Food for thought?

Experimental Evidence on the Learning Impacts of a Large-Scale School Feeding Program

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

Abstract

There is limited experimental evidence of the impact of large-scale, government-led school meals programs on educational achievements. We report results from a nationwide randomized trial of the Government of Ghana's school feeding program. After two years, program availability led to moderate increases in test scores for the average pupil, and to remarkable learning and cognitive gains for girls, and children from poorest households and regions. Increases in enrolment, attainment, and shifts in time spent at school constituted mechanisms for impact. The program combined social protection with equitable human capital accumulation, thus contributing to the "learning for all" sustainable development agenda.

Jel codes: *I28, I24, I00*

Keywords: school feeding; school meals; learning; cognition; Ghana; field experiment; primary school education; randomized controlled trial

[1] Aurino: Department of Management and Centre for Health Economics and Policy Innovations. Imperial College Business School, Tanaka Building. South Kensington Campus, Imperial College London. London SW7 2AZ, UK (e.aurino@imperial.ac.uk)

Gelli: Poverty, Health and Nutrition Division, International Food Policy Research Institute. 201 Eye St., NW, Washington, DC 20005-3915 USA (a.gelli@cgiar.org)

Adamba: School of education and leadership & Institute of Statistical, Social, and Economic Research. P.O. Box LG 1181 University of Ghana, Legon, Ghana (cadamba@ug.edu.gh)

Osei-Akoto: Institute of Statistical, Social, and Economic Research. P.O BOX LG 74, University of Ghana, Legon, Ghana (ioseiak@ug.edu.gh)

Alderman: Poverty, Health and Nutrition Division, International Food Policy Research Institute. 201 Eye St., NW, Washington, DC 20005-3915 USA (h.alderman@cgiar.org)

Corresponding author: Aurino

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 Pedro Rosa Dias, Jan Ross, Franco Sassi, Jacob De Hoop, Chung Truong Hoang, and participants at seminars at IPA Conference on Education in Ghana, Italian Conference of Public Economics, and UNICEF Innocenti Research Centre for helpful comments.

The authors acknowledge funding from the Bill and Melinda Gates Foundation, the World Bank, and Dubai Cares. Elisabetta Aurino gratefully acknowledges the financial support of the Imperial College Research and Guido Cazzavillan Research Fellowships and the hospitality of Prof. Felix Asante and his team at ISSER/University of Ghana during initial analyses.

The trial has received IRB approval from the University of Ghana. This study is registered in the AEA RCT Registry and the unique identifying number is AEARCTR-0003629. The trial is also registered in the ISRCTN registry with the following identifying number ISRCTN66918874.

1 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 inequalities 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). Coherently 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 of large-scale, government-led interventions on human capital, especially in SSA. One of such interventions is school feeding, which ranks amongst the world's most common forms of social protection. Every day, about 368 million children receive some form of school feeding globally, for an estimated investment of $70 billion a year (Honorati, Gentilini, and Yemtsov 2015; WFP 2013). In SSA, since the early 2000s, a large number of countries have invested in school feeding as a multisectoral strategy involving education, health, and agriculture (Alderman and Bundy 2012; Drake et al. 2017). These programs can be relatively expensive to operate: at an average cost of US$54 and US$83 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 considerable (Gelli and Daryanani 2013). Quantifying the impact of school feeding on educational achievements is thus critical for policy makers evaluating intervention options to attain SDG4 in resource-constrained settings.

School feeding has a robust track record in increasing school participation (Alderman and Bundy 2012; Drake et al. 2017; Kristjansson et al. 2015). Yet, experimental evidence on the effectiveness of school feeding on educational achievements is more limited, and with mixed findings (see Appendix A and Snilstveit et al. [2015] for a meta-analysis). Of importance, in Africa and elsewhere, experiments have evaluated programs implemented as part of international food assistance, usually by the World Food Programme (WFP) or other international NGOs, at a relatively limited scale (Alderman, Gilligan, and Lehrer 2012; Kazianga, de Walque, and Alderman 2012, 2014; Neumann et al. 2007; Omwami, Neumann, and Bwibo 2011; Powell et al. 1998; van Stuijvenberg et al. 1999; Vermeersch and Kremer 2005). In these studies, school feeding was usually undertaken in restricted geographic contexts 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).

Government programs reaching daily large populations may suffer from a number of additional challenges as compared to smaller-scale interventions implemented by international NGOs, including market equilibrium effects and spillovers (Acemoglu 2010; Cunha, De Giorgi, and Jayachandran 2018; Filmer et al. 2018); endogenous political economy reactions (Bold et al. 2018); heterogeneity by site or implementer characteristics (Allcott and Mullainathan 2012); and implementation challenges related to scale, including poor monitoring (Linnemayr and Alderman 2011). These issues may all hamper the generalizability of evidence based on smaller-scale, internally-valid trials to "real-world" programs reaching large populations (Banerjee et al. 2016; Bold et al. 2018; Deaton and Cartwright 2018; Nandi et al. 2017; Pritchett and Sandefur 2015). Also, existing literature has seldom analyzed mechanisms underpinning treatment effects, which are instead critical to assess external validity. Analysis of heterogeneity in treatment effects between average and marginalized learner groups constitutes a further key evidence gap (Bashir et al. 2018; Evans and Yuan 2018).

We tackle these gaps by providing novel experimental evidence on the impact of the Ghana School Feeding Programme (GSFP) on child educational achievements. The GSFP provides a free, hot-cooked daily meal to over 2 million pupils in government primary schools across Ghana. In collaboration with the Government of Ghana, we conducted a randomized control trial around the re-targeting and scale-up of the GSFP to the country's most food insecure districts, which were previously not covered by the program[2]. While the overall trial aimed at assessing program impacts on education, nutrition, and agriculture[3] (see Gelli et al., 2016), here we report treatment effects on per-protocol educational attainments outcomes, including child learning (maths, literacy), cognition (working memory, problem-solving ability), and overall achievements[4]. We also assess treatment effects mechanisms and heterogeneity by per-protocol population subgroups.

Ghana represents an ideal context to study the learning effects of government school meals programs, as its 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

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

[3] The results of the analyses on child anthropometrics and community agriculture will be published separately.

[4] These are three composite scores of, respectively: maths and literacy, cognition and working memory, and of all test scores.

word or perform a two-digit subtraction (World Bank 2018). Second, wide inequalities in achievements exist by gender, poverty, and place of residence (World Bank 2018). Third, despite rapid economic growth, food insecurity and poverty are widespread, particularly in rural areas. In this context, 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, including the opportunity cost of schooling, than alternative social protection measures that target households, such as cash transfers or generalized food assistance (Aurino et al. 2018). This mechanism could be particularly effective among the most vulnerable groups of learners. 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 similar settings.

Following the methodology outlined in our study protocol (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 increases in all scores considered by about 0.1 to 0.16 standard deviations (∂, hereafter) for the average pupil. We note that these findings are likely to correspond to lower bounds of potential effects, as program take-up was imperfect and implementation challenges were present. The latter mostly related to severe delays in financial disbursements to the caterers, which are in charge of procuring food, and cooking and serving the meals. Turning to heterogeneity, we find strong variation in program impact in favor of the most marginalized groups of learners. Girls' maths, literacy, and learning composite scores increased by 0.2∂ in school feeding communities. 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 0.25∂ and 0.3∂ across all scores. The offer of school feeding led to increases in school enrolment, grade attainment, and, for the poorest children, shifts in child time use from housework toward increased time in school, which we interpret as likely impact pathways. Overall, these results underscore the social protection-*cum*-human capital accumulation of a large-scale, government-led school feeding program, which appeared particularly effective in raising learning outcomes among the most vulnerable learner subgroups in Ghana.

We believe that this paper contributes to the literature in multiple ways. To the best of our knowledge, this is the first field experiment from Africa investigating the effects of a government-led, nationwide program on a rich set of educational attainments. Therefore, it adds to an experimental literature on school feeding that, as noted, is overwhelmingly focused on smaller-scale trials based on programs financed and implemented by international actors (see Appendix A). So far, only a few studies have focused on the impact of government-run school meals programs on learning using quasi-experimental methods. In India, Chakraborty and Jayaraman (2016) exploited staggered program implementation to identify positive effects of the local "midday-meal" scheme on maths and literacy. Recently, two trials have used the infrastructure of the Indian school feeding program to scale up food fortification, with mixed effects on learning (Berry et al. 2018; Krämer, Kumar, and Vollmer 2018). In Chile, a middle-income country with little undernutrition, McEwan (2013) applied a regression discontinuity design to administrative data, without finding any impacts of the local school feeding program on fourth-grade test scores. Belot and James (2011) evaluated the educational effects of improvements in the quality of school meals in a borough of London by employing difference-in-difference. They found evidence of improved English and science scores and decreases in absenteeism due to sickness.

Additionally, we provide evidence on treatment effects heterogeneity for the most marginalized learners, which is, surprisingly, an under-investigated topic in educational research in LMICs (Bashir et al. 2018; Evans and Yuan 2018; World Bank 2018). Our findings suggest that in contexts characterized by wide educational inequalities such as Ghana, school feeding programs may be able to "level the playing field" by raising learning outcomes, especially among children at the margin (Jukes, Drake, and Bundy 2008).

More broadly, this study speaks to a body of literature employing randomized trials to evaluate the learning and cognitive impacts of educational, health, and social protection interventions, particularly in SSA (Evans and Popova 2015; Glewwe and Muralidharan 2016; De Groot et al. 2015). In the case of social protection, existing evidence is overwhelmingly focused on schooling (enrolment and attendance), rather than on learning achievements (Baird et al. 2014; De Groot et al. 2015).

This paper is organized as follows. The next section presents the background and the study design. Then, Section 3 illustrates the data and identification strategy. Sections 4 and 5, respectively, present the ITT estimates and potential mechanisms for impact. Section 6 concludes, including discussion of costs.

2 Background and Study Design

2.1School feeding and child learning: potential pathways for impact

Improved school participation, changes in time use, and enhanced health and nutrition constitute potential channels through which school feeding can affect learning (Adelman, Gilligan, and Lehrer 2009). First, school meals may have an incentive effect on enrolment and attendance by subsidising the cost of education through the provision of a free meal conditional on attendance, which may contribute to shifting household decisions toward increased schooling. In turn, increased attendance may have a positive effect on learning. Also, the school feeding transfer may free up resources that households could use to purchase additional educational inputs for improved child learning, such as books, other educational materials, or private tuition.

Second, beyond fostering attendance through subsidising the costs of education, school feeding may contribute to changing children's time use through a shift from labor (particularly in agriculture) toward increased time devoted to educational activities (Alderman, Gilligan, and Lehrer 2012; Kazianga, De Walque, and Alderman 2009). A meta-analysis has shown that increases in time spent at school may be an important pathway to better learning achievements (Snilstveit et al. 2015).

Also, by addressing short-term hunger and micronutrient deficiencies, including iron and iodine, school feeding can positively affect children's learning and cognition via reduced morbidity-related absenteeism, better micronutrient status, and increased concentration in the classroom. It may be plausible that teachers can be more motivated by interacting with more 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 (Ahmed 2004; Chakraborty and Jayaraman 2016; Jacoby 2002; Kazianga et al. 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).

Investments in children's education, and related academic attainments, can vary by child and household characteristics (e.g., uneducated parents, poorest or remote households, etc.), and by school quality (e.g., teachers availability and experience, educational supplies, and facilities, etc.). It may be the case that marginalized learner groups such as girls, lower-ability children, and children from economically-disadvantaged households and areas may benefit disproportionally more from the transfer in terms of human capital investments than less disadvantaged pupils (Akresh, De Walque, and Kazianga 2013; Björkman-Nyqvist 2013).

2.2 Educational setting and the GSFP

Ghana's basic education is compulsory between 5 and 15 years. During the 2000s, the country prioritised school participation through various initiatives, including the GSFP. These efforts resulted in a substantial expansion of basic education, with primary enrolment moving 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). As per educational access, wide inequalities in achievements exist by gender, poverty, and northern regions (World Bank 2018).

The Government of Ghana initiated the GSFP 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).

2.3 Evaluation design

The trial was designed around the scale-up of the GSPF to new intervention districts, based on the 2012 retargeting exercise. The retargeting was guided by the development of poverty and food insecurity rankings to assess potential priority districts in which the GSFP could be rolled-out. Rankings were used to generate district-level indices on the share of national poverty and food insecurity, through which 58 priority districts were identified for the scale-up of GSFP (see Gelli et al., [2016], for details).

The trial focused on assessing program effects on both child-level (education, health) 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 hyphotesised to affect mostly child education and health) was designed to be delivered at the school-level, while the agriculture-related acivities 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). First, districts identified from the retargeting exercise were randomly assigned to two treatment arms: the standard GSFP and a "home-grown" school feeding (HGSF) pilot including an integrated package of activities aimed at enhancing the impact of the GSFP on community agriculture and food security by sourcing the produce for the meals directly from the local farmers (Gelli et al. 2016). At the second stage, two schools (and relative school-catchment areas, which we refer to as "communities" therein) were randomly assigned within each district to a school feeding or control arm. A protocol was designed in order to ensure that the sampled schools were comparable based on Education Management Information System (EMIS) data (for details, see Gelli et al. [2016]). Following a household census at baseline, approximately twenty-five households with children in the 5-15 target-age group were randomly selected for interview from each community. Control communities were planned 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[5].

[FIGURE 1 ABOUT HERE]

3 Data and Sample Description

3.1 Timeline and Sample

A baseline survey was undertaken in 93 communities between June and September 2013. The follow-up survey was conducted in February-March 2016. Implementation in treatment communities started in the academic year 2014/15, due to burocratic delays.

