Lower sample sizes reflect covariates that are missing or not applicable. Column 1 shows probabilities of child being followed-up by treatment assignment; column 2 presents odd ratios by baseline learning and cognition, while column 3 interacts randomized assignment with key variables by which heterogeneity analysis was conducted throughout the paper. Household poverty is a dichotomous indicator having the value of one if the household had baseline per capita consumption levels falling below the national consumption poverty line in 2013. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Table 3.

Descriptive statistics and test of balance of raw and age-standardized test scores, by survey round and treatment arm, longitudinal sample

	Baseline			Endline	
	Control (N=1,404)	School feeding (N=1,579)	Treatment-control difference (SE)[a]	Control (N=1,186)	School feeding (N=1,343)
Math (raw score)	1.57	1.68	0.073	3.62	4.03
	(2.04)	(2.07)	(0.146)	(3.37)	(3.43)
Math (age-standardized score)	-0.13	-0.09	0.051	-0.04	0.11
	(0.82)	(0.83)	(0.067)	(0.96)	(0.98)
Literacy (raw score)	1.81	1.97	0.106	3.87	4.33
	(2.43)	(2.56)	(0.211)	(3.51)	(3.57)
Literacy (age-standardized score)	-0.15	-0.12	0.034	-0.06	0.11
	(0.77)	(0.78)	(0.067)	(0.92)	(0.97)

[a] The school feeding-control difference column reports the school feeding coefficient of a basic OLS regression of each outcome over school feeding arm and controlling for child age in months. Standard errors are clustered at the community level. Lower sample sizes in the cognitive scores (as compared to the full longitudinal sample) reflect missing values in those scores.

Table 4.

Treatment effects: Full sample and heterogeneity by child gender, household poverty, and geographical areas

	Math	**Literacy**	**Composite learning score**
Panel A. All children			
School feeding	0.147	0.132	0.168
	[-0.003 - 0.298]	[-0.012 - 0.277]	[0.002 - 0.333]
Unadjusted p-value	0.055*	0.072*	0.047**
R-W p-value	0.098*	0.098*	
Observations	2,278	2,274	2,314
R-squared	0.068	0.130	0.139
Mean Treatment Endline	0.112	0.105	0.194
Mean Control Endline	-0.0435	-0.0562	0

	Math		**Literacy**		**Composite learning score**	
Panel B: Gender	Girls	Boys	Girls	Boys	Girls	Boys
School feeding	0.242	0.0645	0.205	0.0758	0.273	0.0797
	[0.0384 - 0.445]	[-0.148 - 0.277]	[0.0451 - 0.365]	[-0.195 - 0.347]	[0.0783 - 0.467]	[-0.185 - 0.345]
Unadjusted p-value	0.024**	0.509	0.018**	0.543	0.011**	0.513
R-W p-value	0.008***	0.573	0.018**	0.573		
Observations	1,071	1,207	1,067	1,207	1,085	1,229
R-squared	0.089	0.062	0.137	0.135	0.150	0.144
Mean Treatment Endline	0.158	0.0700	0.120	0.0923	0.237	0.156
Mean Control Endline	-0.0878	-0.00534	-0.102	-0.0170	-0.0509	0.0440

Table 4 (ctd.)

	Math		Literacy		Composite learning score	
Panel C: Household Poverty	Below the poverty line	Non-poor	Below the poverty line	Non-poor	Below the poverty line	Non-poor
School feeding	0.309 [0.163 - 0.454]	0.101 [-0.0567 - 0.260]	0.233 [0.0163 - 0.449]	0.0883 [-0.150 - 0.327]	0.328 [0.142 - 0.515]	0.114 [-0.114 - 0.341]
Unadjusted p-value	0.001***	0.181	0.038**	0.424	0.004***	0.287
R-W p-value	0.008***	0.410	0.021**	0.410		
Observations	539	1,739	537	1,737	542	1,772
R-squared	0.090	0.071	0.089	0.151	0.123	0.151
Mean Treatment Endline	0.136	0.104	0.0704	0.116	0.177	0.199
Mean Control Endline	0.186	5.20e-05	0.203	-0.0115	-0.177	0.0539

	Math		Literacy		Composite learning score	
Panel D: Place of residence	North	South	North	South	North	South
School feeding	0.253 [0.0244 - 0.482]	0.0508 [-0.186 - 0.288]	0.243 [-0.311 - 0.797]	0.0308 [-0.166 - 0.228]	0.297 [-0.118 - 0.711]	0.0496 [-0.174 - 0.273]
Unadjusted p-value	0.041**	0.619	0.200	0.715	0.091*	0.607
R-W p-value	0.069*	0.803	0.069*	0.803		
Observations	1,083	1,195	1,087	1,187	1,096	1,218
R-squared	0.043	0.099	0.098	0.163	0.090	0.193
Mean Treatment Endline	0.179	0.0418	0.169	0.0382	0.270	0.115
Mean Control Endline	-0.0639	-0.0273	-0.0629	-0.0509	0.0160	0.0127

Notes: *** p<0.01, ** p<0.05, * p<0.1. Confidence intervals clustered at community levels in squared brackets. R-W p-values were adjusted for multiple testing using the Romano-Wolf (2005, 2016) step-down method with 2,000 iterations and standard errors clustered at community level.

The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and place of residence. Models were estimated through OLS. For each outcome, the model controls for its baseline value, a dichotomous variable related to the randomized assignment to school feeding, and region dummies. Math and literacy scores are age-standardized. The composite index of learning was computed as the average of the math and literacy scores and then they were standardized to the control group within each round. Household poverty is a dichotomous indicator having the value of one if the household had baseline per capita consumption levels falling below the national consumption poverty line in 2013. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Table 5.

Treatment effects of school feeding on schooling, full sample, and heterogeneity by child gender, household poverty, and geographical areas

	Enrolment	Attendance	Grade
Panel A. All children			
School feeding	0.027	0.044	0.145
	[-0.015 - 0.069]	[-0.044 - 0.132]	[-0.009 - 0.298]
Unadjusted p-value	0.207	0.324	0.065*
R-W p-value	0.372	0.372	0.168
Observations	2,371	2,109	2,254
R-squared	0.030	0.038	0.671
Mean Treatment Endline	0.932	4.685	4.506
Mean Control Endline	0.884	4.665	4.269

	Enrolment		Attendance		Grade	
Panel B: Gender	Girls	Boys	Girls	Boys	Girls	Boys
School feeding	0.042 [-0.0034 - 0.087]	0.015 [-0.060 - 0.089]	0.041 [-0.081 - 0.162]	0.051 [-0.065 - 0.166]	0.031 [-0.161 - 0.223]	0.238 [0.070 - 0.406]
Unadjusted p-value	0.07*	0.563	0.507	0.384	0.750	0.006***
R-W p-value	0.192	0.614	0.753	0.614	0.753	0.015**
Observations	1,097	1,274	988	1,121	1,056	1,198
R-squared	0.043	0.026	0.051	0.035	0.652	0.692
Mean Treatment Endline	0.951	0.915	4.662	4.706	4.487	4.524
Mean Control Endline	0.885	0.882	4.643	4.684	4.313	4.230

	Enrolment		Attendance		Grade	
Panel C: Household Poverty	Below the poverty line	Non-poor	Below the poverty line	Non-poor	Below the poverty line	Non-poor
School feeding	0.053 [0.003 - 0.104]	0.020 [-0.028 - 0.0682]	0.019 [-0.176 - 0.214]	0.056 [-0.034 - 0.145]	0.096 [-0.145 - 0.337]	0.170** [0.0123- 0.328]
Unadjusted p-value	0.039**	0.401	0.847	0.221	0.429	0.035**
R-W p-value	0.097*	0.409	0.841	0.384	0.646	0.093*
Observations	551	1,820	489	1,620	524	1,730
R-squared	0.025	0.040	0.076	0.046	0.666	0.676
Mean Treatment Endline	0.952	0.926	4.589	4.715	4.679	4.452
Mean Control	0.861	0.891	4.643	4.671	3.975	4.355

Endline

Table 5 (ctd.)

