Forthcoming: Journal of Development Economics

SCHOOL FEEDING AND LEARNING ACHIEVEMENT: EVIDENCE FROM INDIA'S MIDDAY MEAL PROGRAM

TANIKA CHAKRABORTY AND RAJSHRI JAYARAMAN

ABSTRACT. We study the effect of the world's largest school feeding program on children's learning outcomes. Staggered implementation across different states of a 2001 Indian Supreme Court Directive mandating the introduction of free school lunches in public primary schools generates plausibly exogenous variation in program exposure across different birth cohorts. We exploit this to estimate the effect of program exposure on math and reading test scores of primary school-aged children. We find that prolonged exposure to midday meals has a robust positive effect on learning achievement. We further investigate various channels that may account for this improvement including complementary schooling inputs, heterogeneous responses by socio-economic status, and intra-household redistribution.

JEL Classification: I21, I25, O12

Keywords: school feeding, learning, midday meal, primary school education

---

Date: September 2018.

1. Introduction

School feeding programs are ubiquitous. The World Food Program estimated that in 2013, 368 million children, or one in five, received a school meal at a total cost of US$ 75 billion (WFP, 2013). There are two main rationales for this sizable investment. The first is to abate hunger and improve health and nutrition. The second is to improve schooling outcomes.

This paper analyzes the latter by evaluating the impact of India's free school lunch program—which we will refer to by its local moniker, "midday meals"—on learning achievements of primary school children. Importantly, we examine the effect of long-term program exposure of up to five years, on learning outcomes. Proponents argue that free in-school feeding programs have a positive impact on learning through two main channels. First, they encourage school participation in the form of school enrollment or attendance. The latter in particular affords children the opportunity to learn in the first place. Second, they improve children's nutritional intake; alleviation of short-term hunger facilitates concentration, and improved health and nutritional status leads to better cognition and lower absenteeism due to illness.

However, these positive effects are not self-evident for three reasons. First, complementary schooling inputs, such as teachers or school infrastructure, are presumably needed in order to translate potential increases in school participation and improvements in nutrition into better learning outcomes. Second, if children are already well-nourished (e.g. because they come from wealthy families), then school feeding may not provide any added benefit. Third, the program may not actually improve a child's nutritional status if school meals induce families to substitute food away from a school-going child towards other family members. In addition to our main treatment effects, we explore each of these channels.

The Indian context we study is important for three reasons. First, the learning deficit in primary schools is large. An ASER (2005) report, for example, revealed that 44% of children between the ages of 7 and 12 who were actually enrolled in school could not read a basic paragraph and 50% could not do simple subtraction. Second, the scale of the intervention is massive: India's midday meal scheme is the largest school nutrition program in the world. In 2006, it provided lunch to 120 million children in government primary schools on every school day (Kingdon, 2007). To put this number in perspective, it accounts for one third of children globally who, according to the WFP (2013), enjoy school feeding programs. Finally, undernutrition is a severe problem in the country. India has some of the worst anthropometric indicators of nutrition in the world. According to the 2005-6 National Family Health Survey (NFHS-3), 48% of children under the age of 5 were considered chronically malnourished. Comparable data from 40 other developing countries covered by the Demographic and Health Surveys (DHS) indicates that this proportion is higher in India than in any of those other countries. Undernourishment during childhood has been well-documented to have deleterious lifetime consequences. In this context, school feeding programs have a potentially vital role to play in combating undernutrition.

In order to identify the causal effect of this program, we exploit its staggered implementation. Briefly, and in more detail later, a 2001 Indian Supreme Court directive ordered Indian states to institute free midday meals in government primary schools. Prior to 2001 only two states, Tamil Nadu and Gujarat, had universal public primary school midday meal provision. Over the subsequent five years, however, state governments across India introduced midday meals. Staggered implementation of the program in primary schools generates variation in the length of potential exposure to the program based on a child's birth cohort and state of residence. Children only enjoyed the program to the extent that they were of primary-school going age—6 to 10 years old—and lived in a state which had instituted midday meals in primary school. Hence, the earlier their state introduced the program and the young enough they were at the time, the longer was the child's potential exposure to midday meals in this intent-to-treat (ITT) framework.

Although our quasi-experimental empirical design has some advantages that we elaborate on below, its obvious disadvantage is that identification is not as clean as it is in experimental studies. Our main identifying assumption is that there are parallel trends in learning achievement within cohorts between early and late program implementers. We present evidence that the timing of implementation was plausibly exogenous to state-level characteristics; the descriptive analysis suggests that the stronger assumption of parallel trends in average outcomes between early and late implementers is plausible; and our results are robust to a number of specification checks pertaining to the timing of implementation. However, in keeping with all difference-in-differences-type empirical strategies, we must concede that we cannot formally test this assumption.

Our data come from the Annual Status of Education Report (ASER) survey, whose goal is to assess the state of education among children in India. It has three unique features that are useful for the purpose of this analysis. First, it has wide geographic coverage, surveying over 200,000 households across India's roughly 580 rural districts. Second, it has been administered annually since 2005. Third, ASER administers learning assessments of basic literacy (reading skills) and numeracy (number recognition and arithmetic skills) to all children aged 5 to 16. These features allow us to capture variation in exposure to treatment across states and time, while correcting for state- and cohort-specific effects as well as state-specific time trends, in order to assess the program's effect on learning. State fixed effects allow for average test scores to vary across different states, accounting for the possibility that children in better or worse performing states may have longer program exposure because their states implemented the program earlier or later. Cohort fixed effects address the concern that older children are likely to have higher test scores than younger children, and also potentially have longer program exposure. Finally, the inclusion of state-specific time trends permits for trends in average test performance to vary from state to state.

We find that exposure to midday meals increases students' learning achievement, albeit at a decreasing rate. Children with up to five years of exposure have reading test scores that are 18% higher, and math test scores that are 9% higher than students with less than a year of exposure. In terms of potential channels, when we explore complementarities, we find that schooling inputs that are directly related to teaching are associated with significantly higher learning when combined with a midday meal, but more general schooling infrastructure is not. At the same time, we find no evidence of heterogeneous treatment effects on the basis of gender or housing assets. Finally, we find limited evidence of intra-household redistribution from eligible children to other family members.

The benefits of the midday meal program in terms of test score improvements are roughly comparable to some recent interventions aimed at improving the test performance of primary school children in India, in particular the introduction of extra teachers (Muralidharan and Sundararaman, 2013) and tutoring (Banerjee et al., 2007). On the one hand, this is impressive given that improved school performance is, if anything, a side benefit of a program whose primary goal was to improve children's nutritional status. On the other hand, this improvement comes at a cost per child that is almost three times higher than these teaching and tutoring inputs. As a consequence, midday meals underperform relative to these programs in terms of test score improvements per dollar spent.

Related Literature. There is a substantial literature on the effect of school feeding programs on school participation and nutritional outcomes; see Alderman and Bundy (2012) for an excellent review. Most of these studies have focused on young, typically primary-school-aged children, and have generally found that there are positive treatment effects on both participation (e.g. higher school enrollment or attendance) and nutritional status (e.g. lower anemia or higher BMI).[1]

The focus of this paper is to examine the effects of school feeding on learning achievement. It speaks to two main strands of this literature. The first strand has examined the effect of school feeding programs on learning achievement in the context of small-scale, relatively short-term, randomized field experiments.[2] Most of these experiments find no effect on cognitive achievement. A handful that reports improvements finds it only for narrowly defined subsets of students on a subset of skills.

The second strand of the literature has explored the effect of India's midday meal program on children's nutritional and schooling outcomes using quasi-experimental methods. To the best of our knowledge, only two studies have examined its effect on learning outcomes, both using variation at the local level.[3] Singh (2008) finds improvements in Peabody Picture Vocabulary Tests in the Young Lives panel of about 500 children in Andhra Pradesh. However, he is cautious in the interpretation of this result since he lacks a control group in the analysis. Afridi et al. (2014) use the extension of midday meals to upper primary school (grades 6-8) in educationally "backward" localities to evaluate its effect on learning outcomes of 400 students in 16 Delhi schools using difference-in-differences. The authors find a significant improvement in classroom attention. However, in the 4-month time frame of their study, they find no improvement in academic test scores.