[5] The original design included a three-level randomization, with a subset of HGSF schools receiving micronutrients constituting a fourth study arm (Gelli et al. 2016). However, the discovery of GSFP already running in 24 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.

Given that the academic year in Ghana usually runs from August to May, the program was evaluated after roughly two academic years of implementation.

Two communities were excluded from the endline survey, due to logistical problems related to episodes of local insecurity. Both rounds of household surveys included detailed modules on household demographics, farm and other assets, expenditures, farming and other economic activities, child anthropometry, and self-reported[6] education indicators for all target-age children in the household, including child enrolment, attendance and grade attainment, and educational achievements. At endline, data on children's time use were also collected. 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. At endline, 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.

3.2 Balance of Baseline Covariates and Attrition at Endline

Table 1 presents descriptive statistics of baseline characteristics of the longitudinal 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

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

difference between the two groups that was statistically significant at 10 percent was age of household heads in the school feeding arm, which were about one-year-and-a-half older than control communities. A similar picture emerged from the balance analysis focusing only on the baseline sample prior to attrition (Appendix B). 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 attrition tests at the child level. We do not observe any unbalance in the probability of remaining in the longitudinal sample based on offer of school feeding, which would have otherwise introduced some bias in the estimation of the treatment effects, affecting internal validity[7]. Column 2 presents analysis of whether baseline test scores were associated with children being more likely to be resurveyed, which did not appear to be the case either. Finally, column 3 investigates whether the treatment was associated with some child characteristics in predicting odds of remaining in the sample, which may introduce some sample composition bias. We did so by interacting the treatment with some of the key characteristics we use for heterogeneity analysis throughtout the paper. We find these interactions were jointly significant at the 5 percent level, mostly due to the fact that children in control areas in northern regions were slightly more likely to be reinterviewed at endline (96 percent of baseline children were followed up in control communities, viz. 93 percent in treatment areas, translating in, respectively, for a total of 36 additional children lost in treatment areas compared to control). To evaluate further the possible effects of attrition on validity of the impact estimates, Appendix C presents the balance of characteristics across treatment groups at endline for the full sample, and for the northern regions only. Across a wide range of child and

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

household backgrounds, there were no differences between school feeding and control arms for the longitudinal sample at endline. Thus, even if there was some evidence of differential attrition by treatment in the case of northern regions, balance was generally maintened, particularly in light of the relatively low levels of attrition overall and especially in the northern regions, lessening concerns of a change in the sampling frame by treatment assignment due to attrition.

[TABLE 2 ABOUT HERE]

3.3 Program Uptake and Implementation

Only 61 percent of children in school feeding areas reported receiving school meals in the previous week at endline, which we refer to as the overall uptake rate. By restricting the sample of children that were enrolled in primary government schools (as opposed to private schools) in treatment communities, the uptake rate was 83 percent, indicating that most children that were still in basic education did in fact receive school meals. We remind that school meals are not served in secondary schools. 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.

Appendix D presents correlates of endline program uptake (independent of primary enrolment status) among children in treatment communities. Coherently with expectations of older children having progressed to secondary school or being out of school at endline and thus being less likely to receive school feeding, children aged 5-11 years at baseline were two times more likely to receive school feeding compared to adolescents (12-15 years at baseline). 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 maths and literacy scores were 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.

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 to bring their food from the school meal to share at home. There were no differences between treatment and control groups in the mean number of school-days in which the child had breakfast (four out of 5 five days)[8].

Data collected from school caterers (N=55) highlighted that 86 percent of them experienced irregular payments, and about a third of them had not received any payment in the three months before the survey (available upon request). No regional differences were evident. Nearly 85 percent of caterers also indicated that often payments are 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).

3.4Measures of Child Learning and Cognition

Child maths, literacy, SPM and digit span are the primary study outcomes per protocol. The same 15-item maths and literacy tests, and 12-item SPM and digit span

[8] Fernandes et al (2017) reports detailed findings from focus groups on children's dietary behaviors in the context of this evaluation. Also, a companion paper related to impact findings on nutrition will explore these issues in more depth.

tests were administered across rounds. Tests were administered at home to ensure that even children out of school were tested, enhancing internal validity. Given the wide age range included in the target sample, the tests aimed to assess a basic set of skills in literacy, maths. and in cognitive development domains. Each section of the test began with basic domain-specific questions that progressively increased in difficulty in order to cover different ability levels, particularly in the light of the diverse age range of eligible children. The maths 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 standardized progressive matrices test (SPM) test was an adaptation of the Raven's progressive matrices test, a commonly-used measure of nonverbal fluid intelligence and problem-solving ability, while the digit span test focuses on assessing working memory. 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. For the digit span test, 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).

Test scores were standardised 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[9]. Later, the

[9] This approach also helped controlling for potential language effects, as unfortunately we do not have information on the specific language of administration of the test, as each interviewer spoke a different local language.

residuals from these regressions were non-parametrically estimated to obtain age-conditional means and standard deviations.

We also generated composite indicators of learning, cognition and across all the outcomes considered to address potential issues related to multiple testing[10]. The summary indices should enhance statistical power to detect effects that go in the same direction (Kling, Liebman, and Katz 2006). We computed those as an average from the normalized test scores, and then standardized again to the control group within each round[11]. In this way, estimated ITT effects can be interpreted as the effect size relative to the control group (Banerjee et al. 2015).

3.5Descriptive Statistics of Learning and Cognition

Table 3 presents descriptive statistics of raw tests scores in the four compentencies 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[12].

The analysis of the raw scores highlights the low achievement levels in each outcome and survey round: at baseline, on average, children were not able to respond to at least two out of 15 questions in the Maths and Literacy tests. This proportion increased

[10] This was not included in our original analysis plan. However, this is a common approach in the literature to create more general conclusions about the impact of a program on a family of outcomes (Banerjee et al. 2007; Kling, Liebman, and Katz 2006; Wydick et al. 2016). Wydick et al. (2016) note that this approach "help addressing the issue of over-testing that could erroneously assign too much importance to a possibly spurious rejection of a single null hypothesis for one variable within a family of outcomes"(p. 18).

[11] Although children were given assessments in all tests, discrepancies in sample sizes across raw and standardised scores reflect inability to convert raw scores into standardized scores (e.g. lack of child age in months). A similar issue is highlighted in (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).

[12] A similar picture emerged from the analysis of baseline differences in raw scores for the baseline sample prior to attrition presented in Supplementary Table A1. This provides a further reassurance about potential biases in treatment effects of school feeding on child learning stemming from nonrandom attrition.

slightly three years later. However, endline scores were still very low, with the average pupil only being able to respond to about 4 out of 15 correct questions for maths and literacy, which reflects Ghana's learning challenges.

By contrast, the mean levels of Digit Span achievements did not change over time, and for SPM they decreased between baseline and endline. This decline was common across all groups, but was wider in the case of the oldest children that were not enrolled at endline (12-15 years at baseline). This may suggest some deterioriation of cognitive skills after school leaving, such as recently reported by Soler-Hampesjsek and coathors (2018).

[**TABLE 3 ABOUT HERE**]

Appendix E presents raw scores at both rounds by child gender, household poverty status and by residence (south vs. north Ghana). At both rounds, there were no strong differences between girls and boys, while gaps between nonpoor and poor children were evident across all outcomes. The largest disparities in baseline 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 all competences. This gap substantially reduced or closed at endline.

Figure 2 presents the non-parametric distributions of raw scores in maths and literacy at both rounds by treatment arm. There were important floor effects for low levels of achievements in both maths and literacy, whereby a significant proportion of children did not respond correctly to any question, indicating the tests were too difficult for them. This was especially marked among younger children at baseline. 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). Given the relatively-low proportion of children that were able to respond correctly to all questions, ceiling effects were not a concern. Also, there was an increase in the variance of learning achievements between baseline and endline. This may reflect widening of educational inequalities in the transition from primary to higher levels of education. At the end of primary school, the most vulnerable children tend to enter the labor market, thus widening pre-existing educational disparities (De Groot et al. 2015). The distribution of 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 maths and literacy. The marked gains for above-average levels of achievements among the school feeding group at endline is clearly observable in Supplementary Figure F1, which plots age-standardized empirical distributions of maths and literacy by treatment arm and round.

[FIGURE 2 ABOUT HERE]

Figure 3 presents the non-parametric distributions of raw test scores in Digit Span and SPM scores. For those outcomes, floor effects were less pronounced than in the case of maths and literacy and their distribution was more evenly distributed across the full range of cognitive abilities. This is likely due to the fact that both tests measured cognitive abilities, not schooling achievement (see, Supplementary Figure F2, and Appendix G for disaggregation of test scores by gender, place of residence, and poverty).

[FIGURE 3 ABOUT HERE]

Finally, we note that the autocorrelations of test scores between baseline and follow-up were low by the standards usually adopted in the literature[13] (maths: ρ =0.23; literacy: ρ =0.31; digit: ρ =0.19; raven: ρ =0.13, all significant at <0.01). This finding

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

may be partially explained by some degree of measurement error, and partly by the three-year lag between the assessments.

3.6 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 different 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 both the post-estimator and the DiD approach. Gains in efficiency are particularly marked when outcomes have low autocorrelation, as in our case. In econometric terms, we estimate the following model:

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

Where y_{it}, j and $y_{i(t-1),j}$ represent, respectively, the endline and baseline test scores (when available)[14] for child *i* residing in community *j*; $SF_{it,j}$ is a dichotomous variable for

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

a child residing in a community randomly assigned to school feeding and thus uncorrelated with $y_{i(t-1),j}$; and θ_j is a vector of region dummies to capture region-specific unobservable characteristics or potential variation in quality of implementation. Standard errors were clustered at the community level, which is the unit of randomisation for school feeding. β_1, the coefficient related to school feeding, provides the estimate of the treatment effects. For completeness with the analysis plan, we also provide DiD coefficients in Appendix H.

4 Impact of School Feeding on Educational Attainments

Table 4, Panel A presents ITT estimates for the full sample. Overall, the randomized offer of school feeding led to significant increases across all test scores of about 0.12∂ to $0.15\ \partial$[15]. We then investigate heterogeneity in program effects. Table 4, Panels B, C, and D, respectively, report treatment effects in models that stratify for gender, poverty, and northern regions. Across the three subgroups, school feeding led to sizeable gains in achievements. In the case of girls, literacy and digit span scores increased by 0.2∂ and maths by 0.24∂ (Panel B). The composite scores also increased by more than 0.2∂. Even larger gains were evident in the case of children from households below the poverty line at baseline (Panel C) and children from northern Ghana (Panel D). For the former, gains in maths scores and in the three composite scores amounted to a third of standard deviation, while the increases in literacy, SPM, and Digit Span were slightly smaller, but still accounting to more than 0.2∂. A similar picture emerges from the findings related to northern regions children.

[15] While the treatment effects arising from both ANCOVA and DiD are in most cases remarkably similar, as anticipated, the former estimator proved being more efficient compared to DiD (Appendix G).

[TABLE 4 ABOUT HERE]

We also investigated variation in treatment effects by age in Appendix I. The latter shows that the effect of school feeding was mostly similar, with the exception of maths, 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 Appendix J. No substantial differences in treatment effects on educational achievements between the GSFP and HGSF were detectable.

5 Mechanisms

This section investigates the role of changes in school participation and child time use as potential mechanisms explaining program effects[16]. Descriptive statistics of all intermediate outcomes are presented in Appendix K.

5.1Changes in Schooling

We start by changes in schooling following the intervention, as increases in school enrolment and attendance constitute critical pathways for learning. Also, if school feeding was successful in increasing learning, we should be able to observe higher grade attainments and lower grade repetition rates (Alderman, Gilligan, and Lehrer 2012). Table 5 presents ITT estimates of school feeding on the following indicators: child school enrolment in any educational level; school enrolment in primary school; school attendance (conditional on enrolment), as measured by the number of days the child attended school out of a five-day week; current grade attended by the child; and a dichotomous variable related to whether a child has ever repeated a grade. All of these variables were measured

[16] Detailed effects on child anthropometrics are reported in a companion paper.

in the household survey with questions directed to the child or her caregiver in both survey rounds.

Panel A reports estimates of school feeding for the full sample[17]. Given the observed heterogeneity in treatment effects by specific population subgroups, Panel B, C, and D report ITT effects for girls, children from poor households, and children from northern regions, respectively. Increases in school enrolment emerge as an important channel for impact, but only for the most disadvantaged groups. This finding is expected in contexts such as Ghana, where primary enrolment is almost universal, and the largest gains in school access are now to be among groups that are more likely to be out of school. Consistently with the results on learning and increases in enrolment, there were increases in grade attainment. Estimates were positive for all groups, but they were only statistically significant for the average child and for children from the northern regions, indicating that children in school feeding areas completed more years of schooling than control. Consistently with the results on grade, there was no evidence of increased repetition in the school feeding arm. This result reassures about potential concerns of children not progressing to secondary school in order to continue receiving free meals, as observed in other SSA settings (Kazianga, De Walque, and Alderman 2012).

[TABLE 5 HERE]

5.2 Changes in Time Use

Table 6 provides ITT effects on child time use for the full sample (Panel A), stratified by child gender, household poverty, and northern regions (Panels B, C, and D, respectively). The indicators measured the average time (in hours), on a typical day, a child spends in the following activities: at school; studying; doing housework; doing farm

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

work or other types of labor; and in leisure. These questions were only collected at endline, therefore treatment effects were estimated through a post-estimator.