	Enrolment		Attendance		Grade	
Panel D: Place of residence	North	South	North	South	North	South
School feeding	0.076 [0.016 - 0.135]	-0.013 [-0.067 - 0.0405]	0.072 [-0.026 - 0.171]	0.020 [-0.122 - 0.162]	0.223 [-0.042 - 0.489]	0.0781 [-0.102 - 0.258]
Unadjusted p-value	0.014**	0.624	0.146	0.778	0.097*	0.387
R-W p-value	0.033**	0.864	0.170	0.864	0.170	0.778
Observations	1,092	1,279	992	1,117	1,030	1,224
R-squared	0.051	0.026	0.005	0.034	0.639	0.695
Mean Treatment Endline	0.932	0.932	4.820	4.545	4.352	4.667
Mean Control Endline	0.827	0.927	4.780	4.586	4.068	4.406

Notes: *** p<0.01, ** p<0.05, * p<0.1. Confidence intervals clustered at community levels in squared brackets. R-W p-values were adjusted for multiple testing using the Romano-Wolf (2005, 2016) step-down method with 2,000 iterations and standard errors clustered at community level.

The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and place of residence. Models were estimated through OLS. For each outcome, the model controls for its baseline value, a dichotomous variable related to the randomized assignment to school feeding, and region dummies. Enrolment is a dichotomous variable indicating whether the child is enrolled to any level of education; attendance is an indicator counting the number of days the child attended by the child in the past school week. The indicator ranges from 0 to 5 days. Current grade provides the educational grade (in years) the child is currently enrolled in. Household poverty is a dichotomous indicator having the value of one if the household had baseline per capita consumption levels falling below the national consumption poverty line in 2013. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Table 6.

Treatment effects of school feeding on child cognitive scores, full sample, and heterogeneity by child gender, household poverty, and geographical areas

	Digit Span	**Standardized progressive matrices**	**Composite cognitive score**
Panel A. All children			
School feeding	0.119	0.129	0.143
	[-0.010 - 0.249]	[0.018 - 0.240]	[0.019 - 0.267]
Unadjusted p-value	0.070*	0.024**	0.024**
R-W p-value	0.070*	0.046**	
Observations	2,305	2,307	2,321
R-squared	0.050	0.034	0.064
Mean Treatment Endline	0.0992	0.119	0.171
Mean Control Endline	-0.0348	-0.0326	0

	Digit Span		**Standardized progressive matrices**		**Composite cognitive score**	
Panel B: Gender	Girls	Boys	Girls	Boys	Girls	Boys
School feeding	0.190 [0.0372 - 0.342]	0.0649 [-0.207 - 0.336]	0.116 [-0.0358 - 0.267]	0.148** [0.0191 - 0.276]	0.175** [0.0215 - 0.329]	0.124* [-0.0223 - 0.271]
Unadjusted p-value	0.015**	0.404	0.133	0.025**	[0.0260]	[0.0957]
R-W p-value	0.032**	0.403	0.122	0.042**		
Observations	1,085	1,220	1,086	1,221	1,091	1,230
R-squared	0.057	0.053	0.043	0.037	0.071	0.068
Mean Treatment Endline	0.0977	0.101	0.0687	0.165	0.140	0.200
Mean Control Endline	0.108	0.0288	-0.0942	0.0208	-0.0828	0.0717

Table 6 (ctd.)

	Digit Span		**Standardized progressive matrices**		**Composite cognitive score**	
Panel C: Household Poverty	Below the poverty line	Non-poor	Below the poverty line	Non-poor	Below the poverty line	Non-poor
School feeding	0.269 [0.108 - 0.429]	0.0723 [-0.172 - 0.316]	0.234 [0.100 - 0.367]	0.0931 [-0.0427 - 0.229]	0.293 [0.147 - 0.440]	0.0946 [-0.0517 - 0.241]
Unadjusted p-value	0.001***	0.339	0.001***	0.177	0.0002***	0.202
R-W p-value	0.002***	0.342	0.002***	0.309		
Observations	540	1,765	537	1,770	542	1,779
R-squared	0.096	0.042	0.071	0.030	0.112	0.056
Mean Treatment Endline	0.127	0.0907	0.138	0.113	0.196	0.164
Mean Control Endline	-0.207	0.0175	-0.161	0.00626	-0.180	0.0547
	Digit Span		**Standardized progressive matrices**		**Composite cognitive score**	
Panel D: Place of residence	North	South	North	South	North	South
School feeding	0.253*** [0.0798 - 0.427]	-0.002 [-0.323 - 0.318]	0.212** [0.040 - 0.384]	0.051 [-0.0922 - 0.194]	0.272*** [0.0994 - 0.445]	0.025 [-0.147 - 0.198]
Unadjusted p-value	[0.00539]	[0.981]	[0.0171]	[0.477]	[0.00291]	[0.772]
R-W p-value	0.009	0.98	0.013	0.726		
Observations	1,093	1,212	1,096	1,211	1,099	1,222
R-squared	0.043	0.064	0.028	0.042	0.055	0.079
Mean Treatment Endline	0.131	0.066	0.176	0.06	0.226	0.114
Mean Control Endline	-0.109	0.024	-0.0498	-0.019	-0.0516	0.040

Notes: *** p<0.01, ** p<0.05, * p<0.1. Confidence intervals clustered at community levels in squared brackets. R-W p-values were adjusted for multiple testing using the Romano-Wolf (2005, 2016) step-down method with 2,000 iterations and standard errors clustered at community level.

The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and place of residence. Models were estimated through OLS. For each outcome, the model controls for its baseline value, a dichotomous variable related to the randomized assignment to school feeding, and region dummies. Household poverty is a dichotomous indicator having the value of one if the household had baseline per capita consumption levels falling below the national consumption poverty line in 2013. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Appendix 1. Literature review on the effects of school feeding on child learning

Considerable experimental evidence exists on the positive effects of school feeding programs on school participation, although there is variation in effect sizes (Alderman and Bundy 2012; Drake et al. 2017). By contrast, fewer experiments have provided evidence on the impacts of school feeding on learning, with mixed findings. This overall result may be driven by the insufficiency of school meals to raise achievements on their own in poor learning environments and high food insecurity settings, where the trade-offs between children's schooling and work may be especially high. Also, important factors shaping the magnitude of the effects include schooling rates, modality of feeding, quality of program implementation, differences in study methodologies and in target populations, as well as the overall country context. The remainder of this section presents existing evidence. For a meta-analysis of learning effects combining different interventions, see Snilstveit et al. (2015). We focus here on experiments evaluating programs that provide effects in terms of learning outcomes in both primary and preprimary schools in low- and middle-income countries. We note that none of those studies evaluate government programs.

After one year of implementation, a field experiment evaluating different implementation modalities of the World Food Programme school feeding program in primary schools showed that the school feeding increased math scores for girls only in internally-displaced people camps in Northern Uganda (Kazianga, de Walque, and Alderman 2012). A randomized trial in 16 rural Jamaican schools showed that primary school children receiving a school breakfast had increases in their math achievements, and that effects were stronger among undernourished children (Grantham-McGregor, Chang, and Walker 1998; Powell et al. 1998). A two-year randomized trial set in 12 schools in one rural district in Kenya focusing on providing meat, milk, and an "energy" meal to 360 primary school-children as a mid-morning snack, documented improved arithmetic test scores for children in the meat and milk groups (Hulett et al. 2014; Neumann et al. 2007). Another study conducted in two districts in Western Kenya documented that a preschool breakfast program run by a Dutch NGO increases in preschooler's curricula test scores, but only for those attending more often and had a more experienced teacher (Vermeersch and Kremer 2005). Recently, two trials have used the infrastructure of the Indian school feeding program to scale up food fortification, with no effects on learning overall (Berry et al. 2018; Krämer, Kumar, and Vollmer 2018). However, Krämer et al. found positive treatment effects of about 0.2

standard deviations on math and reading for pupils who had more than 80 or 90 percent school-attendance.

Table A.1.

Summary characteristics of existing randomized control trials of school feeding in low- and middle-income countries

Study	Location	Study duration	Treatment	Implementer	Geographical area
Ahmed (2004)	Bangladesh	14 months	Snack	WFP and partner NGOs	9 districts
Alderman et al., (2012)	Refugee camps in Northern Uganda	28 months	School feeding or take-home rations	WFP	10 refugee camps
Andang'o et al. (2007)	Kenya	5 months	Fortified maize	Researchers	4 schools
Ash et al., (2003)	Tanzania	6 months	Fortified beverages	Researchers	6 primary schools
Berry et al. (2018)	Odisha (India)	3 years	Micronutrient supplementation in existing program	Researchers	1 district
Grantham-McGregor et al., (1998)	Jamaica	21 weeks	Breakfast	Researchers	4 rural primary schools
Kazianga et al., (2009)	Northern Burkina Faso	1 year	School feeding or take-home rations	WFP	46 schools
Kramer et al. (2018)	Bihar (India)	1 year	Micronutrient supplementation in existing program	Researchers	Two blocks of Jenahabad district
Moretti et al. (2006)	Bangalore (India)	9 months	Fortified rice	Researchers	1 school
Muthayya et al., (2007)	Urban Bangalore, India`	1 year	Snack	Researchers	69 children
Osendarp et al., (2007)	Central Jakarta, Indonesia	1 year	Snack	Researchers	6 schools
Radhika et al. (2011)	Andhra Pradesh (India)	8 months	Fortified rice	Researchers	1 school
van Stuijvenberg et al., (1999)	South Africa	year	Snack and beverage	Researchers	1 school
Vermeersch and Kremer (2005)	Western Kenya	2 years	Breakfast	NGO	25 schools in 2 districts
Whaley et al., (2003); Hulett et al. (2014); Neumann et al. (2007)	Kenya	21 months	School feeding	Researchers	12 schools
Zimmermann et al. (2003)	Morocco	9 months	Fortified salt	Researchers	2 schools

Notes: the table above provides characteristics in terms of implementer, duration, program type and geographical areas of existing randomized trials of school feeding. Studies were first identified through the systematic review conducted in 2012 by Lawson (Lawson 2012), which the authors had updated.