We contribute to these two strands of literature in a couple of ways. First, whereas extant evidence has focused on relatively short-term effects—anywhere from a few weeks to at most two years—we explore the effect of up to five years of program exposure. Second, we use a large dataset. This not only affords us statistical power but also allows us to study an intervention that has been implemented on a massive scale. The fact that the intervention is in no way "gold-plated", combined with our use of a large data set that is representative of rural India also arguably adds to the generalizability of our findings.

Our results are in line with the extant literature, in that we find negligible or no positive effects of midday meals on learning outcomes in the short run. However, in contrast to most of this literature,

[1]See, for example, Jacoby et al. (1998), Powell et al. (1998), Van Stuijvenberg et al. (1999), Jacoby (2002), Neumann et al. (2003), and Bhattacharya et al. (2006), who generally find positive effects of school feeding programs on children's health and nutritional status. Jacoby et al. (1998), Powell et al. (1998), Ahmed (2004), Kremer and Vermeersch (2005), Belot and James (2011), Kazianga et al. (2012) find positive effects of school feeding programs on school participation.

[2]See Kazianga et al. (2012), Powell et al. (1998), Van Stuijvenberg et al. (1999), Grantham-McGregor et al. (1998), Adelman et al. (2008), Neumann et al. (2007), Kremer and Vermeersch (2005), and Whaley et al. (2003). In a rare non-randomized evaluation of an extant national program, McEwan (2013) uses a regression discontinuity design to study the effect of Chile's long-established school feeding program on (among other things) fourth-grade test scores. The discontinuity comes from the fact that students received meals with different caloric content depending on a school-level "vulnerability" index cutoff. McEwan (2013) finds that there is no difference in test performance when the caloric content of meals is increased.

[3]Others have examined the effect of midday meals in India on school participation and nutritional outcomes; see Afridi (2010), Singh et al. (2014), Afridi (2011) and Jayaraman and Simroth (2015).

we find an unambiguously positive effect of school feeding after the almost five-year exposure period, measured by both reading and math test scores. There are three possible explanations for the difference between our findings and the previous literature. First, small sample sizes may account for imprecise estimates of positive effects in earlier studies. Our large dataset allows us to evade this problem. Second, the gains that we observe due to long-term exposure may not be fully captured in many of the shorter-term interventions that have been evaluated to date. The empirical design of this paper is not equipped to provide a definitive answer for *why* this discrepancy between short- and long-term exposure exists; that is left to future research. Conceptually, however, cognitive ability has been shown to be associated with cumulative nutrition, as measured in height or height-for-age, and it is likely that the long-term exposure we investigate here captures just that.[4] Moreover, as King and Behrman (2009) have argued in the context of natural policy settings such as ours, there may be lags in implementation, or learning and adjustment to social programs. On the provider side, setting up and operating a school feeding program is logistically challenging in terms of both physical inputs such as food delivery and cooking, and personnel management of teachers and cooks. On the beneficiaries side, learning and adoption by parents and students may also take time. Both of these lags may mean that a program which is otherwise effective in the long-run, may not appear so in the short-run.

The third explanation for why we find positive treatment effects while others often haven't, is the context we study. Our data come from villages in rural India where nutritional deficiency is a chronic problem. Using National Sample Survey (NSS) data, Deaton and Drèze (2009) calculate that in 2004-5—the first year of observation in our data—almost 80% of the rural population lived in households with a per capita calorie consumption below the rural poverty line of 2,400. It is conceivable that the sizable gains in learning achievement that we find reflect the fact that the target population of this intervention in our sample is extremely nutritionally disadvantaged to begin with. This would be in line with the findings of Powell et al. (1998), Van Stuijvenberg et al. (1999) and Grantham-McGregor et al. (1998). It is also consistent with the fact that we find no heterogeneous treatment effects in terms of gender or household assets, in that the bulk of children in our data are likely to suffer from substantial economic disadvantage.

The rest of the paper is organized as follows. Section 2 furnishes the policy background. Section 3 describes our data and empirical model. The main result—the effect of midday meals on test scores—is presented in Section 4. Section 5 delves into some channels that may drive the increase in test scores. Section 6 describes a series of robustness checks on our main result from Section 4, and Section 7 concludes with a cost-benefit analysis.

2. Policy Background

The Indian central government has a long-standing commitment to on-site school feeding programs.[5] In 1995, the central government mandated free cooked meals in all public primary schools via the National Program of Nutritional Support to Primary Education. In India, the central government's role

[4]See, for example, Case and Paxson (2008a,b), Schick and Steckel (2010), Karp et al. (1992) and in the Indian context, Spears (2012).

[5]The description of the natural experiment in this section draws, in part, from Jayaraman and Simroth (2015), who exploit the same policy setting that we do. However, the question we address, the data, the empirical strategy, and the analysis are completely different.

in school education lies in issuing policy guidelines and providing funding. Policy implementation is the prerogative of state governments, and not a single state responded to this universal mandate.[6]

Half a decade later, India witnessed a sea change. In early 2001 there was a severe drought in 7 districts, to which the press and many civil society organizations attributed a number of starvation deaths.[7] In April 2001, the People's Union for Civil Liberties (PUCL) took the government of India to court, arguing in its writ petition that, "while on the one hand the stocks of food grains in the country are more than the capacity of storage facilities, on the other there are reports from various states alleging starvation deaths."[8] The PUCL documented that it was perfectly feasible for the government to widen a number of statutory food and nutrition programs, including the moribund midday meal scheme in schools. In response to this petition, on November 28, 2001, the Indian Supreme Court issued an interim order stating that "Every child in every government and government-assisted school should be given a prepared midday meal".[9]

Implementation of this and other Supreme Court orders lies in the hands of the relevant executive branch of government, which in this instance was state governments (Desai and Muralidhar, 2000). Midday meal implementation did not take place immediately or all at once, but over the next 5 years states across India implemented the program until, by 2006, every Indian state had instituted a free school lunch in primary schools. Appendix Table 1 documents the month and year of policy implementation in the 24 states and union territories used in our main analysis; the map in Figure A1 of the Online Appendix depicts geographic variation in the timing of implementation. Tamil Nadu, Gujarat, Puducherry and Kerala are excluded from this sample since their program implementation preceded the 2001 mandate, but we show in robustness checks that their inclusion does not alter our results. The table shows that there is considerable variation in the timing of implementation across different states. As we explain later, this will be key to our identification strategy.

It is worth noting that there appears to be no significant correlation between the timing of implementation and a number of observed state-level variables. This is evident in Table 1, which uses state-level data from 2001 to 2006—the window over which the program was introduced and implemented. The first 2 columns present the results of an OLS regression for states in our main sample in 2001.[10] This was the year of the Supreme Court directive, and the year before the earliest states implemented the program. The first 2 columns regress the year of midday meal implementation on a number of state-level covariates in this baseline year. These include economic indicators such as state domestic product (SDP) and education expenditure, as well as social and civic indicators such as the gender ratio, literacy rates, and voter turnout. None are statistically significant at the 5% level (only literacy is statistically significant at the 10% level, but the coefficient is close to zero.)

[6]Kerala responded with an opt-in program for public primary schools, leading to partial coverage. Tamil Nadu and Gujarat had, in 1982 and 1984 respectively, already instituted universal primary school midday meal programs. Most other states provided raw wheat or rice grains to enrolled children who attended school. Most accounts indicate that this system did not function well: grains were of poor quality, conditional attendance requirements were on paper alone (see for example Probe Team (1999)), and the system was plagued with leakage (see for example Muralidharan (2006)).

[7]There were 7 drought-affected states in 2001: Gujarat, Rajasthan, Maharashtra, Orissa, Madhya Pradesh, Chhattisgarh, and Andhra Pradesh (Down to Earth, Vol. 10, Issue 20010615, June 2001). They include both early and late implementers of midday meals.