While there were no significant changes in time use for the average child, there were increases in the time spent in school for all vulnerable subgroups. The increase in time spent in school ranged from about 30 additional minutes per day for girls, up to 50 additional minutes per day for children from poor households. Given that the average serving time for school feeding is about 30 minutes, it is plausible that the additional net time spent in school in the case of poor children is employed in instructional activities, and thus constitute a pathway to increased test scores. Qualitative evidence from field visits and focus group discussions conducted as part of this evaluation highlighted that children in school feeding communities tended to stay at school in the afternoon, while otherwise they would have not returned to school after having lunch at home (Fernandes et al. 2017). Also, teachers reported extending the instructional time by about 45 minutes in the afternoon to accommodate the time lost earlier in the day due to the serving and eating of the meals. Similar increases in time spent in school during the afternoon shift due to school feeding were observed in Northern Uganda (Alderman, Gilligan, and Lehrer 2012).

For children from the household below the poverty line, there was also a significant decrease in time spent in housework, with a decrease of about 20 minutes per day. We speculate that the indirect income transfer from school feeding has led poorest households, which usually tend to rely more on children for household chores and care, to forego part of the benefits stemming from children's housework in favor of additional schooling time. An analogous finding was reported for rural Burkina Faso, where school feeding led to shifts from farm labor to housework, which may be more compatible with school attendance (Kazianga, De Walque, and Alderman 2012).

[TABLE 6 HERE]

6 Concluding Remarks

Most governments globally offer some form of food at school as a large-scale strategy to enhance children's education and health. Given the pervasiveness of school meals programs worldwide, and size of global educational investments in such intervention, understanding whether large-scale, government-led school feeding is effective in raising achievements, and whether it works for marginalized groups of learners, is a critical policy question and evidence gap.

We report treatment effects estimates from a randomized control trial focusing on a program reaching daily two million children across the poorest districts of Ghana. Program impacts were evaluated after two academic years of implementation in randomized communities across learning and cognitive attainments. We show that the offer of school feeding in randomized communities led to moderate gains in learning and cognitive levels for the average pupil. The magnitude of the effects (about 0.15∂) was comparable to estimates from a recent meta-analysis of school feeding (Snilstveit et al. 2015). Beyond average effects, the program had remarkable impact among the groups that are more vulnerable to poor educational 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 and cognition – ranging between 0.2∂ to 0.3∂. All estimates are likely to represent lower bounds due to imperfect uptake. The positive effects on learning and cognition are especially remarkable when contextualized to the implementation challenges related to the delayed disbursements to the caterers. Regularity in the provision of school meals and their quality 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 type or not attending at all). Overall, the noted challenges add to the generalizability of our findings to "real-world" large-scale school feeding interventions implemented by governments in similar contexts, which may also face financial, implementation, and monitoring constraints.

To the best of our knowledge, this is the first study providing experimental evidence on a government's school feeding program implemented nationwide in a LMIC, with an investigation of the effects of the intervention for most vulnerable learners and impact mechanisms. By subsiding the cost of schooling and providing an indirect social protection transfer to the household, school feeding positively affected educational achievements through increased enrolment rates for all children, with larger impacts for children of poorest families and regions, increases in time spent at school (presumably beyond the meal serving time among the poorest children), and increases in grade attainments. In SSA, reductions in the cost of education may be particularly effective in raising the educational outcomes of vulnerable population segments (Aurino et al. 2018; Björkman-Nyqvist 2013; Kazianga, de Walque, and Alderman 2012). Further, findings from the companion nutrition paper highlight similar pro-poor gains, which highlights the greater effectiveness of the program among the most disadvantaged groups of children (Gelli et al. 2018). Contextualized to Ghana, these findings address concerns raised in a 2010 Ministerial review of GSFP, according to which program investments targeted disproportionally better-off districts (Balwanz and Darvas 2013; WFP 2013). Based on our results, it appears that the retargeting of the program to the country's most food-insecure districts had been able to address these concerns of regressivity.

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 (Gelli et al. 2014). Also, even if we would decide to restrict the focus on educational achievements as single measure of outcomes, this exercise may be still unsatisfactory, as the life course and intergenerational effects of gains from increased schooling achievements are not yet fully known[18]. 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[19]. This figure falls within the range of the average cost per child of school meals in LMICs: the latest data available highlight that in 2008, school feeding cost US$54 and US$83 in low- and middle-income countries, respectively (Gelli and Darayani 2013). Taking inflation into account[20], the GSFP compares well with other programs in LMICs in terms of costs. Also, it is worth noting that most available estimations of program costs are 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

[18] For instance, in the context of Ghana, Duflo et al. (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 maths and literacy, had better health behaviors, and girls had less children.

[19] 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).

[20] 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.

accumulation, particularly for marginalized groups of learners. 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 can be an important policy 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.

Adelman, Sarah, Daniel O. Gilligan, and Kim Lehrer. 2009. "How Effective Are Food-for-Education Programs? A Critical Reassessment." in *The Poorest and the Hungry*, edited by J. von Braun, R. Vargas Hill, and R. Pandya-Lorch. Washington, DC: International Food Policy Research Institute.

Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2013. "School Meals and Classroom Effort: Evidence from India School Meals and Classroom Effort: Evidence from India."

Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2016. "The Mixture as Before? Student Responses to the Changing Content of School Meals in India." (9924).

Ahmed, Akhter U. 2004. "Impact of Feeding Children in School: Evidence from Bangladesh." *International Food Policy Research Institute* 2033:20006–1002.

Akresh, Richard, D. De Walque, and Harounan Kazianga. 2013. *Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality*. The World Bank.

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.

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.

Allcott, Hunt and Sendhil Mullainathan. 2012. “External Validity and Partner Selection Bias.” *NBER Working Paper 18373* 1–44.

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.* Florence.

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.

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.

Banerjee, Abhijit et al. 2015. “A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries.” *Science* 348(6236).

Banerjee, Abhijit V. et al. 2016. *From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application.*

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.

Bashir, Sajitha, Marlaine Lockheed, Elizabeth Ninan, and Jee-Peng Tan. 2018. *Facing*

Forward: Schooling for Learning in Africa. Washington, DC: World Bank.

Belot, Michèle and Jonathan James. 2011. "Healthy School Meals and Educational Outcomes." *Journal of Health Economics* 30(3):489–504.

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

Björkman-Nyqvist, Martina. 2013. "Income Shocks and Gender Gaps in Education: Evidence from Uganda." *Journal of Development Economics* 105:237–53.

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* 168:1–20.

Chakraborty, Tanika and Rajshri Jayaraman. 2016. "School Feeding and Learning Achievement: Evidence from India's Midday Meal Program." (10086).

Cunha, Jesse M., Giacomo De Giorgi, and Seema Jayachandran. 2018. "The Price Effects of Cash Versus In-Kind Transfers." *The Review of Economic Studies*.

Das, Jishnu et al. 2013. "School Inputs, Household Substitution, and Test Score." *American Economic Journal: Applied Economics* 5(2):29–57.

Deaton, Angus and Nancy Cartwright. 2018. "Understanding and Misunderstanding Randomized Controlled Trials." *Social Science and Medicine*.

Drake, Lesley et al. 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.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2017. *The Impact of Free Secondary Education : Experimental Evidence from Ghana*. Mimeo.

Evans, D. K. and F. Yuan. 2018. "Commentary." Pp. 232–34 in *Learning at the bottom* *of the pyramid.*, edited by D. A. Wagner, S. Wolf, and R. F. Boruch. Paris: UNESCO.

Evans, David and Anna Popova. 2015. *What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews.*

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, 1–18.

Filmer, Deon P., Jed Friedman, Eeshani Kandpal, and Junko Onishi. 2018. *General Equilibrium Effects of Targeted Cash Transfers: Nutrition Impacts on Non-Beneficiary Children.* Washington, D.C.

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.

Gelli, A., Aurino, E., Folson, G. et al. 2018. "The national school meals programme in Ghana improves height-for-age during mid-childhood, in girls and in children from poor households: A cluster randomised trial."*mimeo.*

Gelli, A. 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).

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." (February):18.

Ghana Institute of Management and Public Administration (GIMPA)/ U. of California at Berkeley Haas School of Business (2011). 2011. *Ghana School Feeding Program: Re-Tooling for a Sustainable Future*. London.

Glewwe, P. and K. Muralidharan. 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications." Pp. 653–743 in *Handbook of the Economics of Education*, vol. 5. Elsevier.

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.

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.

De Groot, Richard et al. 2015. *Heterogeneous Impacts of an Unconditional CashTransfer Programme on Schooling: Evidence from the Ghana LEAP Programme.*

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

Honorati, Maddalena, Hugo Gentilini, and Rusian G. Yemtsov. 2015. *The State of Social Safety Nets 2015.* Washington, D. C.

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.

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.

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* 106:15–34.

Kazianga, Harounan, Damien De Walque, and Harold Alderman. 2009. "Educational and Health Impacts of Two School Feeding Schemes Evidence from a Randomized Trial in Rural Burkina Faso." *World Bank Policy …* 4976(30):46–52.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2006. "Experimental Analysis of Neighborhood Effects." *Econometrica* (June):1–34.

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

Kristjansson, Elizabeth 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* (3).

Linnemayr, Sebastian and Harold Alderman. 2011. "Almost Random: Evaluating a Large-Scale Randomized Nutrition Program in the Presence of Crossover." *Journal of Development Economics* 96(1):106–14.

Mcewan, Patrick. 2013. "The Impact of Chile ' s School Feeding Program on Education Outcomes." *Economics of Education Review* 32:122–39.

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

Ministry of Education & RTI International. 2014. *Ghana 2013 National Education*

Assessment - Technical Report | Global Reading Network.

Nandi, Arindam, Jere R. Berhman, S. Bhalotra, Anil B. Deolalikar, and R. Laxminarayan. 2017. "The Human Capital and Productivity Benefits of Early Childhood Nutritional Interventions." in *Disease Control Priorities 3*. Washington, D. C.: The World Bank.

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.

Omwami, Edith Mukudi, Charlotte Neumann, and Nimrod O. Bwibo. 2011. "Effects of a School Feeding Intervention on School Attendance Rates among Elementary Schoolchildren in Rural Kenya." *Nutrition* 27(2):188–93.

Powell, C. A., S. P. Walker, S. M. Chang, and S. M. Grantham-McGregor. 1998. "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.

Pritchett, Lant and Justin Sandefur. 2015. "Learning from Experiments When Context Matters." in *American Economic Review*.

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

Snilstveit, Birte et al. 2015. *Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review*. London.

Soler-Hampejsek, Erica et al. 2018. "Reading and Numeracy Skills after School Leaving in Southern Malawi: A Longitudinal Analysis." *International Journal of Educational Development* 59:86–99.

van Stuijvenberg, M. E. et al. 1999. "Effect of Iron-, Iodine-, and Beta-Carotene-

Fortified Biscuits on the Micronutrient Status of Primary School Children: A Randomized Controlled Trial.” *The American Journal of Clinical Nutrition* 69(3):497–503.

UNDP Ghana. 2015. *2015 Ghana Millennium Development Goals Report.* Accra.

Vermeersch, Christel and Michael Kremer. 2005. *School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation.* Washington, D. C.

WFP. 2013. *State of School Feeding Worldwide 2013.* Rome.

World Bank. 2017. “World DataBank: Education Statistics - All Indicators.” Retrieved October 17, 2017 (http://databank.worldbank.org/data/reports.aspx?source=education-statistics-~-all-indicators#).

World Bank. 2018. *World Development Report 2018: Learning to Realize Education’s Promise*. Washington, D.C.

Wydick, Bruce et al. 2016. “Shoeing the Children: The Impact of the TOMS Shoe Donation Program in Rural El Salvador.”