References appendix

Acemoglu, Daron. 2010. "Theory, General Equilibrium, and Political Economy in Development

Economics." *Journal of Economic Perspectives* 24 (3): 17–32. https://doi.org/10.1257/jep.24.3.17.

Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2013. "School Meals and Classroom Effort: Evidence from India." *International Growth Centre.* Vol. 25. http://www.theigc.org/wp-content/uploads/2014/09/Afridi-Et-Al-2013-Working-Paper1.pdf.

———. 2019. "Hunger and Performance in the Classroom."

Ahmed, Akhter U. 2004. "Impact of Feeding Children in School: Evidence from Bangladesh." *International Food Policy Research Institute* 2033: 20006–1002. http://www.lcgbangladesh.org/FSN/reports/IFPRI Final Report_School Feeding in Bangladesh.pdf.

Akresh, Richard, D De Walque, and Harounan Kazianga. 2013. "Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality." *World Bank Policy Research*. Policy Research Working Papers. The World Bank. https://doi.org/10.1037/e547172013-001.

Alderman, Harold, Jere R. Behrman, David R. Ross, and Richard Sabot. 1996. "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labour Market." *Oxford Bulletin of Economics and Statistics*. https://doi.org/10.1111/j.1468-0084.1996.mp58001003.x.

Alderman, Harold, and Donald Bundy. 2012. "School Feeding Programs and Development: Are We Framing the Question Correctly?" *World Bank Research Observer* 27 (2): 204–21. https://doi.org/10.1093/wbro/lkr005.

Alderman, Harold, Ugo Gentilini, and Ruslan Yemtsov. 2018. *The 1.5 Billion People Question: Food, Vouchers, or Cash Transfers?* Edited by Harold Alderman, Ugo Gentilini, and Ruslan Yemtsov. Washington, D. C.: The World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/27907/9781464810879.pdf.

Alderman, Harold, Daniel O. Gilligan, and Kim Lehrer. 2012. "The Impact of Food for Education Programs on School Participation in Northern Uganda." *Economic Development and Cultural Change* 61 (1): 187–218. https://doi.org/10.1086/666949.

Allcott, Hunt, and Sendhil Mullainathan. 2012. "External Validity and Partner Selection Bias." *NBER Working Paper 18373*, 1–44. https://doi.org/10.1093/qje/qjv015.Advance.

Andang'o, Pauline EA, Saskia JM Osendarp, Rosemary Ayah, Clive E. West, David L. Mwaniki, Corine A. De Wolf, Rob Kraaijenhagen, Frans J. Kok, and Hans Verhoef. 2007. "Efficacy of Iron-Fortified Whole Maize Flour on Iron Status of Schoolchildren in Kenya: A Randomised Controlled Trial." *Lancet*. https://doi.org/10.1016/S0140-6736(07)60817-4.

Ash, Deborah M., Simon R. Tatala, Edward A. Frongillo, Godwin D. Ndossi, and Michael C. Latham. 2003. "Randomized Efficacy Trial of a Micronutrient-Fortified Beverage in Primary School Children in Tanzania." *The American Journal of Clinical Nutrition*.

Aurino, Elisabetta, Jean-Pierre Tranchant, Amadou Sekou Diallo, and Aulo Gelli. 2018. "School Feeding or General Food Distribution? Quasi-Experimental Evidence on the Educational Impacts of Emergency Food Assistance during Conflict in Mali." Innocenti Working Paper. Florence. https://www.unicef-irc.org/publications/pdf/Aurino et al for WEB.pdf.

Baird, Sarah, Francisco H.G. Ferreira, Berk Özler, and Michael Woolcock. 2014. "Conditional, Unconditional and Everything in between: A Systematic Review of the Effects of Cash Transfer Programmes on Schooling Outcomes." *Journal of Development Effectiveness* 6 (1): 1–43. https://doi.org/10.1080/19439342.2014.890362.

Balwanz, David, and Peter Darvas. 2013. *Basic Education beyond the Millennium Development Goals in Ghana: How Equity in Service Delivery Affects Educational and Learning*

Outcomes. The World Bank. https://doi.org/10.1596/978-1-4648-0098-6.
Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236). https://doi.org/10.1126/science.1260799.
Banerjee, Abhijit V., Shawn Cole, Esther; Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics* 122 (3): 1235–64. https://doi.org/10.1162/qjec.122.3.1235.
Bashir, Sajitha, Marlaine Lockheed, Elizabeth Ninan, and Jee-Peng Tan. 2018. *Facing Forward: Schooling for Learning in Africa*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1260-6.
Berry, James, S. Metha, P. Mukherjee, H. Ruebeck, and G. Kartini Shastry. 2018. "Inputs, Monitoring and Crowd-Out in India School-Based Health Interventions." S-89206-INC-2. London.
Björkman-Nyqvist, Martina. 2013. "Income Shocks and Gender Gaps in Education: Evidence from Uganda." *Journal of Development Economics* 105 (November): 237–53. https://doi.org/10.1016/j.jdeveco.2013.07.013.
Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2018. "Experimental Evidence on Scaling up Education Reforms in Kenya." *Journal of Public Economics* 168 (December): 1–20. https://doi.org/10.1016/J.JPUBECO.2018.08.007.
Bütikofer, Aline, Eirin Mølland, and Kjell G. Salvanes. 2018. "Childhood Nutrition and Labor Market Outcomes: Evidence from a School Breakfast Program." *Journal of Public Economics* 168 (December): 62–80. https://doi.org/10.1016/j.jpubeco.2018.08.008.
Chakraborty, Tanika, and Rajshri Jayaraman. 2019. "School Feeding and Learning Achievement: Evidence from India's Midday Meal Program." *Journal of Development Economics*, April. https://doi.org/10.1016/J.JDEVECO.2018.10.011.
Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Score." *American Economic Journal: Applied Economics* 5 (2): 29–57. https://doi.org/10.1257/app.5.2.29.
Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55. https://doi.org/10.1257/jel.48.2.424.
Drake, Lesley, Meena Fernandes, Elisabetta Aurino, Josephine Kiamba, Boitshepo Giyose, Carmen Burbano, Harold Alderman, Lu Mai, Arlene Mitchell, and Aulo Gelli. 2017. "School Feeding Programs in Middle Childhood and Adolescence." In *Disease Control Priorities 3*, edited by D. Bundy, N. De Silva, S. Horton, D. Jamison, and G.C. Patton. Washington, D.C.: The World Bank. http://dcp-3.org/sites/default/files/chapters/DCP3 CAHD_Ch 12.pdf.
Duflo, Esther;, Pascaline Dupas, and Michael Kremer. 2017. "The Impact of Free Secondary Education: Experimental Evidence from Ghana." https://www.poverty-action.org/sites/default/files/publications/DDK_GhanaScholarships.pdf.
Evans, D. K., and F. Yuan. 2018. "Commentary." In *Learning at the Bottom of the Pyramid.*, edited by Daniel A. Wagner, Sharon Wolf, and Robert F. Boruch, 232–34. Paris: UNESCO.
Fernandes, Meenakshi, Gloria Folson, Elisabetta Aurino, and Aulo Gelli. 2017. "A Free Lunch or a Walk Back Home? The School Food Environment and Dietary Behaviours among Children and Adolescents in Ghana." *Food Security*, October 1, 2017.