[8]Rajasthan PUCL Writ in Supreme Court on Famine Deaths, PUCL Bulletin, November 2001.

[9]Supreme Court Order of November 28, 2001, Record of Proceedings Writ Petition (Civil) No. 196 of 2001.

[10]Daman & Diu as well as Manipur are excluded from this sample since they lack education expenditure data.

Program implementation:	Year[a]		Dummy[b]		
	(1)	(2)	(3)	(4)	(5)
Per capita SDP[c,g]	-0.139	0.077	0.003	-0.019	-0.036
	(1.921)	(2.147)	(0.008)	(0.028)	(0.028)
Per capita Education Expenditure[d,g]	0.003	0.000	-0.000	0.000	0.001
	(0.057)	(0.063)	(0.000)	(0.001)	(0.001)
Gender ratio[e,h]		0.065			
		(0.048)			
Literacy rate[h]		0.001*			
		(0.001)			
Voter turnout[e]		-0.017			
		(0.025)			
Year Fixed Effects	NO	NO	NO	NO	YES
State Fixed Effects	NO	NO	NO	YES	YES
Observations	22	22	132	132	132
R-squared	0.104	0.188	0.001	0.009	0.111

Table 1. Differences Between Early and Late Implementing States. *Notes.* This table presents regression estimates for a model in which the unit of observation is a state in columns 1-2 and state-year in columns 3-5, for the years 2001-2006 and 22 of the 24 states listed in Table 1. Data on education expenditure for Daman & Diu and Manipur were unavailable. Columns 1-2 present OLS estimates for the year 2001; column 3 presents a linear probability model; and columns 4-5 present (state) fixed effects estimates. *a.* Dependent variable = Year of program implementation $\in \{2002, 2003, 2004, 2005, 2006\}$ *b.* Dependent variable = 1 following midday meal implementation (0 otherwise). *c.* State Domestic Product. *d.* Revenue expenditure on education. *e.* Male/Female Ratio. *f.* Voter turnout in 2009 Lok Sabha (national parliament) elections. Election Commission of India. *g.* Source: National Accounts data from the Reserve Bank of India in 2004 prices, Population data from the Census of India, 2001. *h.* Source: Census of India, 2001

The timing of implementation in columns 3-5 is captured through a dummy variable equal to 1 following midday meal implementation (and 0 otherwise), and the data in these columns cover the full 2001-2006 window. Again, SDP and education expenditure are not significantly correlated with midday meal implementation, whether estimated via pooled OLS (column 3), state fixed effects (column 4) or state and year fixed effects (column 5).

In keeping with standard education policy practice, most of the funding for midday meals comes from the central government. Officially, the center and states share the total midday meal budget at a ratio of 75:25, respectively. In reality, central funds are supplemented to varying degrees by state governments. A systematic relationship between state supplements and the timing of program implementation would bias our treatment effect estimates upward (downward) if early implementers provided higher (lower) supplementary funds than late implementers. Poor documentation prevents us from actually estimating this correlation. However, anecdotal evidence suggests that there is no obvious relationship between state level supplements and timing of midday meal implementation. For example, Andhra Pradesh (which implemented the program in 2003) contributed Rs. 1 per child per day towards cooking costs in 2005, whereas in that year Rajasthan and Chattisgarh, which implemented earlier than Andhra Pradesh, contributed little towards cooking costs (Secretariat of the Right to Food Campaign, 2005). In another instance, Tamil Nadu—a 1982 implementer—and Uttar

Pradesh—a 2004 implementer—were both reported to have allocated the largest (state) budgets per child per day towards midday meals in 2009-2010 (Centre for Policy Research, 2013).

In what follows, we describe central government funding provisions, which are more transparently documented. Food grains are provided by the Food Corporation of India (FCI), an institution set up in 1964 to support the operation of the central government's food policies. Midday meal guidelines stipulate that each student be provided 100 grams of wheat or rice, 20 grams of pulses, 50 grams of vegetables and 5 grams of fat per day, amounting to a targeted total of 300 kilo calories and 8-12 grams of protein (MHRD, MHRD). As of 2009, this cost approximately Rs. 2.5 per student per day, including cooking costs.[11] In addition to the direct cost of food, the cost of labor and management, which include salaries paid to cooks and helpers, adds another Rs. 0.40, for a total cost of food equal to Rs. 2.90 per child per day. Of this, the central government provides Rs 2.17 and states are left to bear the remaining Rs. 0.63. Further, the central government provides a transport subsidy to carry grains from the nearest FCI warehouse to the primary school, up to a maximum of Rs. 75 per quintal, amounting to an average transport subsidy of Rs. 0.075 per child per school day. An additional budget of approximately 2% of total cost is assigned by the central government for the management, monitoring and evaluation of the program, amounting to an additional Rs. 0.045 per child per day. The total value of the central government subsidy therefore amounted to Rs. 2.30. This is approximately 5 U.S. cents per child per school day, or 10 USD per child per year.[12]

While the overall responsibility for program implementation rests with state governments, day-to-day operations lie in the hands of local government bodies, typically village governments (panchayats), who sometimes delegate implementation to local Parent Teacher Associations (PTAs) or NGOs. The meal itself is not extravagant. It is cooked at schools by cooks and their helpers, who are hired for this purpose. At around noon, children are served cooked rice or wheat, depending on the local staple, mixed with lentils or jaggery, and sometimes supplemented with oil, vegetables, fruits, nuts, eggs or dessert at the local level. The menu varies from place to place, but anecdotal evidence suggests that children generally enjoy the opportunity to sit with their peers and eat their midday meals (see, for example, Drèze and Goyal (2003)).

3. Data and Empirical Strategy

3.1. **Data.** Our data come from the Annual Status of Education Report (ASER), a yearly survey devoted to documenting the status of education among children in rural India. The data comprise a repeated cross-section. Each year, the survey is conducted around October and covers a random sample of 20-30 households per village in 20 villages in each of India's roughly 580 rural districts.

What makes ASER truly unique is that it tests all children in the household between the ages of 5 and 16 for reading and math proficiency using rigorously developed testing tools. The fact that the survey is administered in households rather than in schools is useful because it enables an assessment of learning outcomes regardless of school participation. Appendix Figure 1 depicts ASER's English language tests in reading and math. In practice the tests are administered in vernacular languages.

[11]The information on cost of providing midday meals was obtained from http://mdm.nic.in/, accessed on 20th March, 2016. The figures quoted here reflect the cost for India excluding the North Eastern States. In the case of North Eastern States, the central government bears a higher fraction of the total costs.

[12]We use the exchange rate for October 2009, 47.4 INR to a USD, to match with the reference date of the cost estimates 10 USD = 0.05×200 school days per year.

The reading assessment has 4 levels of mastery: letters, words, a short paragraph (grade 1 level text), and a short story (grade 2 level text). Similarly, the math assessment consists of four levels: single-digit number recognition, double-digit number recognition, two-digit subtraction with carry over, and three digit by one digit division. For both tests separately, the child is marked at the highest level he or she can do comfortably with scores ranging from 0 to 4. A score of 0 means that the child cannot do even the most basic level and a score of 4 means that he or she can do level 4 in the respective subject.

We use data from 8 available household cross-sections, from 2005–2012. Our sample comprises primary school-aged children. In India, primary school typically runs from grade 1 to grade 5 and officially corresponds to children aged 6-10. We restrict our sample to this age group for two reasons. First, the Supreme Court mandate pertained to primary schools, which cover precisely this age group. Second, some localities offer feeding programs to younger children, in "Anganwadis" that care for preschool children, or older children in secondary schools. While there is no systematic pattern across states in these offerings, including younger and older children in the sample would run the risk of "mis-allocating" children to the control group when, in fact, they received school feeding.[13]

Since the program itself is likely to have had an enrollment effect, we include all children in this age group who were either enrolled in public school or were not enrolled in school. Dropping non-enrolled children does not change our results since almost all children are in fact enrolled in school over this sample period, but doing so would subject our results to sample selection bias in this ITT framework. We do not include private school students in our sample, both because the policy mandate did not apply to this group and because previous work has found that the program introduction in public schools did not result in children switching from public to private schools (see Jayaraman and Simroth (2015)). Our results are, however, robust to the inclusion of private school students in our main sample.