Tables and Figures

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)
School feeding	0.798	0.792	1.638
	(0.201)	(0.206)	(0.983)
Age-standardized Maths score		0.949	
		(0.144)	
Age-standardized literacy score		1.134	
		(0.184)	
Age-standardized Raven score		1.011	
		(0.095)	
Age-standardized Digit span score		1.094	
		(0.111)	
Child aged 12-15 years *School feeding			0.980
			(0.190)
Child aged 5-11 years * Control			0.849
			(0.202)
Female * School feeding			0.916
			(0.160)
Male * Treatment			1.354
			(0.253)
Southern regions * School feeding			0.675
			(0.291)
Northern regions * Control			2.400***
			(0.798)
Above poverty line * School feeding			0.697
			(0.254)
Below poverty line * Control			0.594
			(0.204)
Constant	13.655***	13.781***	11.619***
	(2.084)	(2.234)	(2.900)
Prob > Wald Chi Squared test	0.3653	0.6064	0.0437
Obs.	3,433	3,132	3,432

*Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents odd ratios of remaining in the longitudinal sample estimated through logistic regression models, with standard errors clustered at the community level. All models include probability of Wald Chi squared test of joint significance of all repressors. N= 3,433 children of target-age prior to attrition. Lower sample sizes reflect covariates that are missing or not applicable. Column 1 shows odd ratios 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 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)**
Maths raw score	1.57	1.68	0.073	3.62	4.03
	(2.04)	(2.07)	(0.146)	(3.37)	(3.43)
Literacy raw score	1.81	1.97	0.106	3.87	4.33
	(2.43)	(2.56)	(0.211)	(3.51)	(3.57)
SPM raw score	3.82	3.96	0.073	2.99	3.24
	(2.61)	(2.57)	(0.225)	(1.98)	(2.02)
Digit-span raw score	4.12	4.35	0.157	4.17	4.37
	(2.23)	(2.21)	(0.184)	(2.49)	(2.51)

[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. SPM stands for standardised progressive matrices test. 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 northern regions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Maths	Literacy	SPM	Digit span	Composite: maths and literacy	Composite: SPM and digit span	Composite: all outcomes
	Panel A: All children						
School feeding	0.147*	0.132*	0.129**	0.119*	0.156*	0.144**	0.154**
	(0.076)	(0.073)	(0.056)	(0.065)	(0.082)	(0.061)	(0.070)
Observations	2,278	2,274	2,307	2,305	2,288	2,312	2,282
R-squared	0.068	0.130	0.034	0.050	0.132	0.064	0.118
	Panel B: Girls						
School feeding	0.242***	0.205**	0.116	0.190**	0.273***	0.175**	0.233***
	(0.082)	(0.082)	(0.076)	(0.077)	(0.093)	(0.077)	(0.088)
Observations	1,071	1,067	1,086	1,085	1,085	1,091	1,092
R-squared	0.089	0.137	0.043	0.057	0.150	0.071	0.138
	Panel C: Poor Households						
School feeding	0.309***	0.233**	0.234***	0.269***	0.328***	0.293***	0.334***
	(0.105)	(0.098)	(0.067)	(0.081)	(0.107)	(0.074)	(0.088)
Observations	539	537	537	540	542	542	543
R-squared	0.090	0.089	0.071	0.096	0.123	0.112	0.140
	Panel D: Northern regions						
School feeding	0.253*	0.243**	0.212**	0.253***	0.297**	0.272***	0.302**
	(0.127)	(0.114)	(0.085)	(0.085)	(0.135)	(0.085)	(0.112)
Observations	1,083	1,087	1,096	1,093	1,096	1,099	1,100
R-squared	0.043	0.098	0.028	0.043	0.090	0.055	0.094

*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and northern regions. Models were estimated through OLS and standard errors were clustered at the community level. For each outcome, the model controls for the baseline value of the outcome, a dichotomous variable related to the randomized assignment to school feeding, and region dummies. Maths, literacy, standardized progressive matrices (SPM) and digit span scores are age-standardized. Composite indices were computed as averages of the standardized 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 indicators, full sample, and heterogeneity by child gender, household poverty, and northern regions

	(1)	(2)	(3)	(4)
	Enrolled	**Attendance**	**Grade attained**	**Repeated a grade**
	Panel A: All children			
School feeding	0.027	0.044	0.145*	-0.02
	(0.021)	(0.044)	(0.077)	(0.024)
Observations	2,371	2,109	2,254	2,059
R-squared	0.03	0.038	0.671	0.095
	Panel B: Girls			
School feeding	0.042*	0.041	0.031	-0.02
	(0.023)	(0.061)	(0.097)	(0.028)
Observations	1,097	988	1,056	962
R-squared	0.043	0.051	0.652	0.089
	Panel C: Poor Households			
School feeding	0.053**	0.019	0.096	-0.005
	(0.025)	(0.098)	(0.121)	(0.048)
Observations	551	489	524	476
R-squared	0.025	0.076	0.666	0.108
	Panel D: Northern regions			
School feeding	0.076**	0.072	0.223*	-0.005
	(0.029)	(0.049)	(0.131)	(0.031)
Observations	1,092	992	1,030	977
R-squared	0.051	0.005	0.639	0.047

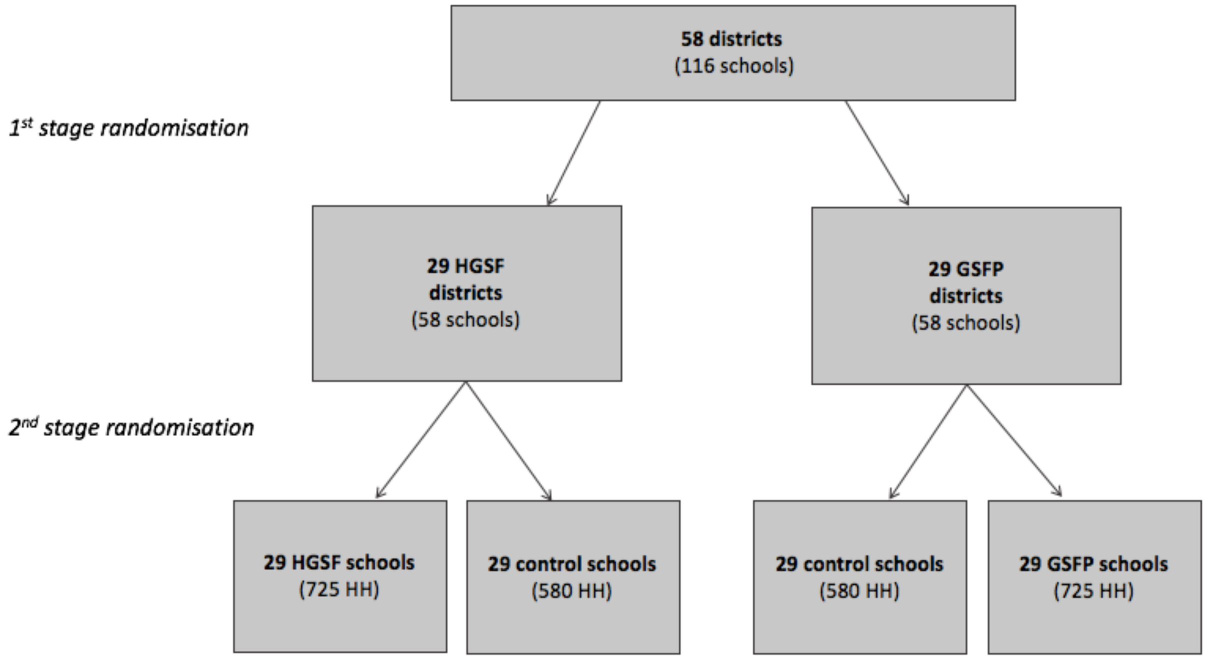
*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and northern regions. Models were estimated through OLS and standard errors were clustered at the community level. For each outcome, the model controls for the baseline value of the outcome, a dichotomous variable related to the randomized assignment to school feeding, a dichotomous variable for missing baseline outcome, 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. Grade repetition is a dichotomous variable of whether the child has ever repeated a grade. Household poverty and northern regions are defined in note below Table 5.*

Table 6. Treatment effects of school feeding on child time use

	(1)	(2)	(3)	(4)	(5)
	Time at school	Time at home studying	Leisure time	Time in household chores/care	Time in paid work/farm work
	Panel A: All children				
School feeding	0.359	0.018	0.044	-0.119	-0.153
	(0.264)	(0.145)	(0.080)	(0.089)	(0.104)
Observations	2,529	2,529	2,529	2,529	2,529
R-squared	0.097	0.088	0.164	0.041	0.054
	Panel B: Girls				
School feeding	0.488**	0.010	0.092	-0.141	-0.130
	(0.225)	(0.113)	(0.092)	(0.132)	(0.101)
Observations	1,173	1,173	1,173	1,173	1,173
R-squared	0.152	0.099	0.176	0.031	0.060
	Panel C: Poor Households				
School feeding	0.885**	0.188	0.033	-0.279*	-0.196
	(0.438)	(0.258)	(0.165)	(0.168)	(0.185)
Observations	591	591	591	591	591
R-squared	0.065	0.105	0.116	0.112	0.065
	Panel D: Northern regions				
School feeding	0.590**	0.063	0.030	-0.120	-0.180
	(0.288)	(0.064)	(0.099)	(0.116)	(0.159)
Observations	1,203	1,203	1,203	1,203	1,203
R-squared	0.212	0.015	0.219	0.048	0.065

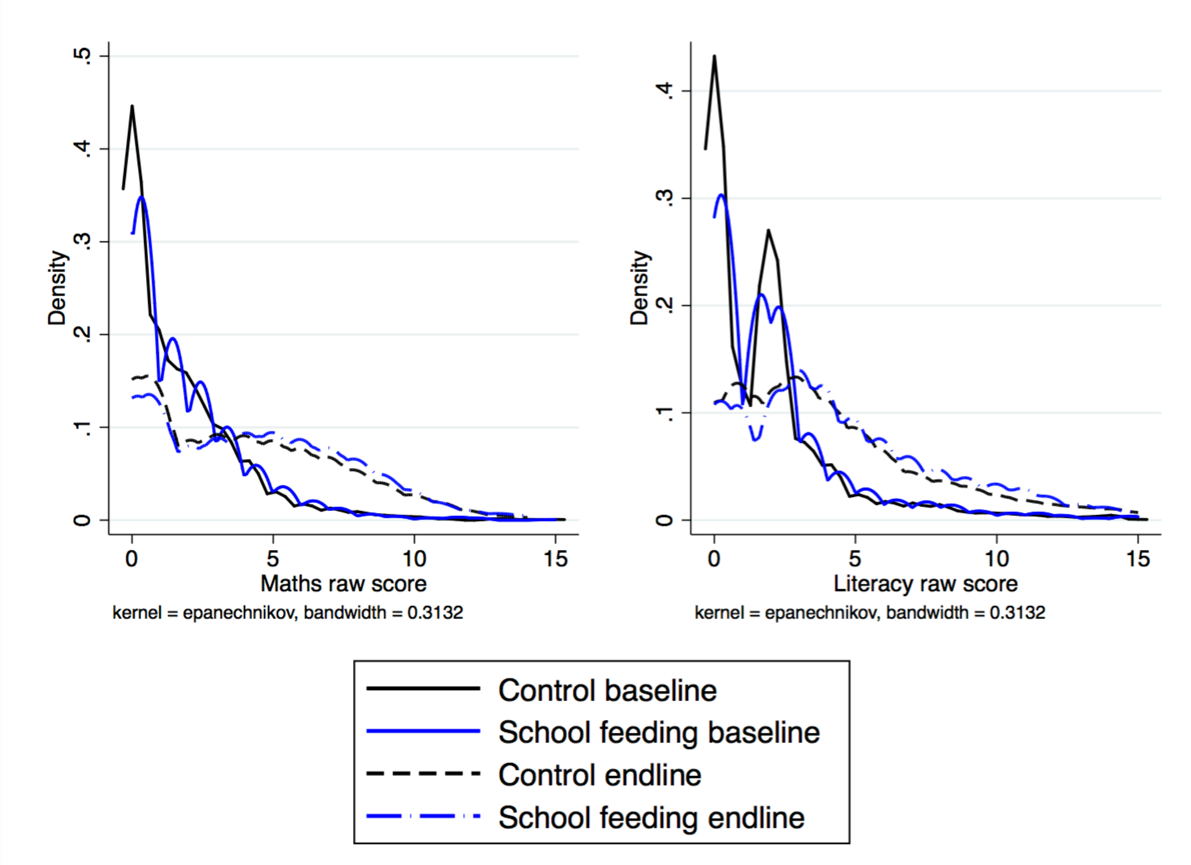
*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table above presents intent-to-treat effects on each outcome for the full sample and stratified by child gender, household poverty, and northern regions. Models were estimated through OLS with region dummies and standard errors were clustered at the community level. Time spent in school is an indicator that measures the average time (in hours) a child spends in school on a typical week day (in the school term); time spent in study is an indicator that measures the average time (in hours) a child spends studying at home or taking private tuition on a typical day; time spent in household chores is an indicator that measures the average time (in hours) a child spends doing household chores or care for other household members on a typical day; time in work is an indicator that measures the amount of time a child spends working in agriculture work, in the family business, or in another work activity on a typical day; leisure time measures the average amount of play and other leisure activities on a typical day. Household poverty and northern regions are defined in note below Table 5.*

Figure 1. Two-level randomisation



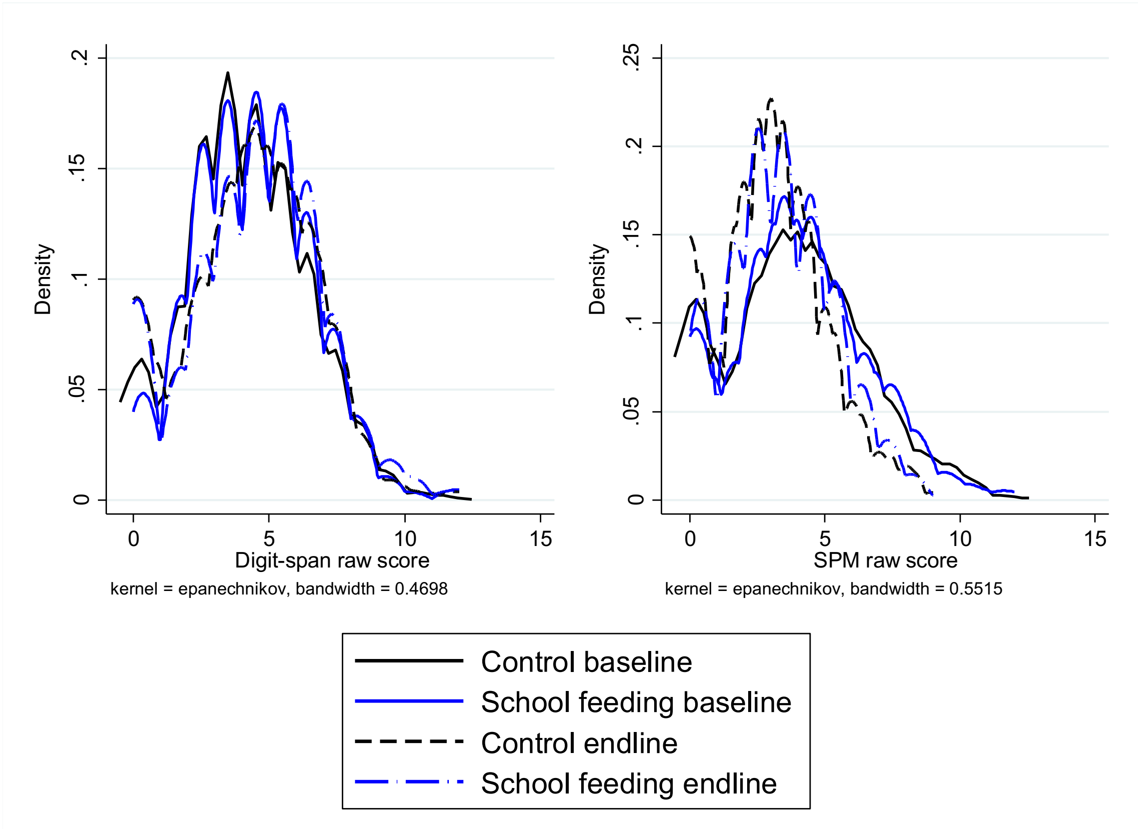
Notes: this figure provides the study design: first, districts were randomly assigned to pilot (HGSF) and standard (GSFP) school feeding through a first-level randomization; second, within each district, two schools (and related households living within the school catchment areas, which were refer to 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. Empirical distributions of raw 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 maths (left-side panel) and literacy (right-side panel) raw scores for the full longitudinal sample of children. Non-parametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel. These data are from the household survey, and therefore do not suffer from selection issues related to school enrolment. Floor effects were present, particularly in the baseline data; highlighting the tests were too difficult, particularly for the younger children. Second, there was an improvement in mean achievement in both competences between baseline and endline, although achievements were widely dispersed across the sample. Third, the distribution of achievements, especially in literacy, 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 achievements.