https://doi.org/10.1007/s12571-017-0712-0.
Figlio, David N., and Joshua Winicki. 2005. “Food for Thought: The Effects of School Accountability Plans on School Nutrition.” *Journal of Public Economics* 89 (2–3): 381–94. https://doi.org/10.1016/J.JPUBECO.2003.10.007.
Filmer, Deon P., Jed Friedman, Eeshani Kandpal, and Junko Onishi. 2018. “General Equilibrium Effects of Targeted Cash Transfers: Nutrition Impacts on Non-Beneficiary Children.” 8377. *Policy Research Working Paper*. Vol. 8377. Washington, D.C. http://documents.worldbank.org/curated/en/989031522077749796/General-equilibrium-effects-of-targeted-cash-transfers-nutrition-impacts-on-non-beneficiary-children.
Frison, Lars, and Stuart J Pocock. 1992. “Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and Its Implications for Design.” *Statistics in Medicine* 11 (13): 1685–1704. https://doi.org/10.1002/sim.4780111304.
Gelli, A., E. Masset, G. Folson, A. Kusi, D.K. Arhinful, F. Asante, I. Ayi, et al. 2016. “Evaluation of Alternative School Feeding Models on Nutrition, Education, Agriculture and Other Social Outcomes in Ghana: Rationale, Randomised Design and Baseline Data.” *Trials* 17 (1). https://doi.org/10.1186/s13063-015-1116-0.
Gelli, Aulo, Elisabetta Aurino, Gloria Folson, Daniel Arhinful, Clement Adamba, Isaac Osei-Akoto, Edoardo Masset, et al. 2019. “A School Meals Program Implemented at Scale in Ghana Increases Height-for-Age during Midchildhood in Girls and in Children from Poor Households: A Cluster Randomized Trial.” *The Journal of Nutrition*, May. https://doi.org/10.1093/jn/nxz079.
Gelli, Aulo, and Roshan Daryanani. 2013. “Are School Feeding Programs in Low-Income Settings Sustainable? Insights on the Costs of School Feeding Compared with Investments in Primary Education.” *Food and Nutrition Bulletin* 34 (3): 310–17.
Gelli, Aulo, Francisco Espejo, Jing Shen, and Elizabeth Kristjansson. 2014. “Putting It All Together: Aggregating Impacts of School-Feeding Programmes on Education, Health and Nutrition: Two Proposed Methodologies.” *WIDER Working Paper Series*. https://ideas.repec.org/p/unu/wpaper/wp2014-036.html.
Ghana Statistical Service. 2018. “Ghana Living Standards Survey (GLSS7): Poverty Trends in Ghana; 2005-2017.” *Ghana Statistical Service*.
Glewwe, Paul, and Michael Kremer. 2006. “Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries.” *Handbook of the Economics of Education* 2: 945–1017. https://doi.org/10.1016/S1574-0692(06)02016-2.
Graff Zivin, Joshua, Solomon M. Hsiang, and Matthew Neidell. 2018. “Temperature and Human Capital in the Short and Long Run.” *Journal of the Association of Environmental and Resource Economists* 5 (1): 77–105. https://doi.org/10.1086/694177.
Grantham-McGregor, S M, S Chang, and S P Walker. 1998. “Evaluation of School Feeding Programs: Some Jamaican Examples.” *The American Journal of Clinical Nutrition* 67 (4): 785S-789S. http://www.ncbi.nlm.nih.gov/pubmed/9537629.
Groot, Richard De, Sudhanshu Handa, Michael Park, Robert Osei Darko, Isaac Osei-Akoto, Garima Bhalla, and Luigi Peter Ragno. 2015. “Heterogeneous Impacts of an Unconditional CashTransfer Programme on Schooling: Evidence from the Ghana LEAP Programme.” Working Paper.
GSFP. 2006. “Ghana School Feeding Programme. Programme Document 2007-2010.” Accra.
Hayes, Richard J., and Lawrence H. Moulton. 2017. *Cluster Randomised Trials*. 2nd ed. Chapman and Hall/CRC.

Hulett, Judie L., Robert E. Weiss, Nimrod O. Bwibo, Osman M. Galal, Natalie Drorbaugh, and Charlotte G. Neumann. 2014. "Animal Source Foods Have a Positive Impact on the Primary School Test Scores of Kenyan Schoolchildren in a Cluster-Randomised, Controlled Feeding Intervention Trial." *British Journal of Nutrition* 111 (05): 875–86. https://doi.org/10.1017/S0007114513003310.

Jacoby, Hanan G. 2002. "Is There an Intrahousehold ' Flypaper Effect '? Evidence from a School Feeding Programme." *Economic Journal* 112 (476): 196–221.

Jukes, M C H, L J Drake, and D A P Bundy. 2008. *School Health, Nutrition and Education for All: Levelling the Playing Field*. Wallingford: CABI Publishing. https://doi.org/http://dx.doi.org/10.1079/9781845933111.0011.

Kazianga, H., D. de Walque, and H. Alderman. 2012. "Educational and Child Labour Impacts of Two Food-for-Education Schemes: Evidence from a Randomised Trial in Rural Burkina Faso." *Journal of African Economies* 21 (5): 723–60. https://doi.org/10.1093/jae/ejs010.

Kazianga, Harounan, Damien de Walque, and Harlod Alderman. 2009. "Educational and Health Impacts of Two School Feeding Schemes Evidence from a Randomized Trial in Rural Burkina Faso." 4976. World Bank Policy Research Working Paper.

Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2014. "School Feeding Programs, Intrahousehold Allocation and the Nutrition of Siblings: Evidence from a Randomized Trial in Rural Burkina Faso." *Journal of Development Economics* 106 (January): 15–34. https://doi.org/10.1016/j.jdeveco.2013.08.007.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2006. "Experimental Analysis of Neighborhood Effects." *Econometrica*, no. June: 1–34. https://doi.org/10.1111/j.1468-0262.2007.00733.x.

Krämer, Marion; Kumar, Santosh; Vollmer, Sebastian. 2018. "Improving Children Health and Cognition: Evidence from School-Based Nutrition Intervention in India." 203. GLO Discussion Paper. Maastricht.

Kristjansson, Elizabeth, Damian K Francis, Selma Liberato, Maria Benkhalti Jandu, Vivian Welch, Malek Batal, Trish Greenhalgh, et al. 2015. "Food Supplementation for Improving the Physical and Psychosocial Health of Socio-Economically Disadvantaged Children Aged Three Months to Five Years (Review)." *Cochrane Database of Systematic Reviews*, no. 3. https://doi.org/10.1002/14651858.CD009924.pub2.

Lawson, Ty M. 2012. "Impact of School Feeding Programs on Educational, Nutritional, and Agricultural Development Goals: A Systematic Review of Literature." Michigan State University. https://www.3ieimpact.org/evidence-hub/systematic-review-repository/impact-school-feeding-programs-educational-nutritional.

McKenzie, David. 2012. "Beyond Baseline and Follow-up: The Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210–21. https://doi.org/10.1016/j.jdeveco.2012.01.002.

Ministry of Education & RTI International. 2014. "Ghana 2013 National Education Assessment - Technical Report | Global Reading Network." https://globalreadingnetwork.net/eddata/ghana-2013-national-education-assessment-technical-report.

Moretti, Diego, Michael B. Zimmermann, Sumithra Muthayya, Prashanth Thankachan, Tung Ching Lee, Anura V. Kurpad, and Richard F. Hurrell. 2006. "Extruded Rice Fortified with Micronized Ground Ferric Pyrophosphate Reduces Iron Deficiency in Indian Schoolchildren: A Double-Blind Randomized Controlled Trial1-3." *American Journal of*