We further restrict our attention to the states listed in Appendix Table 1, which were subject to the Supreme Court mandate; in additional robustness checks, we add earlier implementers. Altogether, our main sample of 6-10 year olds comprises roughly 1.24 million children in 24 states and union territories, averaging about 150,000 observations in each cross section. The data display rich temporal and geographic coverage; see Online Appendix Table A1. State-wise sample sizes obviously vary, reflecting their differing populations, but robustness checks in which we drop states one-by-one indicate that no one state is driving our results.

Table 2 furnishes summary statistics for each of the 8 survey years. The first two rows denote average reading and math scores. These scores measure learning achievement and will be the main outcomes of interest in our empirical analysis. The scores take integer values ranging from 0 (inability to do even the most basic level) to 4 (mastery of the highest level). In our main analysis, the dependent variable will be this raw integer test score. In other words, we treat test scores as interval scales. Later, in our robustness checks, we discuss the limitations of this approach and use an alternative measure, which accounts for the ordinal nature of the test score variable.

The average scores for both reading and math hover at around 2 during the observation period. Concretely this number means that, on average, primary school-aged children can read words but not a

[13]When we extend the sample to include 5-11 year-olds, our estimates are qualitatively similar in terms of sign and statistical significance, and the point estimates indicate, if anything, larger treatment effects for each of the specifications in both reading and math.

	Survey Year								
	2005	2006	2007	2008	2009	2010	2011	2012	Overall
Reading Score	2.05	2.05	2.20	2.14	2.21	2.17	1.96	1.80	2.09
	(1.47)	(1.39)	(1.33)	(1.36)	(1.31)	(1.32)	(1.35)	(1.40)	(1.37)
Math Score	1.91	2.06	2.06	1.96	2.08	2.00	1.77	1.61	1.95
	(1.42)	(1.40)	(1.24)	(1.25)	(1.23)	(1.21)	(1.19)	(1.16)	(1.28)
Enrollment	95.56	94.21	97.41	96.89	97.38	97.89	97.87	97.94	96.82
	(20.59)	(23.35)	(15.88)	(17.37)	(15.97)	(14.37)	(14.43)	(14.21)	(17.56)
Dropout	1.24	1.15	0.97	0.94	0.77	0.69	0.54	0.56	0.88
	(11.06)	(10.68)	(9.80)	(9.63)	(8.73)	(8.25)	(7.36)	(7.44)	(9.32)
Never Enrolled	3.20	4.63	1.62	2.18	1.85	1.42	1.58	1.51	2.31
	(17.59)	(21.02)	(12.62)	(14.59)	(13.48)	(11.85)	(12.48)	(12.18)	(15.02)
Age	8.16	8.15	8.18	8.18	8.22	8.22	8.19	8.17	8.18
	(1.42)	(1.43)	(1.40)	(1.41)	(1.41)	(1.41)	(1.41)	(1.41)	(1.41)
Female	46.02	46.59	46.65	47.47	46.62	46.85	48.09	49.52	47.12
	(49.84)	(49.88)	(49.89)	(49.94)	(49.89)	(49.90)	(49.96)	(50.00)	(49.92)
Household	7.34	7.82	6.69	6.83	6.44	6.41	6.59	6.73	6.87
Size	(4.23)	(4.64)	(3.49)	(3.14)	(2.80)	(2.85)	(3.08)	(3.25)	(3.55)
Exposure in Months	15.00	20.51	26.31	29.01	30.65	30.69	30.23	30.05	26.49
	(9.60)	(11.26)	(13.53)	(15.78)	(16.93)	(16.90)	(16.95)	(16.88)	(15.73)
Exposure in Years	0.72	1.26	1.78	2.04	2.22	2.22	2.19	2.17	1.82
	(0.96)	(0.96)	(1.10)	(1.27)	(1.41)	(1.41)	(1.41)	(1.41)	(1.34)
No. Observations	125,960	184,628	198,321	173,711	162,829	149,564	132,768	111,000	1,238,781

Table 2. Summary Statistics by Survey Year. *Notes.* Each cell in this table (except in the last row) contains means, with standard deviations in parentheses. Enrollment, Dropout, Never Enrolled and Female are reported as percentages.

grade 1 level paragraph or grade 2 level story; they can recognize double digit numbers but cannot do two digit subtraction or divide a 3 digit number by a 1 digit number.

In addition to administering these tests, ASER collects information regarding the child's current school enrollment status (though, unfortunately, not class attendance), as well as some basic demographic information pertaining to their age in years, gender and household size. Consistent with official estimates, net enrollment remains at a pretty steady 97% during the observation period. Among out-of-school children, approximately one-third are dropouts and two-thirds have never enrolled in school. The average age of children in our sample is around 8. Just under half are female, and the average household size is roughly 7.

3.2. **Empirical Strategy.** We use an ITT framework in which we examine the effect of potential midday meal exposure on test scores. This, rather than the average treatment effect on the treated (ATT), is particularly relevant in this context because it examines the overall effect in the full population of eligible cohorts, for an extant program. Variation in potential program exposure is jointly determined by a child's age at the time of program implementation, and their age at the time of observation. Both depend on the child's birth cohort. For a given birth cohort, the latter depends on the survey year; the former depends on the timing of implementation in the child's state.

For example, consider a child born in 1996 in Andhra Pradesh, which implemented the program in 2003. This child was 7 years old when the program was implemented. In 2005 (the first survey year), she was 9 years old. At that stage, she had up to 3 years of program exposure (between grades 2 to 4). In 2006, she had up to 4 years of exposure, which is the maximum she could have had, since this marks her final year of primary school. Another child, also born in 1996 but in a different state—say, Rajasthan, which implemented the program in 2002—would have had up to 4 years of exposure in 2005, and her maximum of 5 years of exposure in 2006. Exogenous variation within the same cohort, observed in the same survey year, comes from the fact that the 2001 Supreme Court mandate was implemented in pubic primary schools in a staggered manner across Indian states between 2002 and 2006.

In fact, since different states introduced the program in different months, we have variation in months rather than simply years of exposure. The Indian school year typically starts in June and children are officially supposed to be enrolled in grade 1 in the year they turn 6. The ASER survey is conducted in September-November each year. The precise month varies and is not systematically recorded, so we take the median, October, which is when most surveys are conducted. Depending on the month and year of program implementation a child can in principle have anywhere between 0 (if a 10 year-old was examined in the survey just after the program was implemented) and 52 months (if the child is 10 years old and has had 4 full years of exposure plus 4 months in grade 5 from June, when the academic year starts, to October when the test is administered) of program exposure. Online Appendix B provides a detailed description of how we construct the months of exposure variable based on a child's current age and her age at the time of midday meal introduction.

The bottom rows of Table 2 show that on average, children in the sample have 26 months, or roughly 2 years, of policy exposure. This is obviously lower for earlier survey years given that the policy was implemented between 2002 and 2006. We will account for this difference in our empirical analysis by including survey year fixed effects.