Figure 3. Empirical distributions of raw test scores for Digit Span (left-side panel) and SPM (right-side panel), by survey round and treatment arm



Notes: this figure presents, by treatment group and survey wave, the non-parametric distributions of digit span (left-side panel) and standardized progressive matrices (SPM) (right-side panel) raw scores for the full longitudinal sample of children. Non-parametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel. Some floor effects were present, particularly in the baseline data. Second, a decrease in SPM test scores was evident for both treatment and control groups at endline. This was mostly driven by the older cohort of children.

Supplementary materials

Appendix A. Literature review on the effects of school feeding on 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 and cognition, 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).

After one year of implementation, a field experiment evaluating different implementation modalities of the World Food Programme school feeding program showed that the school feeding increased math scores for girls in internally-displaced people camps in Northern Uganda (Kazianga, de Walque, and Alderman 2012). No significant effects on cognitive tests, including Raven matrices and digit span tests were found. A randomized trial in 16 rural Jamaican schools showed that primary school children receiving a school breakfast had increases in their Maths 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 one rural district in Kenya focusing on providing meat, milk, and an "energy" meal to primary school-children as a mid-morning snack, documented improved Raven and arithmetic test scores for children in the meat and milk groups, but no differences in terms of digit span (Hulett et al. 2014; Neumann et al. 2007). Another study conducted in two districts in Western Kenya documented that a preschool breakfast program increases in preschooler's curricula test scores, but only for those attending more often and had a more experienced teacher (Vermeersch and Kremer 2005). A trial of fortified biscuits in a poor South African community reported positive effects on the Digit Span (van Stuijvenberg et al. 1999).

Appendix B. Development of poverty and food insecurity rankings

Poverty rankings were developed using the Ghana Living Standards Survey and Core Welfare Indicators Questionnaire carried out in 2005/2006 and 2003, respectively. Food consumption scores were calculated using the Comprehensive Food Security and Vulnerability Assessment 2008/2009 and spatial data variables computed by the World Food Programme (WFP). The data were subsequently used to generate district-level composites for the share of national poverty and food insecurity that were then used to allocate program resources (Gelli et al. 2016).

Appendix C. Balance of raw test scores for baseline sample prior to attrition

	Balance by treatment assignment				Balance by treatment assignment and attrition			
	Maths	Literacy	SPM	Digit span	Maths	Literacy	SPM	Digit span
School feeding	0.073	0.106	0.073	0.157	0.141	0.049	-0.236	-0.220
	(0.146)	(0.211)	(0.225)	(0.184)	(0.252)	(0.282)	(0.379)	(0.280)
Child in longitudinal sample					0.265	0.342	-0.019	0.062
					(0.177)	(0.228)	(0.285)	(0.216)
School feeding * Longitudinal sample					-0.070	0.068	0.337	0.412
					(0.273)	(0.303)	(0.375)	(0.304)
Constant	-1.492***	-1.304***	0.804***	1.152***	-1.734***	-1.618***	0.821**	1.094***
	(0.165)	(0.209)	(0.228)	(0.199)	(0.199)	(0.264)	(0.314)	(0.244)
Obs.	3,262	3,262	3,262	3,262	3,262	3,262	3,262	3,262
R-squared	0.204	0.144	0.125	0.164	0.205	0.146	0.125	0.166
Baseline control	1.54	1.78	3.81	4.11				
	(2.01)	(2.40)	(2.60)	(2.22)				
Baseline treatment	1.67	1.93	3.93	4.31				
	(2.05)	(2.51)	(2.55)	(2.20)				
Baseline control - lost to follow-up					1.13	1.28	3.66	3.88
					(1.51)	(1.88)	(2.50)	(2.14)
Baseline control - longitudinal sample					1.57	1.81	3.82	4.12
					(2.04)	(2.43)	(2.61)	(2.23)
Baseline treatment - lost to follow-up					1.68	1.97	3.96	4.35
					(2.07)	(2.56)	(2.56)	(2.21)
Baseline treatment - longitudinal sample					1.51	1.58	3.66	3.90
					(1.85)	(1.83)	(2.48)	(2.13)

*Notes: *** p<0.01, ** p<0.05, * p<0.1. Columns 1-4 test whether there were differences in raw scores by treatment arm prior to attrition. This is achieved through a OLS regression in which each child test score at baseline is regressed on a treatment dummy. Columns 1-4 add a dummy if the child was in the longitudinal sample and an interaction between longitudinal sample and treatment in order to* *investigate survey attrition bias. All models include child age in months and standard errors are clustered at the community level. Raw test scores appeared balanced by treatment arm prior to attrition, and there is no evidence of differential attrition by treatment being associated with raw test scores. At the bottom of the table, mean values by group are reported, with standard errors in parentheses.*

Appendix D. Balance of child and household characteristics at endline, full sample and northern regions

	Child age in months	Child aged 5-11 at baseline	ale	umber of children under 5 years in the household	ouse-hold size	ead of the house-hold is male	Mother's age	Wealth index	as sold any produce in the past year	ivestock	Child has been sick in the past week
Full longitudinal sample											
School feeding	1.782	-0.008	0.009	0.081	0.177	0.036	0.910	-1.117	0.008	0.024	-0.021
	(1.418)	(0.017)	0.020)	0.067)	0.336)	0.037)	(0.787)	(3.238)	0.054)	0.040)	(0.016)
Constant	130.809***	0.837***	.540***	.802***	.501***	.821***	40.534***	30.070***	.320***	.748***	0.099***
	(1.038)	(0.013)	0.016)	0.044)	0.236)	0.027)	(0.579)	(2.414)	0.038)	0.025)	(0.012)
Obs.	2,570	3,203	,602	,203	,203	,201	3,044	3,201	,200	,200	2,274
R-squared	0.001	0.000	.000	.002	.001	.002	0.002	0.001	.000	.001	0.001
Northern regions only											
School feeding	2.563	0.002	.002	.022	0.002	0.056	1.401	1.816	0.031	0.009	-0.010
	(2.024)	(0.027)	0.026)	0.111)	0.546)	0.043)	(1.047)	(1.504)	0.075)	0.041)	(0.020)
Constant	129.107***	0.829***	.541***	.736***	.685***	.901***	39.417***	16.090***	.299***	.830***	0.089***
	(1.510)	(0.022)	0.020)	0.067)	0.351)	0.026)	(0.720)	(1.100)	0.063)	0.019)	(0.015)
Obs.	1,236	1,495	,251	,495	,495	,495	1,440	1,495	,495	,495	1,087
R-squared	0.002	0.000	.000	.000	.000	.007	0.005	0.013	.001	.000	0.000

*Notes: * p<0.1. This table reports balance in endline covariates by treatment assignment obtained through an OLS regression with each covariate as an outcome and randomized assignment to school feeding as regressor. Standard errors were clustered at the community level. For each outcome or background characteristics, the estimated school feeding coefficient provides the difference between the school feeding and control group in a child's backgrounds and its standard errors.*

Appendix E. Predictors of program uptake

	(1)	(2)	(3)	(4)
	Uptake at endline			
Child age in months	0.998	0.996	0.996	0.989
	(0.003)	(0.003)	(0.003)	(0.007)
Male	1.128	1.052	1.048	1.161
	(0.130)	(0.147)	(0.149)	(0.287)
Child aged 5-11 years at baseline	2.176***	2.316***	2.307***	1.105
	(0.428)	(0.483)	(0.481)	(0.359)
Below poverty line	2.062***	1.800***	1.778***	1.280
	(0.376)	(0.300)	(0.300)	(0.312)
Northern regions	2.593***	2.378**	2.433**	1.042
	(0.908)	(0.816)	(0.870)	(0.564)
Maths standardized scores at baseline		0.815*	0.818*	0.845
		(0.090)	(0.093)	(0.109)
Literacy standardized scores at baseline		0.662***	0.667***	0.809
		(0.075)	(0.075)	(0.124)
SPM standardized scores at baseline		0.835**	0.835**	0.886
		(0.072)	(0.073)	(0.106)
Digit span scores at baseline		1.031	1.032	0.940
		(0.095)	(0.095)	(0.110)
GSFP program (HGSF is baseline)			1.143	
			(0.419)	
Grade at baseline				1.003
				(0.106)
Child is enrolled in primary school				5.721***
				(1.507)
Private school				0.011***
				(0.005)
Child has fallen sick in the past week				0.756
				(0.294)
Number of children under 5 years				0.941
				(0.158)
Household size				1.050
				(0.064)
Head of the household is male				0.873
				(0.272)
Mother's age				0.980*
				(0.010)
Wealth index				0.965**

				(0.017)
Household sold produce in the past year				1.416
				(0.428)
Livestock				2.040**
				(0.588)
Constant	0.666	0.904	0.842	9.074**
	(0.361)	(0.509)	(0.500)	(9.592)
Obs.	1,361	1,258	1,258	1,044

*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table presents results odd ratios estimated through logistic regression models of predictors of school feeding uptake at endline for children in the treatment arm. We regress uptake on a set of variables, including key predictors of heterogeneity (Col.1); baseline scores in the various competencies (col. 2); modality of school feeding (col.3) and child- and household-level characteristics. All models clustered the standard errors at the community level. In the last model, enrolment in primary school (as opposed to secondary) and livestock ownership positively predict a child's self-reported uptake of school feeding. By contrast, mother's age, wealth index and enrolment in private school predict lower odds of receiving the meal, suggesting that higher socio-economic status households tend to opt out from the programme.*

Appendix F. Descriptive statistics of raw test scores at baseline and endline, by child gender, household poverty status, and Northern regions