Clinical Nutrition. https://doi.org/10.1093/ajcn/84.4.822.
Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at Scale." *Journal of Economic Perspectives* 31 (4): 103–24. https://doi.org/10.1257/jep.31.4.103.
Muthayya, Sumithra, Tinku Thomas, Krishnamachari Srinivasan, Kirthi Rao, Anura V. Kurpad, Jan Willem van Klinken, Gail Owen, and Eveline A. de Bruin. 2007. "Consumption of a Mid-Morning Snack Improves Memory but Not Attention in School Children." *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2006.09.025.
Neumann, Charlotte G, Suzanne P Murphy, Connie Gewa, Monika Grillenberger, and Nimrod O Bwibo. 2007. "Meat Supplementation Improves Growth, Cognitive, and Behavioral Outcomes in Kenyan Children." *The Journal of Nutrition* 137 (4): 1119–23. http://www.ncbi.nlm.nih.gov/pubmed/17374691.
Osendarp, Saskia J.M., Katrine I. Baghurst, Janet Bryan, Eva Calvaresi, Donna Hughes, Mahdin Hussaini, Elvina Karyadi, et al. 2007. "Effect of a 12-Mo Micronutrient Intervention on Learning and Memory in Well-Nourished and Marginally Nourished School-Aged Children: 2 Parallel, Randomized, Placebo-Controlled Studies in Australia and Indonesia." *American Journal of Clinical Nutrition*.
Powell, C A, S P Walker, S M Chang, and S M Grantham-McGregor. 1998a. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." *The American Journal of Clinical Nutrition* 68 (4): 873–79.
———. 1998b. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." *The American Journal of Clinical Nutrition* 68 (4): 873–79. http://www.ncbi.nlm.nih.gov/pubmed/9771865.
Radhika, Madhari S., Krishnapillai M. Nair, Rachakulla Hari Kumar, Mendu Vishnuvardhana Rao, Punjal Ravinder, Chitty Gal Reddy, and Ginnela N.V. Brahmam. 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*. https://doi.org/10.3945/ajcn.110.007179.
Romano, Joseph P., and Michael Wolf. 2005. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica*. https://doi.org/10.1111/j.1468-0262.2005.00615.x.
———. 2016. "Efficient Computation of Adjusted P-Values for Resampling-Based Stepdown Multiple Testing." *Statistics and Probability Letters*. https://doi.org/10.1016/j.spl.2016.02.012.
Ruel, M.T. 2003. "Operationalizing Dietary Diversity: A Review of Measurement Issues and Research Priorities." *Journal of Nutrition* 133: 3911S-3926S.
SEND-Ghana. 2013. "Ghana School Feeding Programme: Way Forward."
Snilstveit, Birte, Jennifer Stevenson, Daniel Phillips, Martina Vojtkova, Emma Gallagher, Tanja Schmidt, Hannah Jobse, Maisie Geelen, Maria Grazia Pastorello, and John Eyers. 2015. "Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review." London.
Stuijvenberg, M E van, J D Kvalsvig, M Faber, M Kruger, D G Kenoyer, and a J Benadé. 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): 497–503.
UNDP Ghana. 2015. "2015 Ghana Millennium Development Goals Report." Accra. http://www.gh.undp.org/content/ghana/en/home/library/poverty/2015-ghana-millennium-

development-goals-report.html.
Vermeersch, Christel, and Michael Kremer. 2005. “School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation.” Policy Research Working Paper 3523. Washington, D. C. https://doi.org/10.2139/ssrn.667881.
Vivalt, Eva. 2015. “Heterogeneous Treatment Effects in Impact Evaluation.” *American Economic Review* 105 (5): 467–70. https://doi.org/10.1257/aer.p20151015.
———. 2019. “How Much Can We Generalize from Impact Evaluations ?” *Forthcoming, Journal European Economic Association*.
WFP. 2013. *State of School Feeding Worldwide 2013*. Rome: World Food Programme.
Whaley, Shannon E., Marian Sigman, Charlotte Neumann, Nimrod Bwibo, Donald Guthrie, Robert E. Weiss, Susan Alber, and Suzanne P. Murphy. 2003. “The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children.” *The Journal of Nutrition*. https://doi.org/10.1093/jn/133.11.3965s.
World Bank. 2012. *Improving the Targeting of Social Programs in Ghana*. Edited by Quentin Wodon. *World Bank*. https://doi.org/10.1596/978-0-8213-9593-6.
———. 2017. “World DataBank: Education Statistics - All Indicators.” 2017. http://databank.worldbank.org/data/reports.aspx?source=education-statistics-~-all-indicators#.
———. 2018. “World Development Report 2018: Learning to Realize Education’s Promise.” Washington, D.C. https://doi.org/10.1596/978-1-4648-1096-1.
Zimmermann, Michael B., Christophe Zeder, Noureddine Chaouki, Amina Saad, Toni Torresani, and Richard F. Hurrell. 2003. “Dual Fortification of Salt with Iodine and Microencapsulated Iron: A Randomized, Double-Blind, Controlled Trial in Moroccan Schoolchildren.” *American Journal of Clinical Nutrition*. https://doi.org/10.1093/ajcn/77.2.425.

[i] The government approved an expansion of the program to over 3 million children by July 2016, but data on actual coverage are not currently available (http://mogcsp.gov.gh/ghana-school-feeding-programme-gsfp/).

[ii] The results of the analyses on child anthropometrics and community agriculture are part of separate analyses, as per our protocol.

[iii] The original design included a three-level randomization, with a subset of HGSF schools receiving micronutrients constituting a fourth study arm (A. Gelli et al. 2016). However, the discovery of GSFP already running in 25 schools after the baseline survey, and their consequent exclusion from the sample, led to the removal of the fourth study arm due to significant loss of power. This decision was also based on substantial delays in delivering micronutrient powders to the schools that were supposed to receive them.

[iv] In the case of young children, the caregiver reported on schooling.

[v] This result did not change when we split treatment in GSFP and HGSF pilots (available upon request).

[vi] Controlling for interviewer dummies is a common practice in similar standardizations. It also helped tackling potential language effects, as unfortunately we do not have information on the specific language of test administration. The interviewer spoke the same language of the child.

[vii] Although children were given assessments in all tests, discrepancies in sample sizes across raw and standardised scores reflect inability to convert raw scores into standardised scores (e.g. lack of child age in months). A similar issue is highlighted in Graff Zivin and co-authors (Graff Zivin, Hsiang, and Neidell 2018). This could be a potential concern if the missing scores correlate with treatment assignment. Regressions of treatment on score availability rules out this

hypothesis, as the coefficients are zero and not statistically significant across all outcomes (result available upon request).

[viii] A similar picture emerged from the analysis of baseline differences in raw scores for the baseline sample prior to attrition presented in Online Appendix 5. This provides a further reassurance about potential biases in treatment effects of school feeding on child learning stemming from nonrandom attrition.

[ix] McKenzie (2012), for instance, posits that low autocorrelation ranges between ρ=0.2 to 0.4.

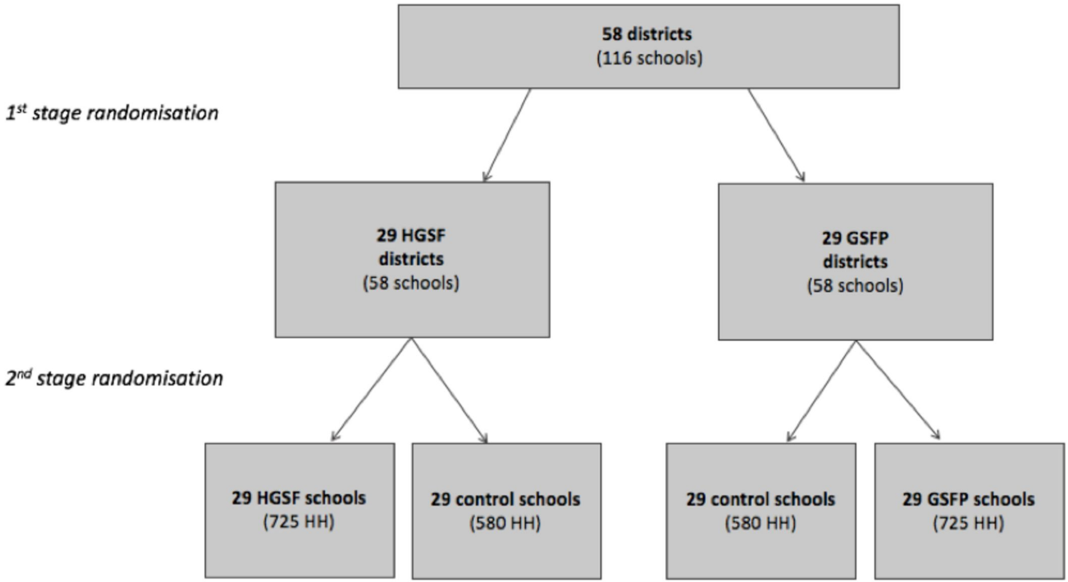
[x] Results are unchanged when a dummy variable for missing baseline test score is included (results available upon request).