Figure 1, which presents a scatter plot of months of potential exposure against birth year, demonstrates that there is considerable variation in year of birth (and thereby, age at first treatment) within each exposure level. Three remarks on this figure are in order. First, all the children in our sample have at least 4 months of program exposure. This follows from the fact that ASER commenced its surveys in 2005 after all major states had already instituted the program. Second, as the megaphone-like shape of the data indicates, older and younger children tend to have less exposure than others. We account for this in our empirical analysis by accounting for birth year. Third, there is a natural "lumpiness" in the data at 4, 16, 28, 40 and 52 months of exposure. Each of these months contain between 14-22% of the children in the main sample. This follows from the fact that the survey was conducted 4 months after the school year commences. So, for example, a 6-year-old child will have had 4 months of potential exposure if the program was instituted before June of the current year, a 7-year old child will have had 16 months of exposure if it was instituted before June of the previous

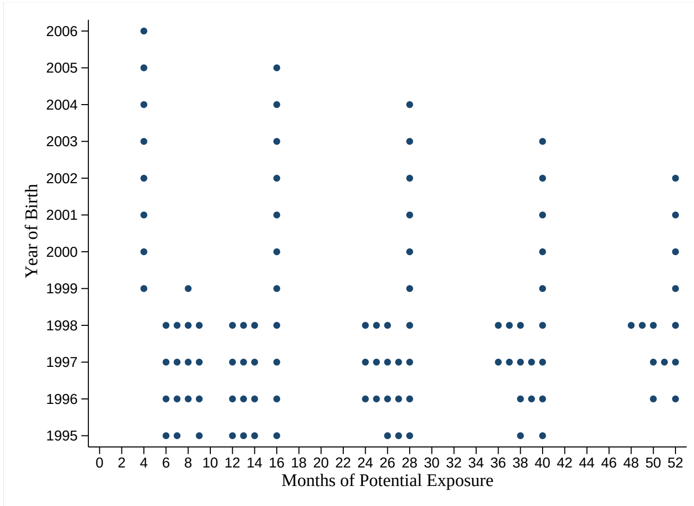


Figure 1. Variation in Months of Potential Exposure by Birth Year. *Notes:* This graph depicts the variation in our data in months of exposure (x-axis) by year or birth (y-axis).

year, and so forth.[14] We start the empirical analysis by exploiting variation in months of exposure. However, in order to account for the lumpiness in the data and reduce the potential for measurement error, we also report results by years of exposure and show that this does not alter our qualitative results.

We begin by estimating the following baseline model, which exploits variation in months of program exposure generated by the survey year (time of observation); a child's birth cohort; and the timing of policy introduction in his or her state:

$$y_{itcs} = \alpha + \beta \cdot \text{Exposure}_{i(tcs)} + \phi \text{Controls}_{itcs} + \delta_t + \delta_c + \delta_s + \gamma_s t + \varepsilon_i \quad (1)$$

where y_{icst} measures the reading or math test score of child i, surveyed in year t, belonging to birth cohort c, and residing in state s. The Exposure variable captures months of *potential* program exposure. In principle, this could vary anywhere from 0 months if the child has never been exposed to the program, and 60 months for children who have the full 5 years of exposure all through primary school. In these data, it varies between 4 and 52 months. Our parameter of interest is β: it is our ITT estimate which captures the treatment effect of potential exposure to midday meals on test scores. Control variables include gender, household size and a dummy variable for whether or not the child's mother attended school. The parameters δ_t, δ_c, and δ_s account for differences in test outcomes by time (i.e. survey year), birth cohort, and state, respectively. The parameter γ_s is a linear state-specific time trend, which allows for the linear evolution of test scores over time to vary by state.

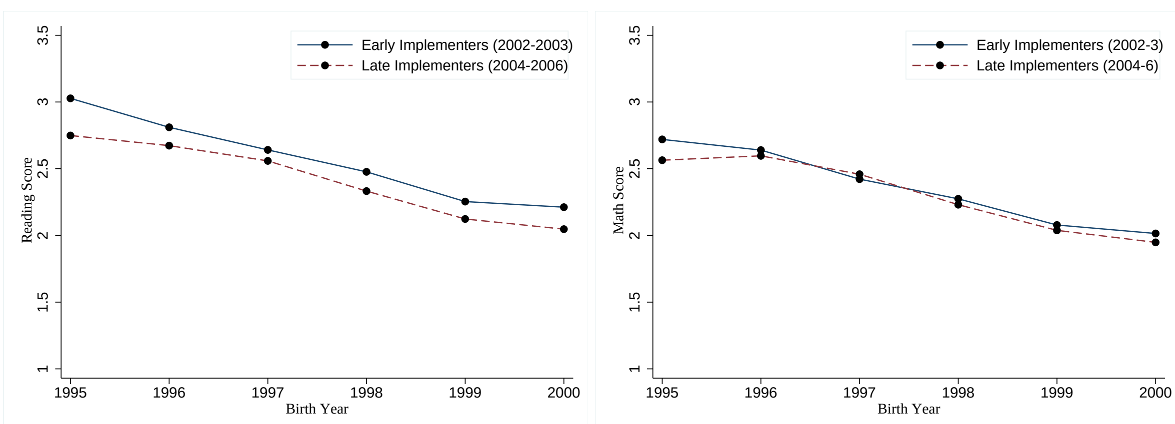
This empirical specification allows us to control for any systematic shocks to outcomes, which are correlated with but not attributable to program exposure across three dimensions. First, survey timing (captured through δ_t) is important because there may be natural variation in test scores over time and, as we saw in Table 2, children surveyed in earlier years naturally have lower levels of program exposure given that midday meals were implemented between 2002 and 2006. Second, cohort effects (δ_c) are relevant because it is natural to expect older children to have more exposure than younger

[14]This can be seen clearly upon examination of row 1 of Online Appendix Table B1.

children and perform better in these tests. Third, differences across states are pertinent because, although Table 1 indicates that there is no correlation between the timing of implementation and many observable state characteristics, there may still be some unobserved differences between early and late implementers. Including state fixed effects (δ_s) captures these unobserved time-invariant differences.

We may still worry that the timing of implementation is correlated with *trends* in test scores. State specific time trends (captured through γ_s) account for this possibility in part. We also provide supportive evidence to allay concerns about underlying and pre-existing state-cohort trends in test scores. First, in Figure 2 we investigate test score trends among children in our main sample who were born prior to the 2001 Supreme Court mandate. It shows that early implementers exhibit slightly better test scores than late implementers. This difference in levels is accounted for by state fixed effects, and is natural since early implementers are likely to be states with better governance and institutions. It also shows that older children have better test scores than younger children. This is accounted for by cohort fixed effects. Importantly for us though, for cohorts born before the Supreme Court Mandate, early and late implementers exhibit parallel trends in test scores.[15]

Second, we conduct a falsification analysis in Section 6 using older cohorts who had completed primary schooling (i.e. were aged 12 and above) at the time of policy implementation. The results of this analysis once again supports the absence of any state-cohort trends in the cohorts which completed primary schooling before the policy implementation. Third, again in Section 6, we show that our results are robust to the inclusion of cohort-state fixed effects, which accounts for the possibility that timing of implementation may have been driven by the (under)performance of particular cohorts within a state.



(a) Reading Score

(b) Math Score

Figure 2. Parallel Trends. *Notes:* This graph depicts the trends in the reading and math score by early implementers (states implementing the policy in 2002-2003) and late implementers (states implementing the policy in 2004-2006) for cohorts born prior to the 2001 policy mandate.

[15]Online Appendix Figures A2 and A3 show that the parallel trend assumption depicted here is robust to the use of alternate learning achievement indicators described in Section 6.1.

We estimate equation (1) using OLS.[16] While this allows for conventional interpretations of the coefficient, we acknowledge that our treatment of the ordinal variable as if it were an interval variable raises a number of issues. We discuss and deal with these in Section 6.

4. The Effect of Midday Meals on Test Scores

In this section, we examine the effect of midday meal exposure on test scores. We present ITT estimates. Standard errors are clustered throughout by state and year of birth; the results are also robust to clustering by state.

In our raw data, program exposure is positively correlated with learning; see Online Appendix Figure A4. Children with the lowest level of program exposure (4 months) have an extremely low average reading and math test scores of about 1.07. Concretely, on average, these children just about read a letter and recognize a one-digit number. Columns 1 and 5 of Table 3 indicate that, from this baseline, average test scores increase steadily by about 0.035 points for reading and 0.030 points for math with each additional month of exposure. Consequently, average test scores for children with 52 months of exposure (the maximum in our sample) are almost 3 times as high as they are for children with only 4 months of exposure: on average, these children can read a short paragraph and conduct two-digit subtraction with carryover.

This positive correlation, though large in magnitude, is likely to be an upward biased estimate of the true causal relationship between midday meal exposure and learning, since it captures differences across time, cohorts, or states. More specifically, children surveyed in later years, belonging to older cohorts, and residing in states which implemented the policy earlier are likely to have both longer exposure and higher test scores.