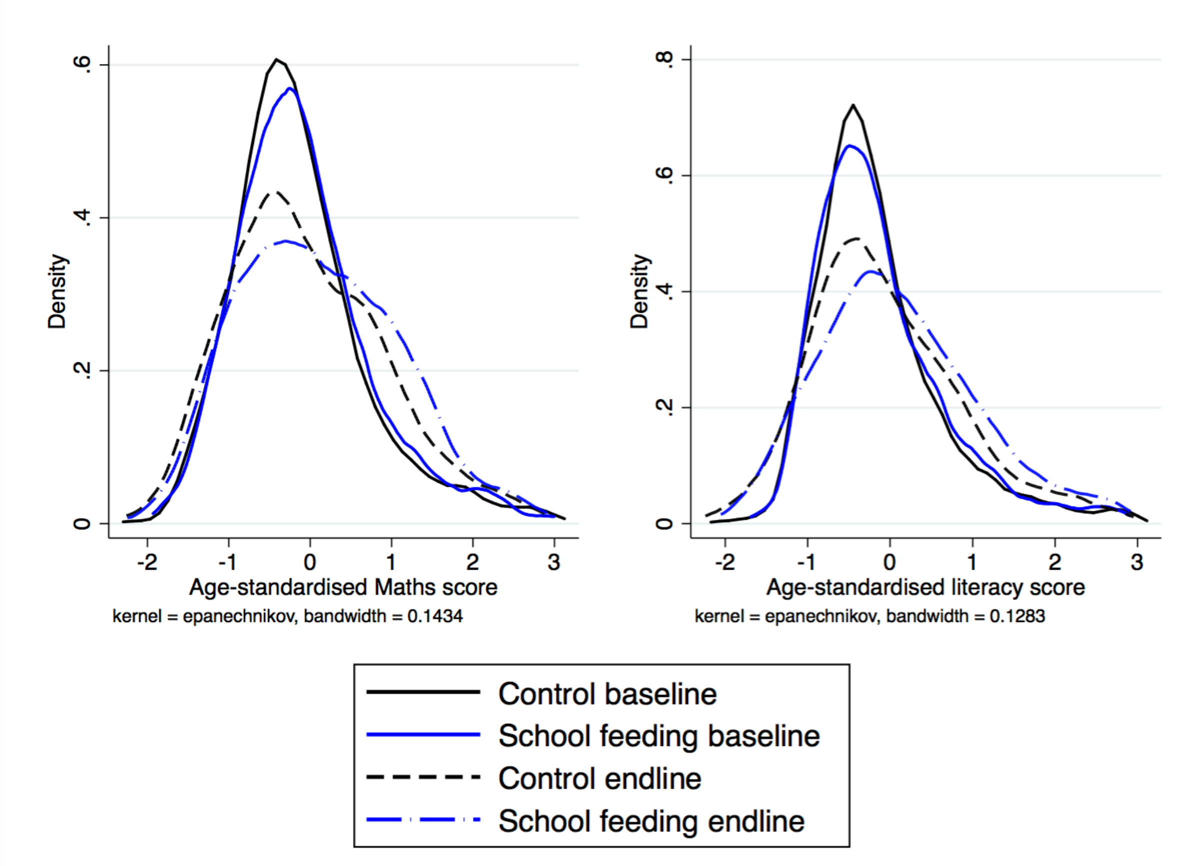
	Panel A. Gender									
	Baseline					Endline				
	Girls		Boys			Girls		Boys		
		Mean/SE		Mean/SE	Difference		Mean/SE		Mean/SE	Difference
Maths	,433	1.585 [0.053]	,580	.675 [0.053]	-0.09	,185	3.882 [0.101]	,372	3.805 [0.090]	0.076
Literacy	,433	1.884 [0.065]	,580	.903 [0.064]	-0.019	,185	4.136 [0.102]	,372	4.106 [0.097]	0.029
Digit span	,433	4.273 [0.058]	,580	.219 [0.057]	0.054	,185	4.231 [0.072]	,372	4.319 [0.068]	-0.087
SPM	,433	3.879 [0.067]	,580	.904 [0.066]	-0.025	,185	3.067 [0.057]	,372	3.179 [0.055]	-0.112
	Panel B. Poverty									
	Baseline					Endline				
	Non-Poor		Poor			Non-Poor		Poor		
		Mean/SE		Mean/SE	Difference		Mean/SE		Mean/SE	Difference
Maths	,324	1.66 [0.043]	88	.541 [0.078]	0.12	,953	3.906 [0.077]	04	3.631 [0.136]	0.275*
Literacy	,324	1.974 [0.054]	88	.626 [0.083]	0.348***	,953	4.225 [0.082]	04	3.78 [0.136]	0.445***
Digit span	,324	4.289 [0.046]	88	.093 [0.082]	0.196**	,953	4.314 [0.057]	04	4.161 [0.102]	0.154
SPM	,324	3.911 [0.055]	88	.831 [0.092]	0.08	,953	3.157 [0.045]	04	3.03 [0.082]	0.127
	Panel C. Region of residence									
	Baseline					Endline				
	Southern regions		Northern regions			Southern regions		Northern regions		

			M		**]**	**D**		**M**		**M**	**D**
			ean/SE		**ean/SE**	**ifference**		**ean/SE**		**ean/SE**	**ifference**
	Mat		1		1	0		3		3	-
hs		,626	.918	,387	.298	.620***	,326	.769	,231	.918	0.149
			[		[			[		[	
			0.055]		0.048]			0.089]		0.102]	
	Liter		2		1	0		4		3	0
acy		,626	.351	,387	.358	.993***	,326	.241	,231	.99	.250*
			[		[			[		[	
			0.069]		0.053]			0.101]		0.097]	
	Digit		4		3	1		4		4	0
span		,626	.758	,387	.643	.115***	,326	.502	,231	.037	.464***
			[		[			[		[	
			0.057]		0.053]			0.065]		0.075]	
	SPM		4		3	1		3		3	0
		,626	.36	,387	.345	.015***	,326	.133	,231	.12	.013
			[		[			[		[	
			0.068]		0.061]			0.056]		0.056]	

*Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. SPM stands for standardized progressive matrices. 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.*

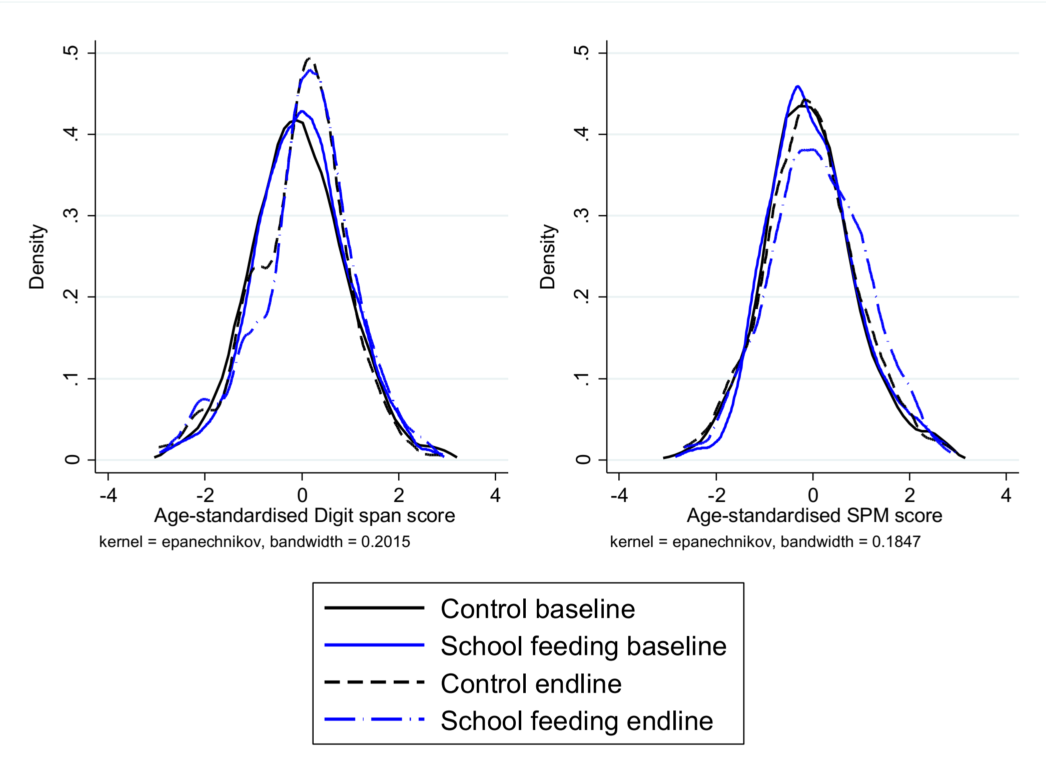
Appendix G. Empirical distribution of age-standardised test scores

Supplementary Figure G1. Empirical distributions of age-standardized maths (left-side panel) and literacy (right-side panel) age-standardized test scores, by survey round and treatment arm



Notes: this figure presents, by treatment group and survey wave, the nonparametric distributions of maths (left-side panel) and literacy (right-side panel) age-standardized scores for the full longitudinal sample of children. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel.

Supplementary Figure G2. Empirical distributions of age-standardized digit span (left-side panel) and SPM (right-side panel) age-standardized test scores, by survey round and treatment arm

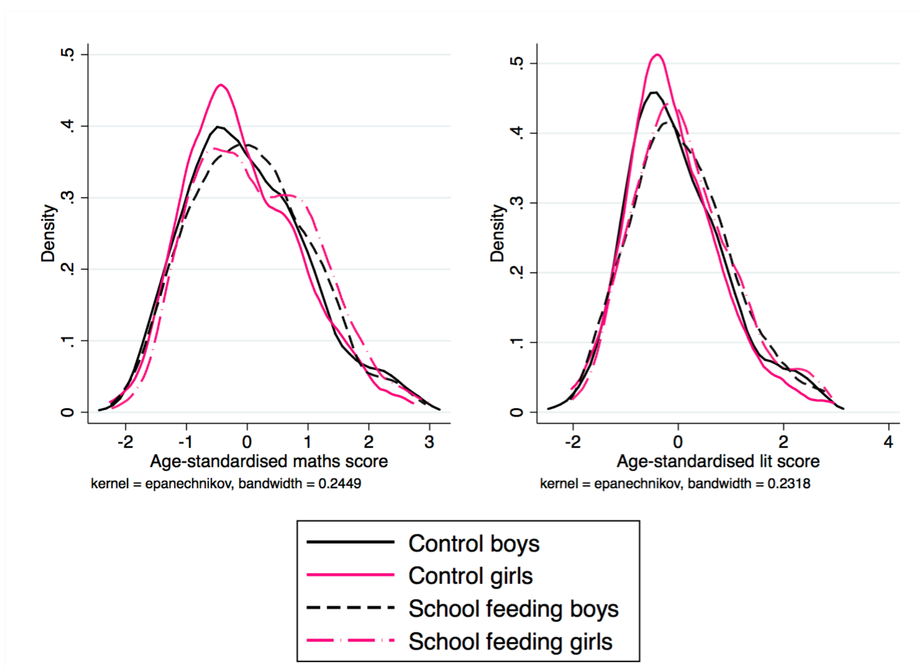


Notes: this figure presents, by treatment group and survey wave, the nonparametric distributions of Digit span (left-side panel) and standardized progressive matrices (SPM, right-side panel) age-standardized scores for the full longitudinal sample of children. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel.

Appendix H. Empirical distributions of age-standardized test scores by gender, poverty, and place of residence

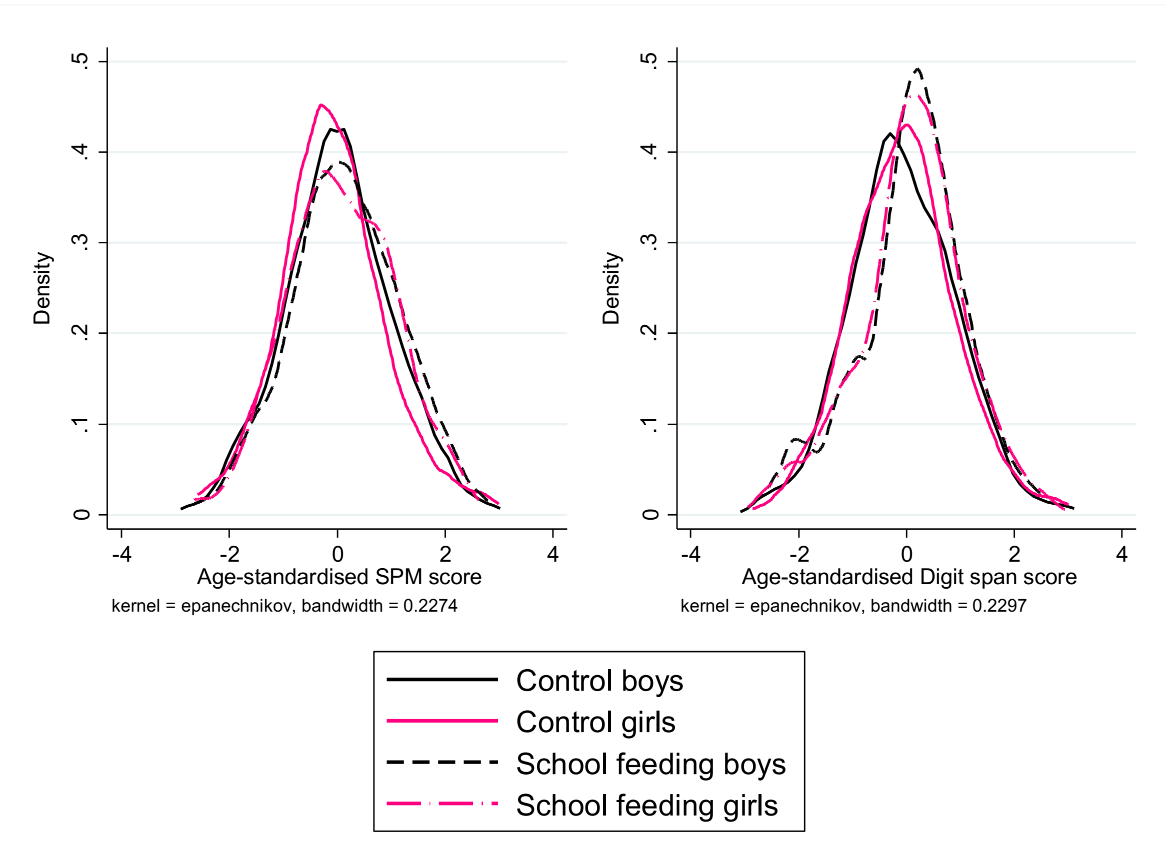
Supplementary Figures H1–H6 provide the nonparametric distributions of age-standardized test scores of achievements by gender, houshold poverty status, and region of residence at endline. While at baseline the distribution of achievements tended to overlap between treatment and control group, highlighting 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.