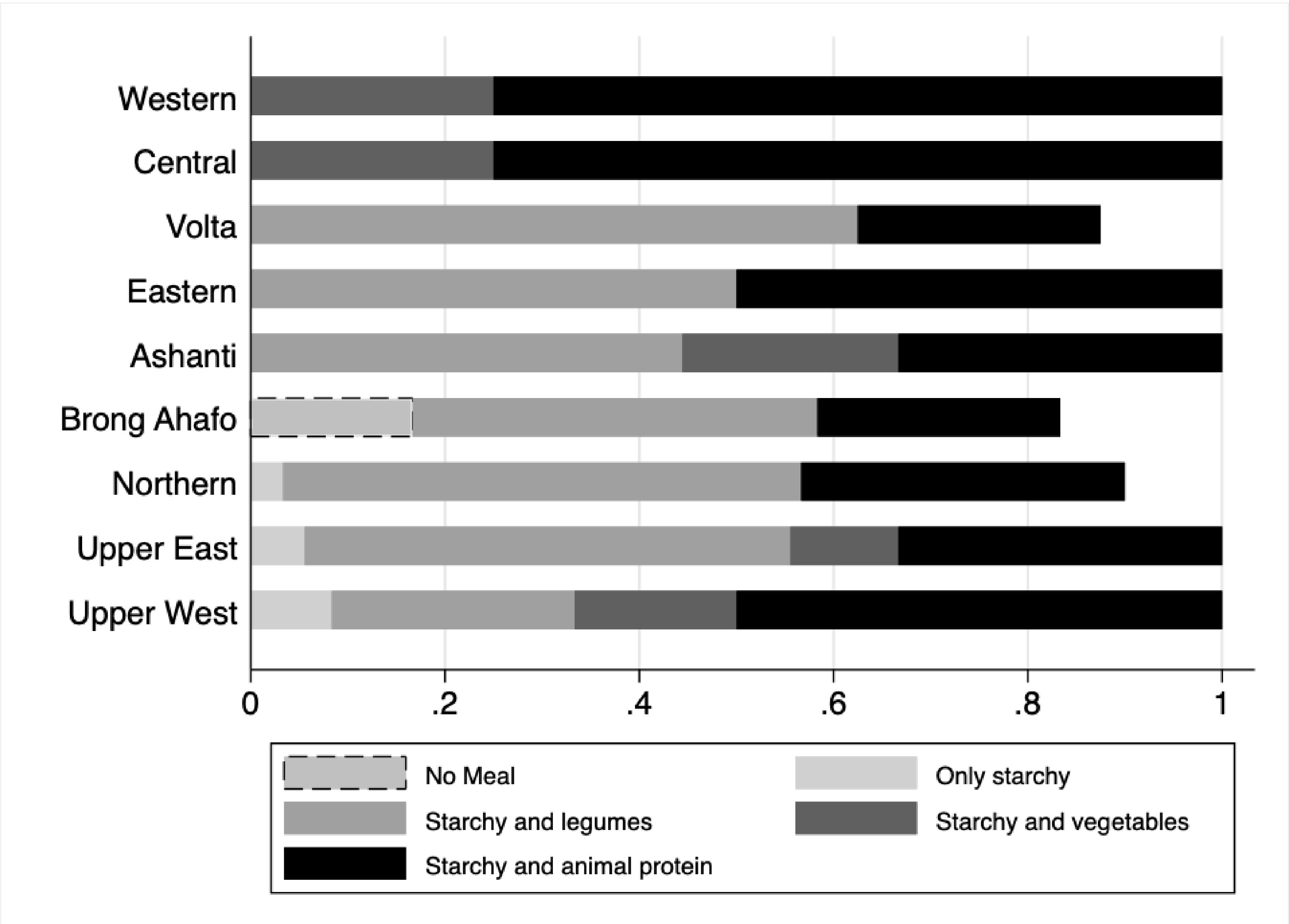
[xi] We opted for this approach, as compared to a different one in which we would interact school feeding with the policy group of interest, for different reasons. First, we wanted to estimate the *total* effect of the policy on each sub-group, and, second, because the stratification has the advantage that the separate regressions allow all parameters to vary by subgroup. Nonetheless, we tested the differential effect in the intervention between each of the comparison groups in a pooled regression model with interactions, and the Romano-Wolf adjusted p-values are around p ~~0.1 in the case of gender and household poverty (results available upon request).

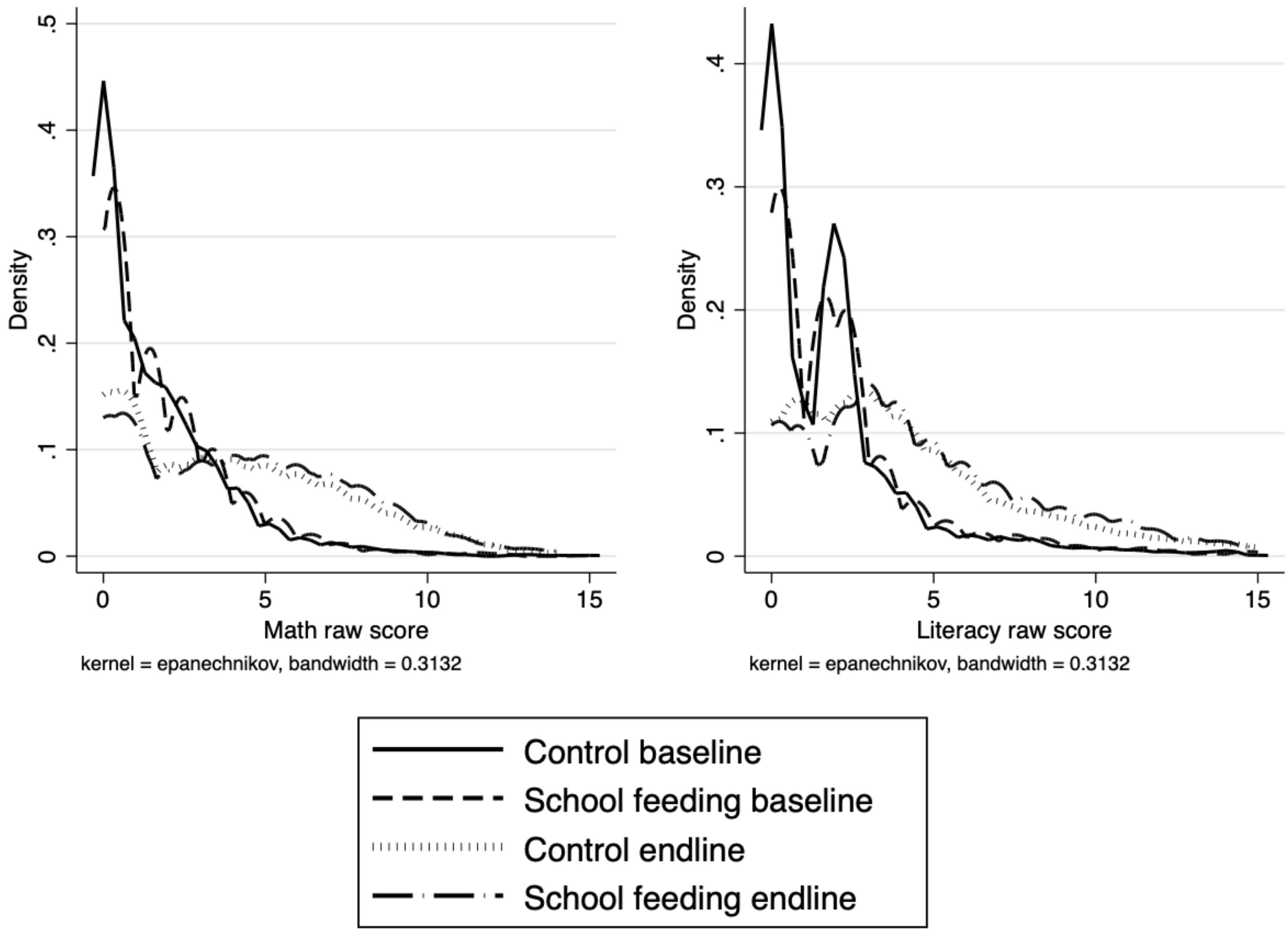
[xii] DiD results are consistent to ANCOVA, and available upon request.