We account for this in columns 2-4 and 6-8 of Table 3, which present OLS estimates for equation (1) for reading and math test scores, respectively. Row 1 presents the ITT estimate, β, which corrects for state, cohort, and time fixed effects, as well as state-level time trends. The effect of midday meals on test scores is positive and statistically significant at the 1% level. In keeping with our priors, this treatment effect is substantially smaller than the simple linear association. The point estimates in columns 2 and 6, which present the baseline treatment effect without controls are roughly one-fourth the size of that in column 1 for reading and one-seventh the size of that in column 5 for math.

This treatment effect is qualitatively robust to the inclusion of additional controls in columns 3 and 7. Controlling for these variables entails sample loss, and differences in point estimates and the loss of statistical significance for the math score are likely due to this. The coefficients of the controls themselves are largely in keeping with our priors. Test scores are lower for girls than they are for boys. Children in larger households perform worse, probably because these households also tend to be poorer. And children whose mothers have attended school do considerably better than children whose mothers haven't. These regressions nevertheless demonstrate that the results are qualitatively robust to the inclusion of these controls. The estimates that follow will therefore use the full sample, eschewing these controls.

[16] Ordered probit and ordered logit and probit estimates produce qualitatively similar results.

	Reading Score				Math Score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure (β)	0.0352***	0.0081***	0.0059***	0.0175***	0.0300***	0.0042**	0.0024	0.0122***
(in months)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Exposure^2				-0.0002***				-0.0001***
				(0.000)				(0.000)
Female			-0.0265***				-0.0552***	
			(0.007)				(0.006)	
Household Size			-0.0025***				-0.0023***	
			(0.001)				(0.001)	
Mother Attended			0.3995***				0.3603***	
School			(0.016)				(0.016)	
State FE	NO	YES	YES	YES	NO	YES	YES	YES
Birth Year FE	NO	YES	YES	YES	NO	YES	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
State×Trend	NO	YES	YES	YES	NO	YES	YES	YES
Mean at 4 months	1.072	1.072	1.061	1.072	1.061	1.061	1.068	1.061
Observations	1,238,781	1,238,781	1,048,509	1,238,781	1,238,781	1,238,781	1,048,509	1,238,781
Adjusted R-squared	0.163	0.273	0.302	0.274	0.136	0.264	0.297	0.265

Table 3. Effect of Midday Meal Exposure on Test Scores *Notes.* This table presents OLS estimates for equation 1. Exposure measures months of potential program exposure. The dependent variable is the reading test score (columns 1-4) and the math test score (columns 5-8); they take integer values ranging from 0 to 4. Each column represents a different regression. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors in parentheses are clustered by state and year of birth.

Columns 4 and 8 allow for a non-linear treatment effect by adding squared months of exposure to the baseline specification.[17] The estimates in rows 1 and 2 show that the effect of program exposure on test performance is increasing in the first 3 years of exposure and then tapers off in the last 2 years of primary school.

To understand the magnitude of these effects, we aggregate exposure in yearly intervals (where 0-1 years is 0-12 months, 1-2 years are 13-24 months, etc.). This has three advantages over the monthly exposure measure. First, it is more natural to think of children in primary school with years as opposed to months of exposure, given that grade promotion occurs annually and primary school extends over the course of 5 years. Second, exposure measured in years rather than months is less "lumpy" (see Figure 1), and this allows us to both avoid out-of-sample predictions for months of exposure for which we have no observations and provides us with enough observations within each year of exposure to estimate confidence intervals for marginal effects. Finally, it facilitates the interpretation of results in the next section, where we explore what may account for the learning effects we estimate in this section.

We estimate the following equation with **Exposure** measured through a vector of 4 dummy variables denoting 1-2, 2-3, 3-4, and 4-5 years of exposure, with 0-1 years of exposure being the exclusion:

$$y_{itcs} = \alpha + \beta' \mathbf{Exposure}_{i(tcs)} + \delta_t + \delta_c + \delta_s + \gamma_s t + \varepsilon_i \tag{2}$$

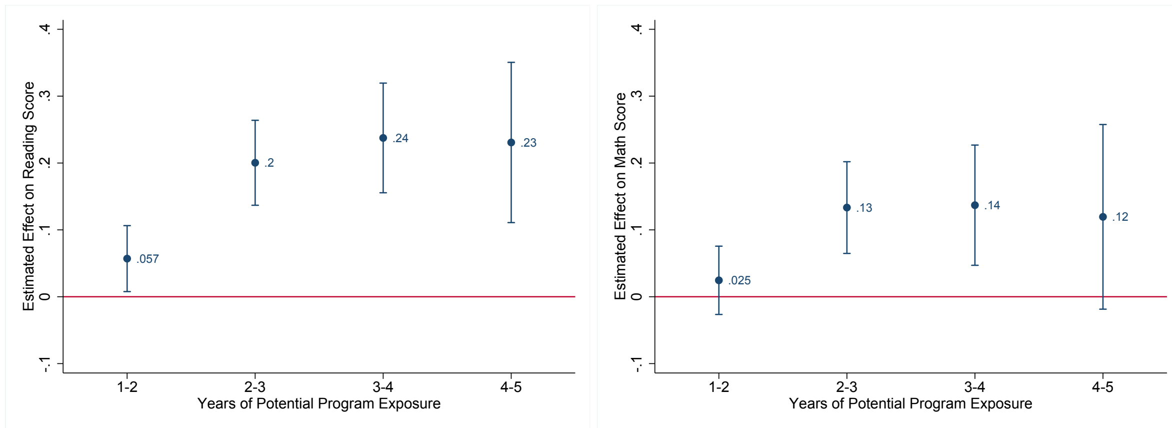
where y_{itcs} is the test score, so the coefficient estimates for the vector of yearly exposure dummies β' capture the change in test scores as a result of up to one additional year of exposure. The remaining variables are defined as in equation (1).

Figure 3 depicts OLS estimates for β' in equation (2) graphically; regression results are presented in Online Appendix Table A2. It confirms what we saw in the final results of Table 3, namely that learning increases, albeit at a decreasing rate, with exposure to midday meals. In the second year of exposure, test scores increase by a statistically significant 0.057 points for reading, which amounts to an approximately 4.4% increase relative to the baseline (children with less than 1 year of exposure). For math, the increase in test scores is half this size and statistically insignificant.

Test scores jump dramatically in the third year with a 0.20 point (15%) increase in reading and a 0.13 point (10%) increase in math, relative to the baseline. This increase in test scores from the second to the third year of exposure is not just economically, but also statistically significant (p=0.0 for both reading and math). The increase jumps slightly to 0.24 (i.e. by 18%) for reading and 0.14 (11%) for math in the fourth year of exposure although the difference relative to three years of exposure is only marginally significant for reading (p=0.08) and statistically insignificant for math (p=0.87).

In the final year of exposure the effect tapers off slightly to 0.23 points for reading and 0.12 points for math. This represents a statistically significant increase relative to the baseline for reading (p=0.0) and math (p=0.09), although the difference is statistically insignificant relative to the previous two years. The larger confidence intervals in the last year of exposure (4-5) reflects a loss of statistical power arising from the smaller sample size in this last group, since survey timing forces us to censor the data at 52 months rather than the 60-month end of the full 5 years of primary school.

[17] The results for this quadratic specification are broadly consistent with the introduction of higher order polynomials in this regression, as well as semi-parametric estimation. (Results not shown.)



(a) Reading Score

(b) Math Score

Figure 3. Effect of Midday Meal Exposure on Test Scores by Years of Potential Exposure. *Notes:* This figure provides a graphical depiction of the OLS estimates for β' in equation (2). The exclusion is 0-12 months (i.e. less than 1 year) of potential exposure; 1-2 years correspond to 13-24 months, 2-3 correspond to 25-36 months, and so on. Coefficient estimates for the change in test scores from up to one additional year of exposure are denoted in the graph, and the bars denote the corresponding 95% confidence intervals, with standard errors clustered by state and year of birth. The full regression results corresponding to this figure are presented in Online Appendix Table A2.