Supplementary Figure H1. Empirical distributions of age-standardized maths (left-side panel) and literacy (right-side panel) age-standardized test scores at endline, by child gender and treatment arm



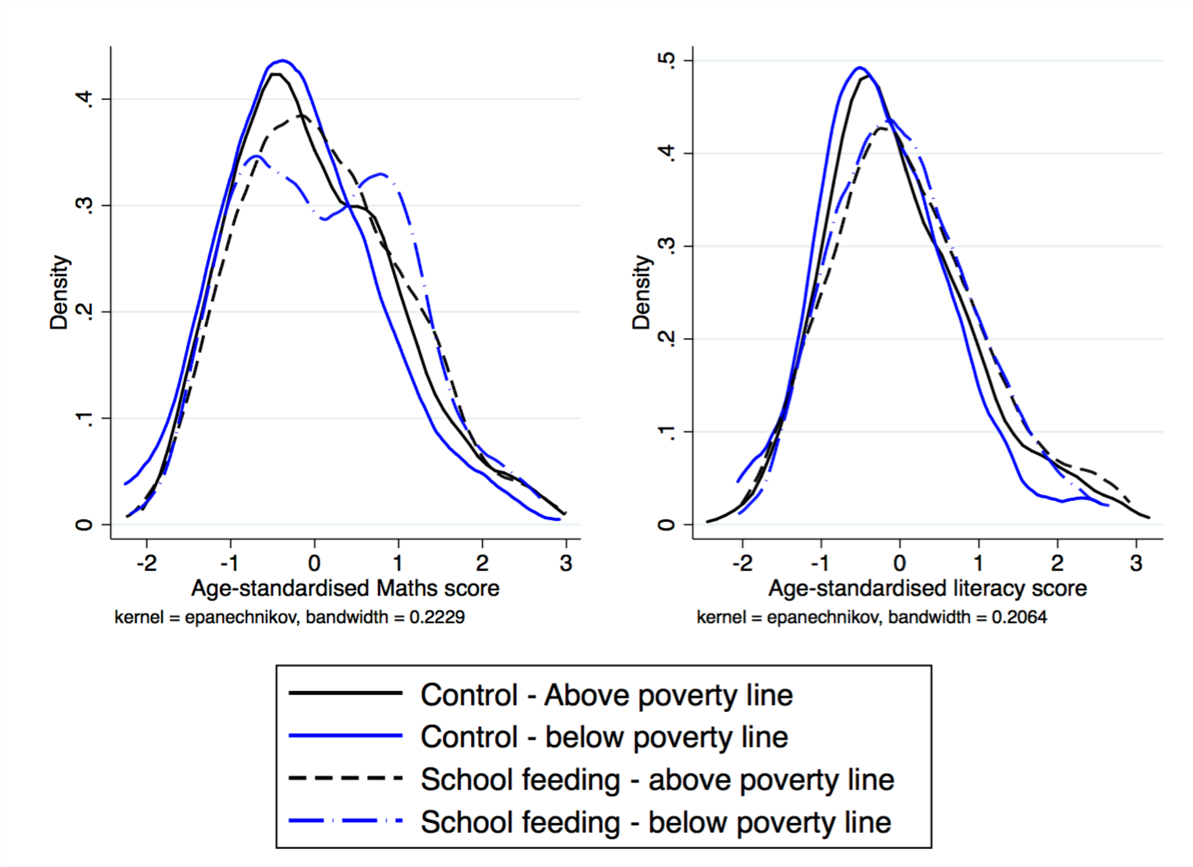
Notes: this figure presents, by treatment group and child gender, the nonparametric distributions of maths (left-side panel) and literacy (right-side panel) age-standardized scores for the full longitudinal sample of children at endline. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel.

Supplementary Figure H2. Empirical distributions of age-standardized SPM (left-side panel) and digit span (right-side panel) age-standardized test scores at endline, by child gender and treatment arm



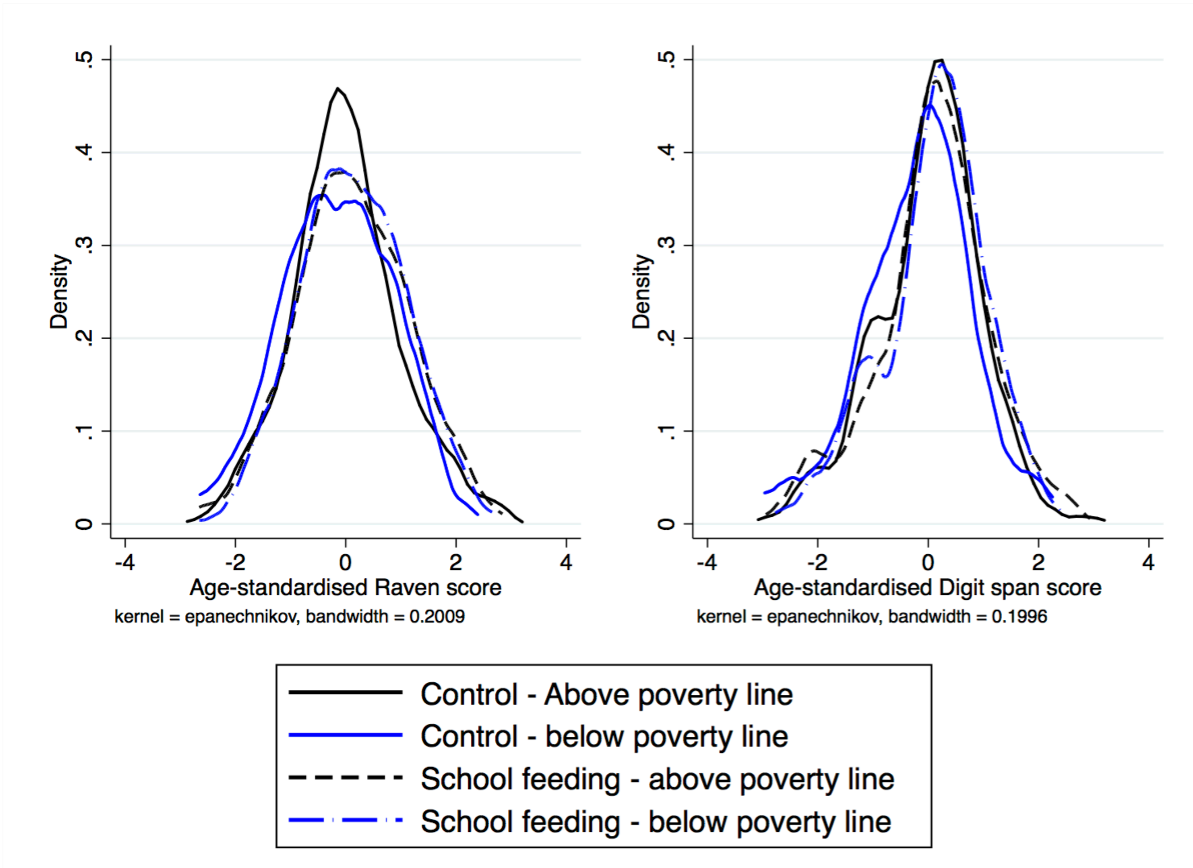
Notes: this figure presents, by treatment group and child gender, the nonparametric distributions of standardized progressive matrices (SPM, left-side panel) and digit span (right-side panel) age-standardized scores for the full longitudinal sample of children at endline. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel.

Supplementary Figure H3. Empirical distributions of age-standardized Maths (left-side panel) and literacy (right-side panel) age-standardized test scores at endline, by household poverty status and treatment arm.



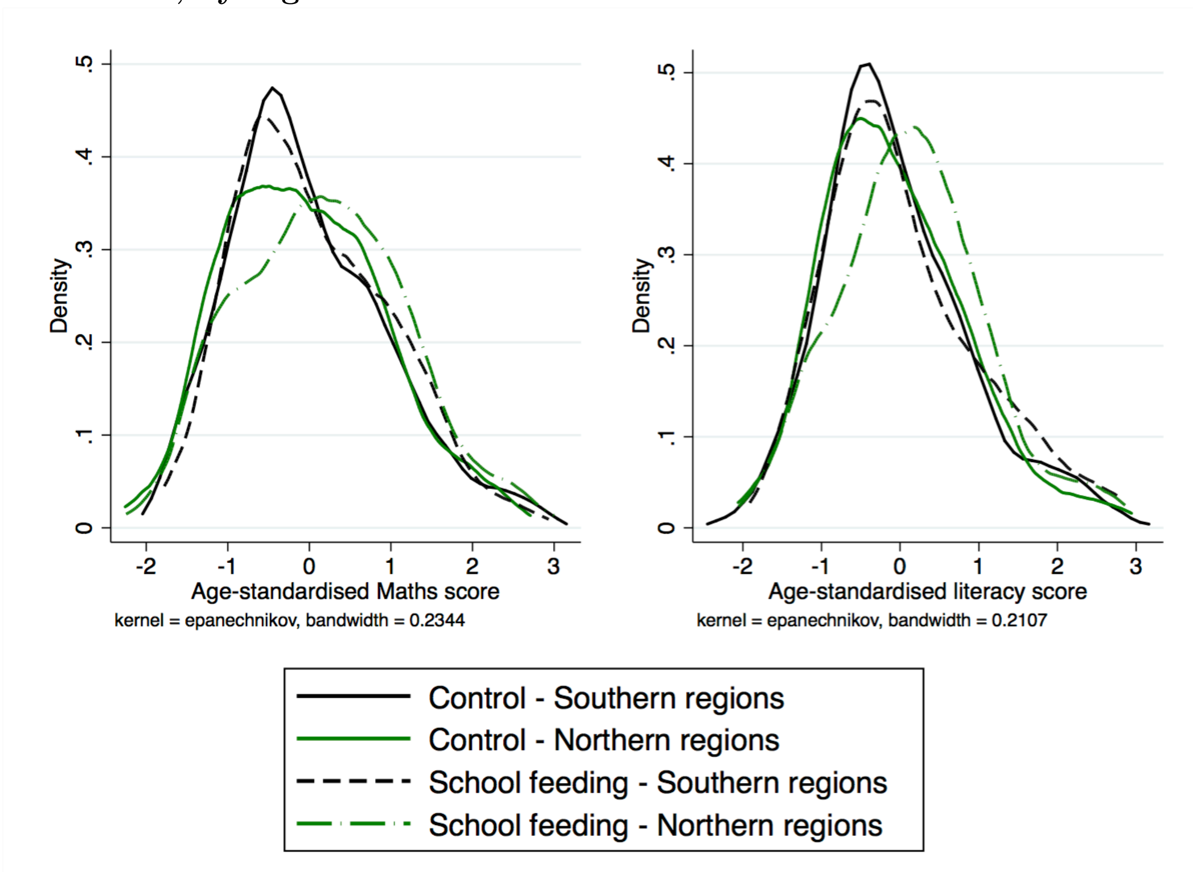
Notes: this figure presents, by treatment group and household poverty status at baseline, the nonparametric distributions of Maths (left-side panel) and literacy (right-side panel) age-standardized scores for the full longitudinal sample of children at endline. Nonparametric distributions were calculated 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.

Supplementary Figure H4. Empirical distributions of age-standardized SPM (left-side panel) and digit span (right-side panel) age-standardized test scores at endline, by household poverty status and treatment arm



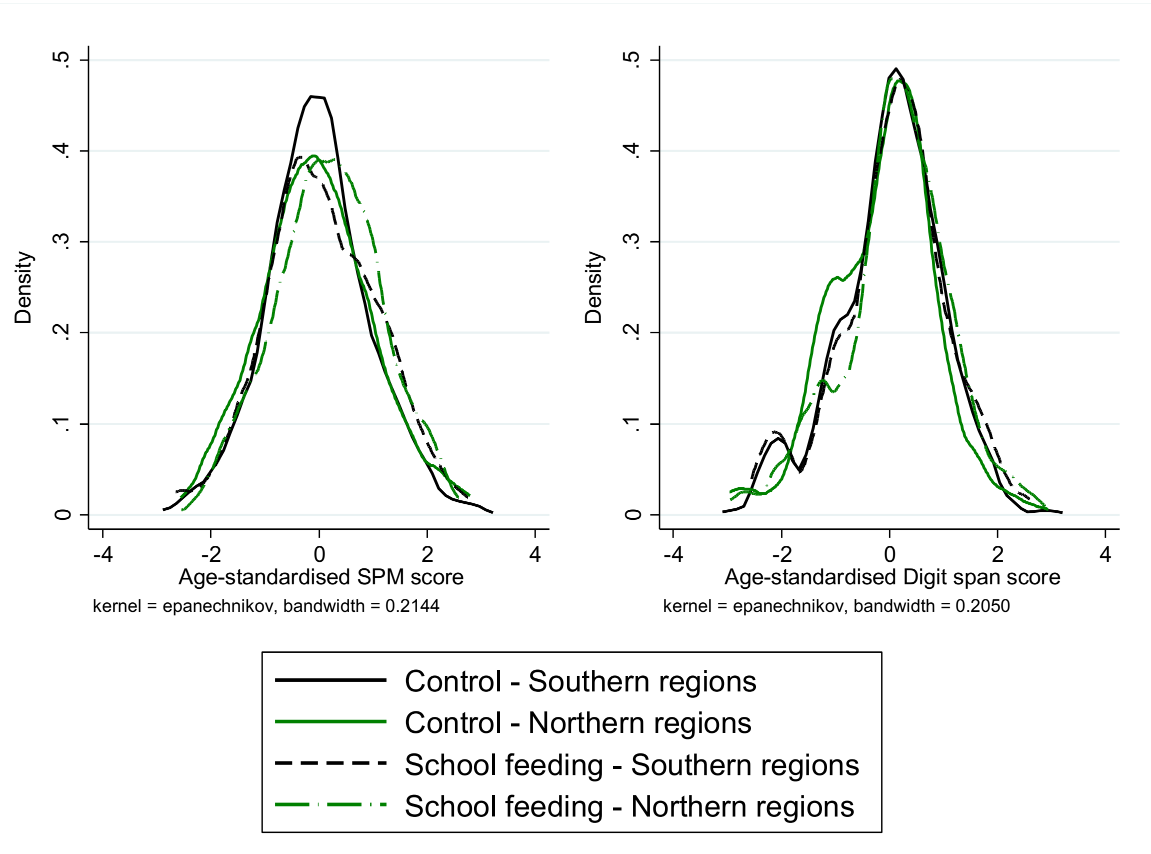
Notes: this figure presents, by treatment group and household poverty status at baseline, the nonparametric distributions of standardized progressive matrices (SPM, left-side panel) and digit span (right-side panel) age-standardized scores for the full longitudinal sample of children at endline. Nonparametric 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.