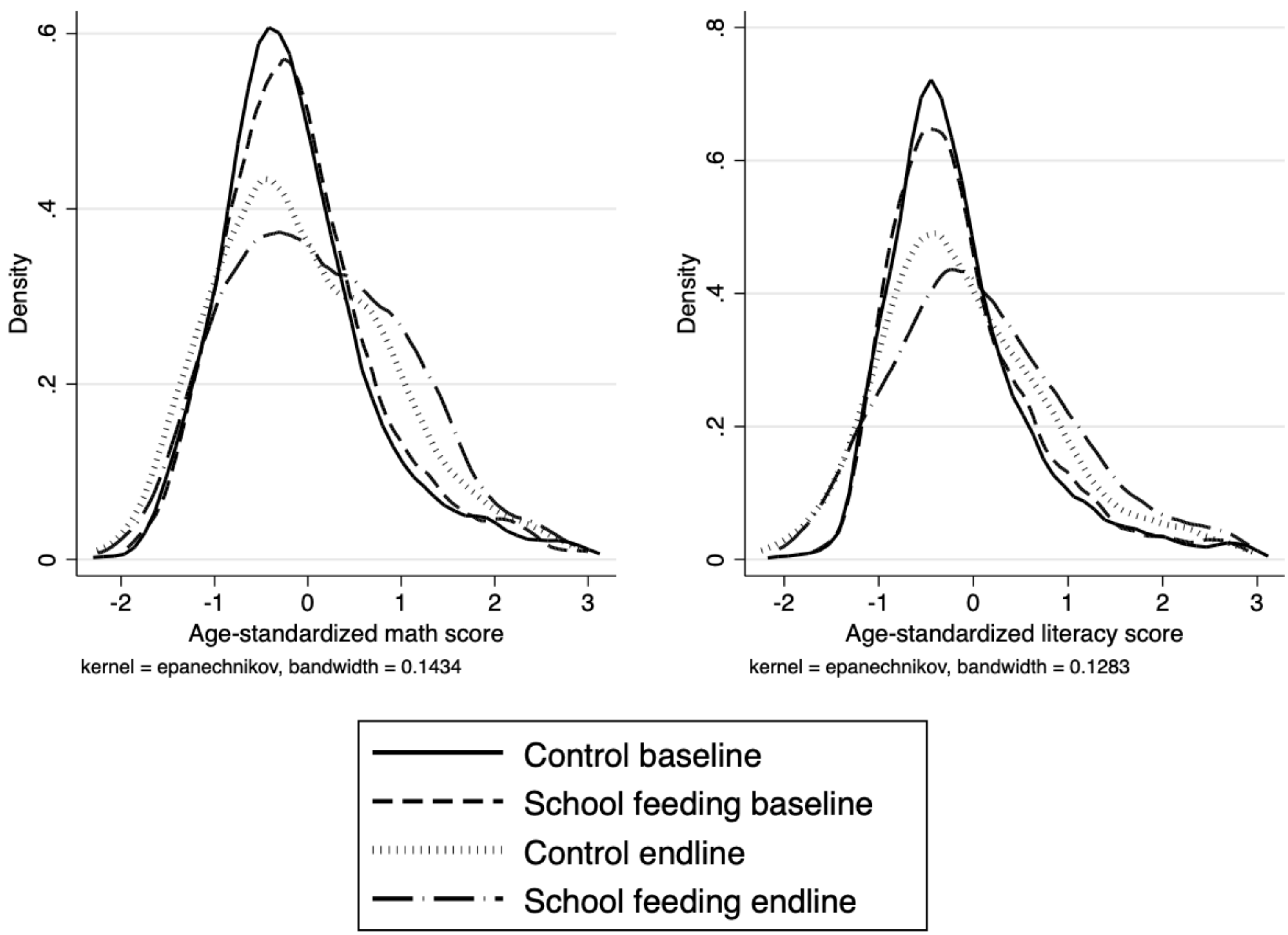
[xiii] In the context of Ghana, Duflo et al. (Duflo, Dupas, and Kremer 2017) have recently assessed the medium-term effects of secondary school scholarships. After eight years, scholarship winners had higher schooling, scored on average 0.15 greater in math and literacy, had better health behaviors, and girls had less children.

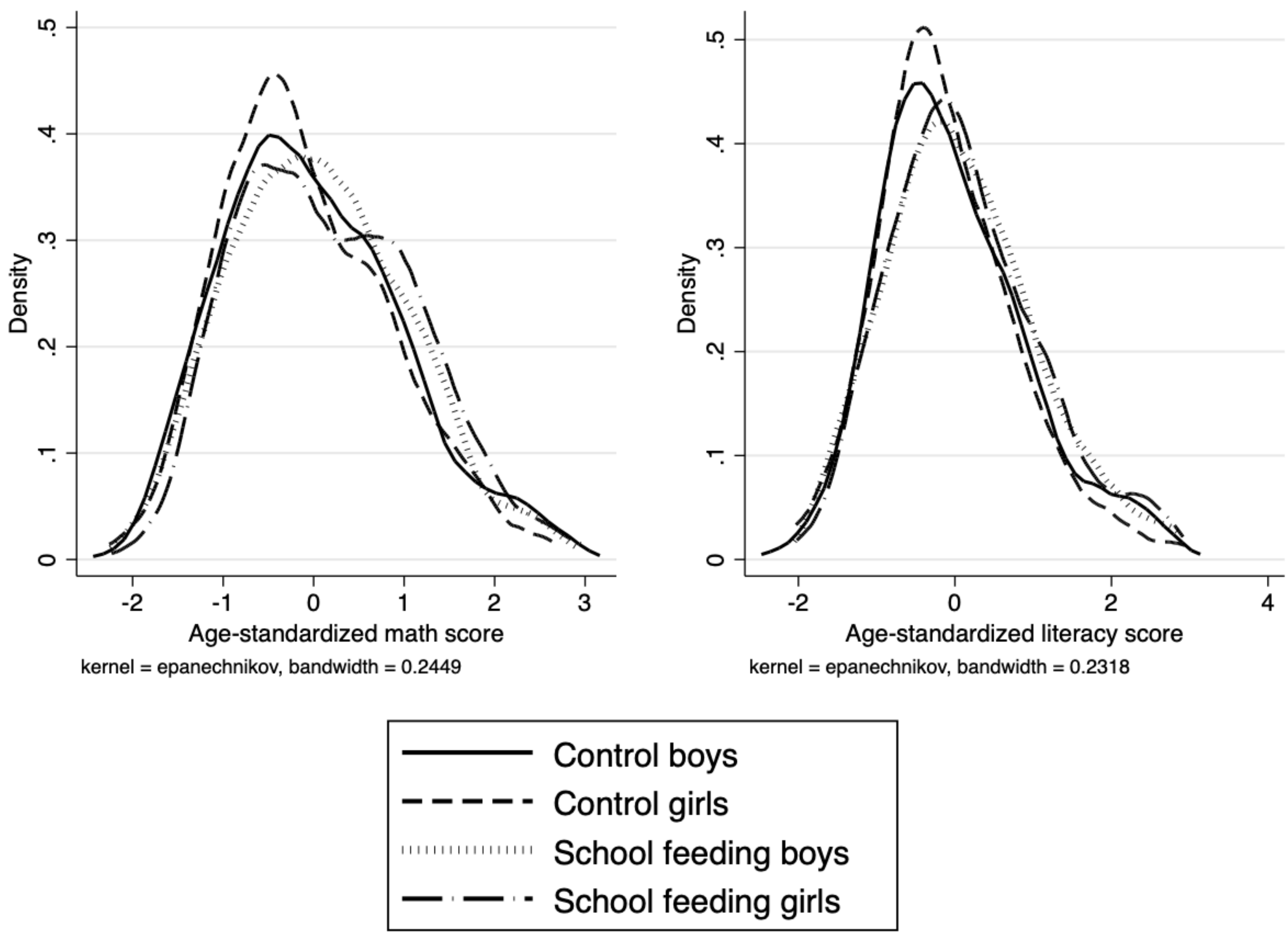
[xiv] For instance, Gelli and Darayani show that between 2005 and 2008, the costs of school feeding increased by 12 percent and 24 percent, on average, in middle- and low-income countries, respectively.