According to these estimates, a child who has been exposed to midday meals throughout primary school has reading test scores that are 18% (0.17σ) higher and math test scores that are 9% (0.09σ) higher than those of a child with less than one year of exposure.[18] As we discuss in more detail in Section 6.1, the increase in reading scores reflects a significant improvement in the proportion of children who have achieved levels 2-4 in this subject, whereas that in math reflects increases in the proportion who have achieved levels 2-3.

In sum, relative to the (up to) one year baseline, the increase in test scores is small in magnitude, and in the case of math, statistically insignificant, after up to two years of exposure. Thereafter, it is large and significantly higher. This is important in view of the negligible learning effects of school feeding programs documented in the literature to date. In particular, the extant literature has—without exception—examined program effects after *at most* two years of exposure. Our results suggest that students may need prolonged exposure in order to reap substantive learning benefits from the program.

[18]Note that interpreting the increase in test scores as percentage improvements imposes the implicit assumption of test scores being on a ratio scale, i.e. the student scoring 2 knows twice as much as the student scoring 1; the student scoring 4 knows twice as much as the student scoring 2 and so on.

5. Accounting for Improved Test Scores

The analysis in the previous section shows that midday meals have a positive and statistically significant impact on learning achievement. The literature has stressed two avenues by which school feeding programs may accomplish this. The first is increased school participation, which provides children the opportunity to learn in the first place. The second is through improved nutrition: better nourished children have more learning capacity and therefore perform better in school.[19]

Unfortunately, two things prevent us from directly exploring these channels. First, we have neither nutrition nor attendance data. Second, over our period of observation, we have no pure control group since all children have at least 4 months of program exposure. In other words, attendance, enrollment, or nutritional status may well have been lower in the absence of this program. However, we cannot identify this effect because we do not have a pure counterfactual. Previous studies, which cover earlier time periods during which some states were yet to introduce midday meals, have estimated substantial effects on school participation.[20] In Online Appendix C we apply these estimates to an accounting exercise which disaggregates the total learning effect into a participation-effect and a nutrition-learning effect, to place an upper bound on the nutrition-learning effect.[21] Beyond this, there is not much we can say with the data at hand.

In the remainder of this section, we explore three further channels which may account for the treatment effects documented in the previous section. The first is complementary inputs. School attendance and better nourishment doesn't automatically foster learning. Children presumably need to learn reading and math in class in order to answer reading and math questions. Section 5.1 explores this by estimating potential complementarities between program exposure and various schooling inputs.

Second, school lunches are likely to be more effective in improving the performance of more disadvantaged children because they are more likely to enjoy nutritional improvements as a result of the program and are likely to have higher marginal benefits of improved nutrition since they start from a lower baseline nutritional status. Section 5.2 explores this by examining heterogeneous treatment effects based on two measures of socio-economic status: gender and housing assets.

Finally, school lunches only improve the nutritional status of children to the extent that families do not fully substitute away food allocations from program recipients to other family members. Section 5.3 explores this by examining whether children living in households that may be more likely to redistribute resources away from them, benefit less from midday meal exposure.

[19]See, for example, Adelman et al. (2008), Kristjansson et al. (2007), Bundy et al. (2009), Behrman et al. (2013), Jomaa et al. (2011), Alderman and Bundy (2012) Lawson (2012), and McEwan (2015), for recent reviews of this literature in the context of developing countries.

[20]Jayaraman and Simroth (2015) report that the introduction of midday meals increased grade 1 enrollment by approximately 25 per cent. Afridi (2010) finds that in Madhya Pradesh, the program increased girls' grade 1 attendance by 10 percentage points.

[21]We estimate the upper bound on the nutrition-learning effect to be 0.32 (0.23σ) for reading and 0.17 (0.13σ) for math. Under the assumptions spelled out there, this is effectively what the (maximum) increase in reading and math scores would be if cognitive skills of newly enrolled children were unchanged, and the improvement in test scores came entirely from a nutrition-learning channel for children who would be enrolled in school whether or not they received a school lunch.

5.1. **Complementary Inputs.** It is unlikely that school lunches work in isolation. For instance, if teachers are frequently absent from school then the program may encourage children to go to school and may improve their nutritional status, but they are unlikely to learn much once they are there. In general, it seems plausible that schooling inputs that directly foster learning—such as teachers, books, or blackboards—serve to translate higher school participation and nutritional status arising from school feeding programs into improved cognitive skills.

From 2009-2012 ASER contemporaneously surveyed a public school in each village where they conducted household surveys.[22] This allows us to explore potential complementarities between schooling inputs and midday meal exposure. Hence, in this subsection we restrict the sample to 2009-2012 and match the school survey to the household survey data at the village level. Fewer survey years and missing information on schooling inputs, has the drawback that we are only able to match roughly 40% of the children in our main sample, that too only for later years.

Nonetheless, using this matched sample we show, in Section 6, that the results of our main specification with months of exposure are robust to the inclusion of a wide array of schooling inputs. Here, for ease of interpretation, we measure exposure linearly in terms of years rather than months, and investigate the presence of potential complementarities between schooling inputs and midday meals by estimating the following model:

$$y_{itcvs} = \alpha + \beta \text{Exposure}_{i(tcs)} + \phi \text{Input}_{vts} + \theta(\text{Exposure}_{i(tcs)} \times \text{Input}_{vts}) + \delta_t + \delta_c + \delta_s + \gamma_s t + \varepsilon_{itcvs} \tag{3}$$

where Exposure $= 1, 2, ...5$ measures the linear years of potential program exposure and Input denotes a schooling input in village v for the government school surveyed in that village. The remaining variables are defined as in equation (1). Our parameter of interest is θ, which captures potential complementarities between program exposure and schooling inputs. If children attend school more frequently and are better nourished on account of midday meal exposure, they are more likely to benefit more from these inputs in the learning process. This would be consistent with $\theta > 0$.

We examine complementarities between program exposure and six separate schooling inputs.[23] Teacher attendance refers to the number of teachers present in school on the day the ASER school survey took place, as a fraction of the total number of appointed teachers. Usable Blackboard is a dummy variable reflecting the presence of at least one usable blackboard in either grade 2 or grade 4. Learning Material indicates the availability (or not) of supplementary learning materials, such as books, in the school. Separate Classroom is a dummy indicating whether grade 2 and grade 4 are taught along with other grades or not. Tap in School indicates whether or not the school has a functioning drinking water tap. No. Classrooms indicates the total number of usable classrooms in the school.

[22]While ASER started the school surveys in 2007, the first round has little comparability to subsequent rounds which provide a much more comprehensive list of schooling variables.

[23]ASER reports a long list of schooling variables from which we choose a subset. Our choice of variables is driven both by the fact that they have the fewest missing observations, and also because they are relevant learning inputs.

Input	Teacher Attendance	Usable Blackboard	Learning Material	Separate Classroom	Tap in School	No. Classrooms
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Reading			
Exposure (β)	0.483***	0.314***	0.313***	0.335***	0.329***	0.442***
	(0.081)	(0.051)	(0.048)	(0.050)	(0.060)	(0.072)
School Input(ϕ)	0.061***	0.009	0.018	0.084***	0.050***	0.011***
	(0.023)	(0.014)	(0.014)	(0.014)	(0.016)	(0.004)
Exposure $\times$ Input (θ)	0.039***	0.023***	0.035***	-0.010	0.008	0.001
	(0.010)	(0.006)	(0.006)	(0.007)	(0.007)	(0.001)
Adjusted R-squared	0.288	0.288	0.289	0.288	0.294	0.293
			Panel B: Math			
Exposure (β)	0.308***	0.180***	0.176***	0.196***	0.220***	0.267***
	(0.074)	(0.051)	(0.049)	(0.051)	(0.064)	(0.059)
Input(ϕ)	0.055***	0.019	0.016	0.074***	0.037**	0.011***
	(0.019)	(0.014)	(0.014)	(0.012)	(0.015)	(0.003)
Exposure $\times$ Input (θ)	0.041***	0.018***	0.031***	-0.005	0.010	0.001
	(0.009)	(0.006)	(0.006)	(0.006)	(0.007)	(0.001)
Adjusted R-squared	0.297	0.294	0.295	0.294	0.299	0.300
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
State×Trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	460,058	435,076	435,076	435,076	340,727	406,240

Table 4. Complementarity with Schooling Inputs *Notes.* This table presents OLS estimates for β, ϕ and θ from equation (3), estimated separately for each schooling input and separately for reading (Panel A) and math (Panel B). Exposure is measured in terms of linear years of potential program exposure. The relevant Input in each row corresponds to the element listed in the column headings. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. Standard errors in parentheses are clustered by state and year of birth.