Supplementary Figure H5. Empirical distributions of age-standardized maths (left-side panel) and literacy (right-side panel) age-standardized test scores at endline, by region of residence and treatment arm



Notes: this figure presents the nonparametric distributions of maths (left-side panel) and literacy (right-side panel) raw scores at endline by treatment group and household region of residence. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo.

Supplementary Figure H6. Empirical distributions of age-standardised SPM (left-side panel) and digit span (right-side panel) age-standardized test scores at endline, by region of residence and treatment arm



Notes: this figure presents the nonparametric distributions of standardized progressive matrices (SPM, left-side panel) and digit span (right-side panel) raw scores at endline by treatment group and household region of residence. Nonparametric distributions were calculated through weighted local polynomial regressions using an Epanechnikov kernel. Northern regions include Upper West, Upper East, and Northern region. Southern regions include Western, Central, Greater Accra, Volta, Eastern, Asanti, Brong Ahafo

Appendix I. Treatment effects estimated through Difference-in-Differences

	Maths	Literacy	SPM	Digit span	Composite: maths and literacy	Composite: SPM and digit span	Composite: all outcomes
Panel A. All children							
School feeding * endline	0.114	0.139*	0.103	0.068	0.149	0.102	0.13
	(0.087)	(0.083)	(0.085)	(0.086)	(0.102)	(0.092)	(0.1)
Observations	5,423	5,415	5,457	5,449	5,466	5,473	5,476
R-squared	0.031	0.043	0.02	0.031	0.044	0.03	0.042
Panel B. Girls							
School feeding * Endline	0.207**	0.194**	0.098	0.099	0.240**	0.124	0.191*
	(0.096)	(0.092)	(0.103)	(0.099)	(0.113)	(0.107)	(0.114)
Observations	2,565	2,560	2,584	2,579	2,584	2,590	2,591
R-squared	0.038	0.037	0.014	0.032	0.042	0.027	0.038
Panel C. Households below the poverty line							
School feeding * Endline	0.261**	0.324***	0.155	0.127	0.330***	0.158	0.272**
	(0.123)	(0.097)	(0.101)	(0.124)	(0.123)	(0.118)	(0.112)
Observations	1,265	1,264	1,264	1,266	1,270	1,270	1,271
R-squared	0.063	0.062	0.039	0.051	0.057	0.05	0.058
Panel D. Northern regions							
School feeding * Endline	0.199	0.197*	0.227**	0.155	0.237	0.230**	0.247**
	(0.135)	(0.116)	(0.098)	(0.096)	(0.142)	(0.097)	(0.12)
Observations	2,557	2,564	2,575	2,570	2,574	2,578	2,579
R-squared	0.034	0.047	0.024	0.029	0.031	0.03	0.036

*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table above illustrates intent-to-treat effects on each outcome estimated for the full sample and for the subgroups through difference-in-differences. Models include a dichotomous variable for treatment assignment, a dummy for endline survey, and the treatment effect relates to the interaction between these two variables. It also includes region dummies. Maths, literacy,* *standardized progressive matrices (SPM) and digit span scores are age-standardized. Composite indices were computed as averages of the standardized scores and then they were standardized to the control group within each round.*

Appendix J. Treatment effects of school feeding on child learning and cognition, by child age group at baseline

	Maths	Literacy	SPM	Digit span	Composite: maths and literacy	Composite: SPM and digit span	Composite: all outcomes
Panel A: Younger children							
School feeding	0.161**	0.132*	0.128**	0.113*	0.177**	0.138**	0.161**
	(0.078)	(0.076)	(0.060)	(0.064)	(0.086)	(0.064)	(0.076)
Observations	2,011	2,006	2,039	2,036	2,045	2,052	2,054
R-squared	0.061	0.124	0.034	0.054	0.135	0.064	0.127
Panel B: Older children							
School feeding	0.040	0.123	0.118	0.132	0.096	0.145	0.126
	(0.126)	(0.108)	(0.114)	(0.124)	(0.128)	(0.127)	(0.131)
Observations	267	268	268	269	269	269	269
R-squared	0.167	0.201	0.064	0.050	0.185	0.079	0.135

*Notes: *** p<0.01, ** p<0.05, * p<0.1. The table above illustrates intent-to-treat effects on each outcome estimated for different age cohorts. The younger and adolescent cohorts refer, respectively, to children aged 5-11 years and 12-15 years at baseline. Both models were estimated through OLS and standard errors were clustered at the community level. For each outcome, the model controls for the baseline value of the outcome, and region dummies. Maths, literacy, standardized progressive matrices (SPM) and digit span scores are age-standardized. Composite indices were computed as averages of the standardized scores and then they were standardized to the control group within each round.*

Appendix K. Heterogeneity of impact by treatment modality

The analysis plan did not include the comparison in treatment effects by school feeding implementation modality. This is because, by design, the comparison between the standard GSFP and the HGSF modality was geared only toward assessing the impact of HGSF on small-holder farmers income and production (see Gelli et al. 2016, for further discussion). However, different implementation modalities may affect program impact on child learning through variation in frequency of delivery of the meal, nutrition content, type of meal, timing of delivery (e.g., breakfast *viz* lunch), etc. The HGSF, through its emphasis on improving the quality of the meal, may, theoretically, lead to better cognition and learning as compared to standard GSFP through enhanced child health (Belot and James 2011).

Table K.1 provides ITT estimates of child learning outcomes by considering GSFP, HGSF, and control as three separate arms. For both ANCOVA and difference-in-differences estimates, the last row in each respective panel includes an F-test that assesses the equality of the treatment effect coefficients related to GSFP and HGSF. In the case of ANCOVA, assignment to the HGSF arm led to significant increases in literacy, SPM, and the three composite indicators, as compared to control. However, in either set of estimates, we were never able to reject the null hypothesis of equality of the treatment effect coefficients between HGSF and GSFP, thus suggesting lack of heterogeneity in impact by program modality. This may be either attributable to insufficient power to detect significant differences (as by design this comparison was not initially pursued) and/or to challenges in the implementation. We tend to lean toward the second explanation, also in light of similar coefficient sizes for both modalities in most cases. We hypothesise that delayed reinmbursements to caterers for the costs incurred in supplying the meals may have prompted deviations from the planned content and portion sizes of the school feeding menu, particularly in the case of the HGSF arm. In this arm, caterers initially committed to adhere to a nutritionally-enhanced menu, but they may have in practice deviated from it, due to the reimbursement delays and other implementation challenges (e.g., monitoring visits highlighted substantial deviations from the guidelines related to food fortification). Thus, the nutritional differences of the meals between the two modalities may have been in practice diluted, making them too limited to have heterogenous impacts on children's academic achievements through the health channel. A similar lack of heterogeneity by modality was evident in the group-disaggregated estimates (available upon request).

Table K.1. Treatment effects of school feeding on child learning and cognition, by school feeding modality

	Maths	Literacy	SPM	Digit span	Composite: maths and literacy	Composite: SPM and digit span	Composite: all outcomes
GSFP	0.159*	0.121	0.086	0.120	0.166*	0.118	0.147*
	(0.085)	(0.084)	(0.070)	(0.073)	(0.094)	(0.073)	(0.083)
HGSF	0.136	0.143	0.169**	0.119	0.169*	0.167**	0.171*
	(0.087)	(0.089)	(0.067)	(0.082)	(0.100)	(0.078)	(0.090)
Constant	-0.229**	-0.291*	-0.262*	-0.199	-0.339**	-0.275*	-0.375**
	(0.104)	(0.154)	(0.156)	(0.134)	(0.151)	(0.165)	(0.162)
Observations	2,278	2,274	2,307	2,305	2,314	2,321	2,323
R-squared	0.068	0.130	0.035	0.050	0.139	0.064	0.127
P(GSFP=HGSF)	0.782	0.823	0.309	0.994	0.983	0.569	0.796

*Notes: *** p<0.01, ** p<0.05, * p<0.1 The table above illustrates intent-to-treat effects on each outcome for the two school feeding modalities. GSFP is a dichotomous variable related to randomized assignment to the standard Ghana school feeding program, HGSF is a dichotomous related to assignment to "home-grown" school feeding pilot, endline is a dummy variable indicating the 2016 survey. Both models were estimated through OLS and standard errors were clustered at the community level. For each outcome,* *the model controls for the baseline value of the outcome, a dichotomous variable related to the randomized assignment to school feeding, and region dummies. The last row presents the p-values of a F-test assessing the equality of coefficients between the intent-to-treat effect related to GSFP and HGSF.*

Appendix L. Descriptive statistics of intermediate outcomes, full sample

		Baseline						Endline					
					Sch			Contr			Sch		
		Control		ool feeding				ol		ool feeding			
			M		ean				M		ean		
		ean						ean					D
		(SE)			SE)	iff.		(SE)			SE)	iff.	
Child is enrolled			0						0				-
	,353	.988		,604	.981	.007	,195	.884		,377	.932	0.049***	
			[						[				
		0.003]			0.003]			0.009]			0.007]		
Child is enrolled in			1						0				-
primary school	94			,034		/A	,057	.861		,284	.868	0.007	
			[						[				
		0.000]			0.000]			0.011]			0.009]		
Days attended over past			4						4				-
week	,291	.881		,508	.828	.052**	,056	.665		,284	.685	0.021	
			[						[				
		0.018]			0.019]			0.030]			0.026]		
Grade attained			2						4				-
	,368	.336		,589	.436	0.100*	,049	.269		,278	.496	0.227***	
			[						[				
		0.039]			0.037]			0.061]			0.055]		
Repeated a grade			0						0				0
	,257	.113		,480	.124	0.011	,057	.257		,284	.223	.035*	
			[						[				
		0.009]			0.009]			0.013]			0.012]		
Time at school									5				-
							,186	.775		,371	.018	0.243**	
									[				
								0.079]			0.068]		
Time for study									0				0
							,186	.615		,371	.581	.034	
									[				
								0.047]			0.041]		
Housework time									1				0
							,186	.816		,371	.685	.131**	
									[				
								0.049]			0.041]		
Labor									0				0
							,186	.691		,371	.563	.128**	
									[				
								0.042]			0.034]		
Hours for leisure									2				-
							,186	.062		,371	.139	0.078	
									[				
								0.042]			0.036]		

*Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. Enrolment and enrolment in primary are dichotomous variables indicating whether the child is enrolled to any level of education or primary school, respectively; 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. Grade repetition is a dichotomous variable of whether the child has ever repeated a grade. Time spent in school is an indicator that measures the average time (in hours) a child spend in school on a typical week day (in the school term); time spent in study is an indicator that measures the average time (in hours) a child spend studying at home or taking private tuition on a typical day; time spent in household chores is an indicator that measures the average time (in hours) a child spends doing household chores or care for other household members on a typical day; time in work is an indicator that measures the amount of time a child spends working in agriculture work, in the family business or in another work activity on a typical day; leisure time measures the average amount of play and other leisure activities on a typical day.*

References supplementary materials

Alderman, H., and D. Bundy. 2012. "School Feeding Programs and Development: Are We Framing the Question Correctly?" *World Bank Research Observer* 27(2):204–221. https://doi.org/10.1093/wbro/lkr005.

Belot, M., and J. James. 2011. "Healthy School Meals and Educational Outcomes." *Journal of Health Economics* 30(3):489–504. https://doi.org/10.1016/J.JHEALECO.2011.02.003.

Drake, L., M. Fernandes, E. Aurino, J. Kiamba, B. Giyose, C. Burbano, H. Alderman, M. Lu, A. Mitchell, and A. 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, DC: The World Bank.

Gelli, A., E. Masset, G. Folson, A. Kusi, D. K. Arhinful, F. Asante, I. Avi, K. M. Bosompem, K. Watkins, L. Abdul-Rahman, R. Agble, G. Ananse-Baden, D. Mumuni, E. Aurino, M. Fernandes, and L. J. Drake. 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.

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.

Hulett, J. L., R. E. Weiss, N. O. Bwibo, O. M. Galal, N. Drorbaugh, and C. 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):875–886. https://doi.org/10.1017/S0007114513003310.

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):723–760. https://doi.org/10.1093/jae/ejs010.

Neumann, C. G., S. P. Murphy, C. Gewa, M. Grillenberger, and N. O. Bwibo. 2007. "Meat Supplementation Improves Growth, Cognitive, and Behavioral Outcomes in Kenyan Children." *The Journal of Nutrition* 137(4):1119–1123.

Powell, C. A., S. P. Walker, S. M. Chang, and S. M. Grantham-McGregor. 1998. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." *The American Journal of Clinical Nutrition* 68(4):873–879.

Snilstveit, B., J. Stevenson, D. Phillips, M. Vojtkova, E. Gallagher, T. Schmidt, H. Jobse, M. Geelen, M. Pastorello, and J. Eyers. 2015. *Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review*. London: International Initiative for Impact Evaluation (3ie).

van Stuijvenberg, M. E., J. D. Kvalsvig, M. Faber, M. Kruger, D. G. Kenoyer, and A. J. Spinner 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.

Vermeersch, C., and M. Kremer. 2005. *School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation*. Policy Research Working Paper 3523). Washington, DC: World Bank. https://doi.org/10.2139/ssrn.667881.