58 districts
(116 schools)
1st stage randomisation
29 HGSF
districts
(58 schools)
29 GSFP
districts
(58 schools)
2nd stage randomisation
29 HGSF schools
(725 HH)
29 control schools
(580 HH)
29 control schools
(580 HH)
29 GSFP schools
(725 HH)

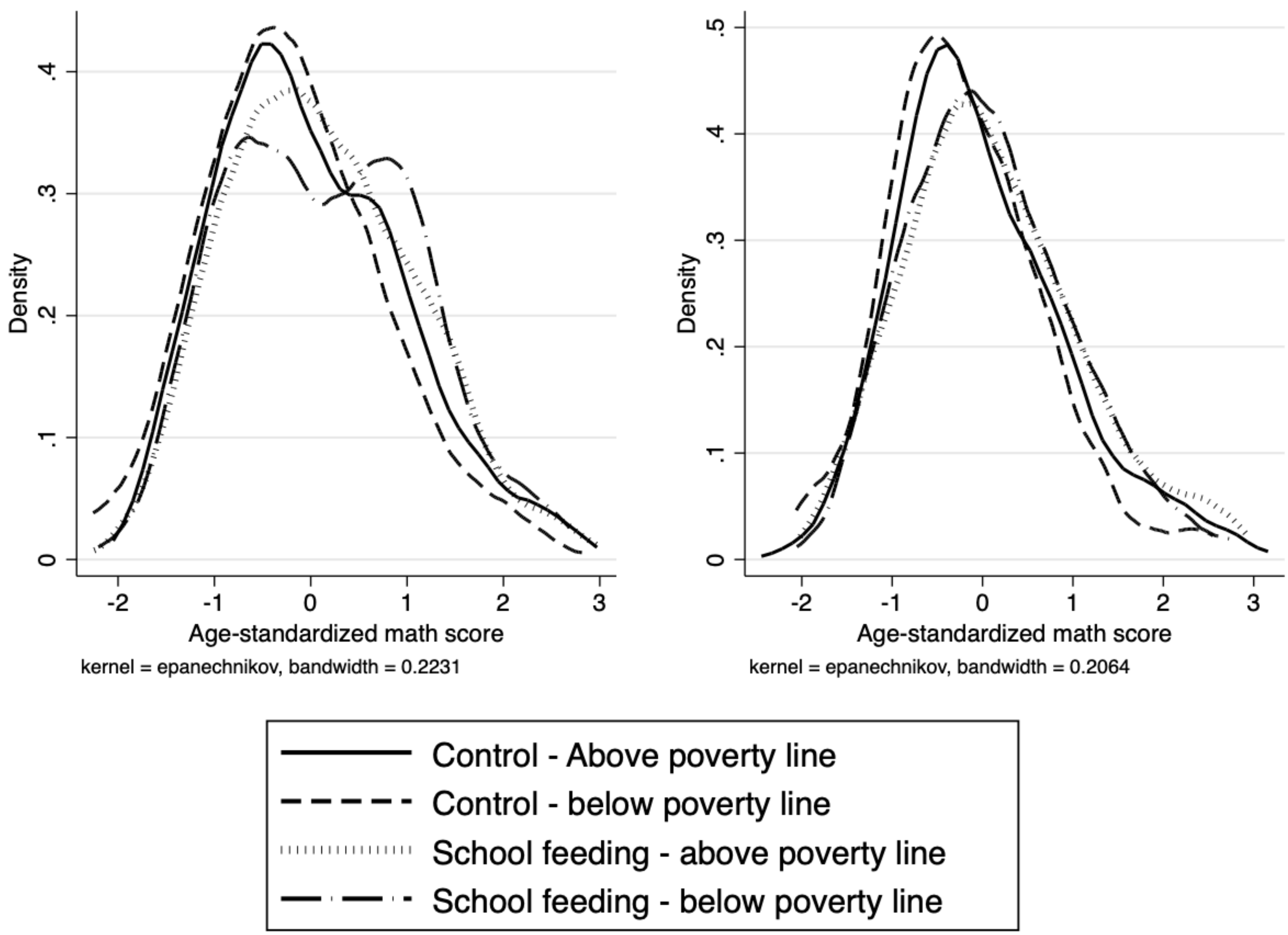

Western
Central
Volta
Eastern
Ashanti
Brong Ahafo
Northern
Upper East
Upper West
0
.2
.4
.6
.8
1
No Meal
Only starchy
Starchy and legumes
Starchy and vegetables
Starchy and animal protein

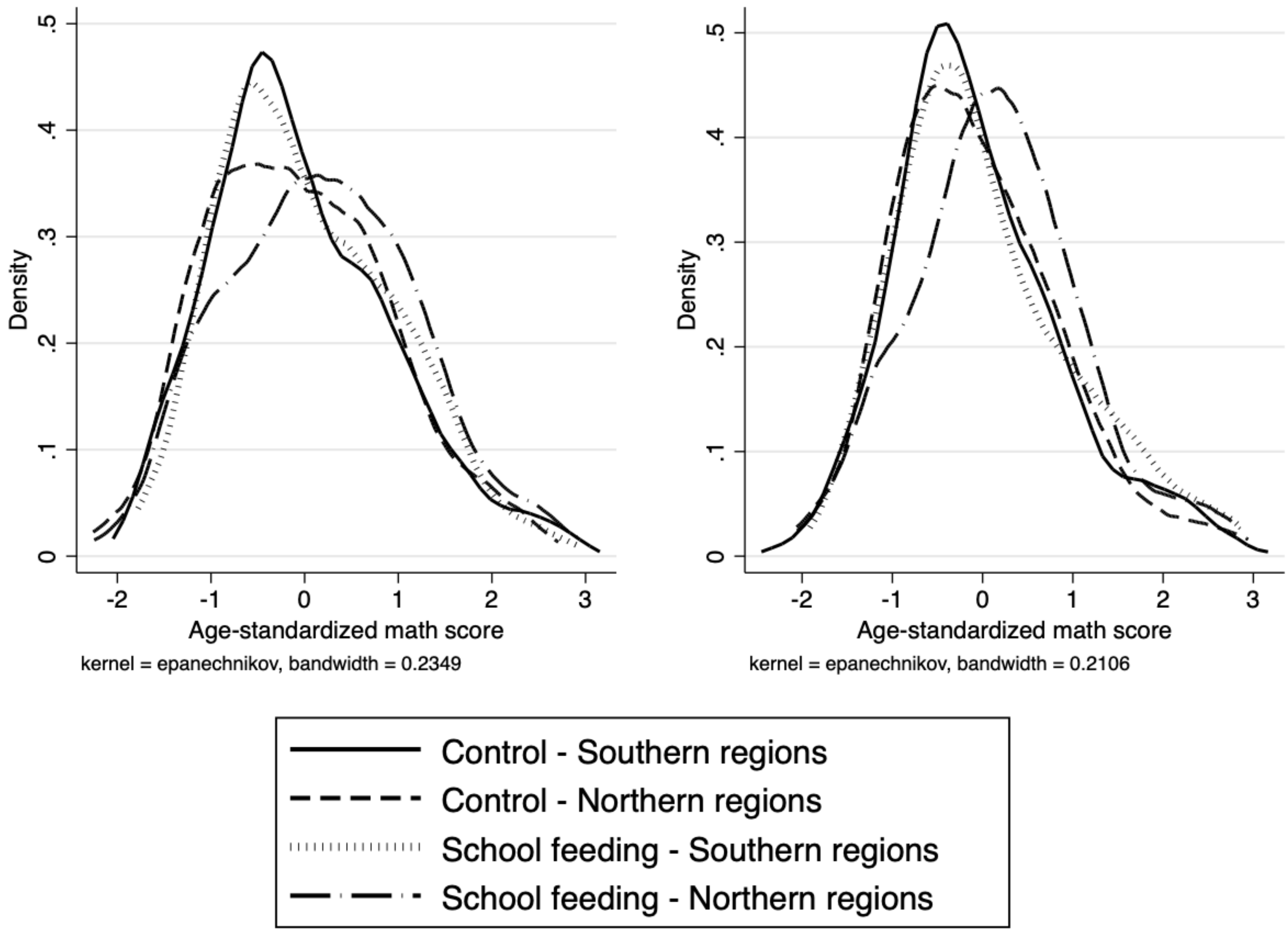

Density
.5
.4
.3
.2
.1
0
0
5
10
15
Math raw score
kernel = epanechnikov, bandwidth = 0.3132
Density
.4
.3
.2
.1
0
0
5
10
15
Literacy raw score
kernel = epanechnikov, bandwidth = 0.3132
Control baseline
School feeding baseline
Control endline
School feeding endline


Density
.6
.4
.2
0
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.1434
Density
.8
.6
.4
.2
0
-2
-1
0
1
2
3
Age-standardized literacy score
kernel = epanechnikov, bandwidth = 0.1283
Control baseline
School feeding baseline
Control endline
School feeding endline



.5
.4
.3
.2
.1
0
Density
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.2449
.5
.4
.3
.2
.1
0
Density
-2
0
2
4
Age-standardized literacy score
kernel = epanechnikov, bandwidth = 0.2318
Control boys
Control girls
School feeding boys
School feeding girls


Density
.4
.3
.2
.1
0
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.2231
Density
.5
.4
.3
.2
.1
0
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.2064
Control - Above poverty line
Control - below poverty line
School feeding - above poverty line
School feeding - below poverty line



Density
.5
.4
.3
.2
.1
0
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.2349
Density
.5
.4
.3
.2
.1
0
-2
-1
0
1
2
3
Age-standardized math score
kernel = epanechnikov, bandwidth = 0.2106
Control - Southern regions
Control - Northern regions
School feeding - Southern regions
School feeding - Northern regions