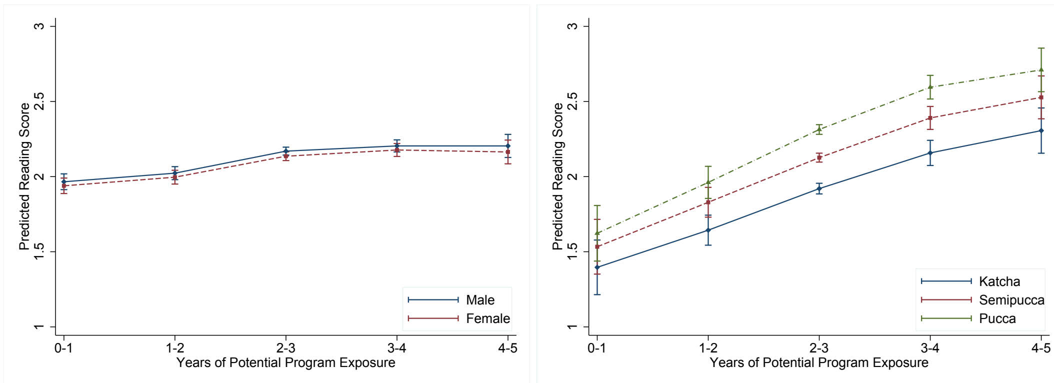
The findings are reported in Table 4, which presents OLS estimates for β, ϕ and θ from equation (3), estimated separately for each schooling input, separately for reading (Panel A) and math (Panel B). These results are robust to correcting for multiple hypothesis testing; See Online Appendix Table A3. They suggest the presence of significant complementarities with respect to those teaching inputs that are directly related to learning opportunities of children. For instance in column 1, we see that a 10 percentage point increase in teacher attendance is associated with a 0.006 (0.005) point increase in reading (math) scores on its own. However, when combined with one additional year of exposure to school lunches, a 10 percentage point increase in teacher attendance is associated with a roughly 0.01 point increase in reading and math scores. Although access to a functional blackboard (column 2) or supplementary learning material (column 3) don't by themselves improve test scores, when combined with midday meals they improve both reading and math test performance.

By contrast, in columns 4-6 we see that more general schooling infrastructure like the availability of separate classrooms, access to drinking water tap, or the total number of classrooms, is not complementary to midday meals. Together these results suggest that schooling inputs that are used in classroom instruction are complements to midday meals, but more general schooling infrastructure is not. What we interpret as complementarity may, however, be a reflection of differential funding for, and quality of, program implementation. More specifically, it is possible that early implementers are also states that have also more generously invested in midday meals and other schooling inputs. This seems unlikely given the anecdotal evidence discussed in Section 2. Furthermore, as we will see in Section 6.6, baseline program effects are not affected by the inclusion of schooling inputs as controls. Nevertheless, we cannot rule this out in the absence of state-level funding information.

5.2. **Heterogeneous Treatment Effects.** The efficacy of school feeding programs in improving learning achievement depends on whether they improve attendance and nutrition, and whether this translates into better school performance. On the one hand, there are a couple of related reasons why disadvantaged children may derive greater benefit from the program's nutritional benefits than more privileged children. First, as Afridi (2010) has documented, midday meals are more likely to increase the nutritional intake of disadvantaged children. Second, since poorer children start from a lower nutritional baseline, the marginal benefits of improved nutrition are likely to be larger for them than for more privileged children who tend to have better nutritional status; see Strauss and Thomas (1998) and Strauss (1986) who document an increasing concave relationship between nutrition and productivity. On the other hand, disadvantaged children may be less well positioned to take advantage of these nutritional benefits because they face larger barriers to (regular) school attendance.

Figure 4 investigates the presence of heterogeneous treatment effects for reading scores along two dimensions, namely, gender and housing assets; analogous results for math scores are presented in Online Appendix Figure A5. Female disadvantage in terms of educational outcomes has been well-documented for India; see for example, Kingdon (2002, 2007). Following the logic outlined above, we would expect baseline test performance to be lower for girls than for boys, but for girls to be more responsive to program exposure. The focus of the ASER survey is on testing children, and as a consequence information on economic status is rudimentary. Still, enumerators do record some proxies for wealth for the years 2008-2012, the most complete of which is housing assets.[24] This

[24]Patterns are similar for other measures of economic status, such as a broader asset index constructed using principal components analysis. However, reduced sample sizes due to missing observations on these indicators preclude the calculation of confidence intervals for marginal effects.



(a) Gender **(b)** Housing

Figure 4. Heterogeneous Responses: Reading. *Notes:* This graph depicts predicted reading test scores for different years of potential exposure by gender (panel a) and housing assets (panel b). Bars denote 95% confidence intervals, with standard errors clustered by state and time.

comprises a record of the material from which a house is made, where "Pucca" denotes a house made of durable materials such as brick, stones or cement; "Kutcha" denotes a house made of less durable materials such as mud, reeds, or bamboo; and "Semipucca" denotes something in between. Hence, Pucca (Kutcha) is a proxy for relatively high (low) economic status. Here again, we expect that children living in Pucca houses have better baseline performance than children living in poorer quality housing, but that the increase in test scores with exposure is larger for the latter, more disadvantaged, group relative to the wealthier former group.

Figure 4 shows that our first prior is confirmed: girls perform worse than boys, as do poorer children (those living in Semipucca or Kutcha housing) relative to wealthier children. However, there is no evidence that disadvantaged children enjoy higher marginal benefits from program exposure. This "negative" result is likely to reflect three realities. First, these are crude measures of disadvantage compared to measures like consumption expenditure or (better yet) baseline caloric intake; this may mask differences in marginal effects of program exposure. Second, these children are starting from a very low baseline in terms of nutritional status. Deaton and Drèze (2009) report that three quarters of the Indian population lives in households whose per capita calorie consumption lies below "minimum requirements" and that even privileged Indian children are mildly stunted. It is possible, in this context, that marginal effects of nutritional input are high, and roughly comparable, for both relatively privileged and relatively disadvantaged children. Finally, these children may not be able to reap potential nutritional benefits of the program because they are unable to attend school with any regularity.

	Reading				Math			
	Full Sample	Full Sample	Full Sample	Female Only	Full Sample	Full Sample	Full Sample	Female Only
R =	HH Size	At least 1 sibling	Sibling without MDM	Male sibling	HH Size	At least 1 sibling	Sibling without MDM	Male sibling
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure (β)	0.076***	0.069***	0.083***	0.070***	0.044**	0.038**	0.050***	0.041**
	(0.016)	(0.016)	(0.015)	(0.016)	(0.018)	(0.018)	(0.017)	(0.017)
Interaction Effects								
R×Exposure	-0.002***	-0.006	-0.036***	-0.010***	-0.002***	-0.004	-0.029***	-0.005
	(0.001)	(0.005)	(0.004)	(0.004)	(0.000)	(0.005)	(0.003)	(0.003)
Level Effects								
R	-0.000	-0.059***	0.054***	0.036***	-0.001	-0.044***	0.054***	0.044***
	(0.001)	(0.014)	(0.010)	(0.007)	(0.001)	(0.013)	(0.009)	(0.007)
Total Siblings			-0.034***	-0.035***			-0.030***	-0.030***
			(0.003)	(0.003)			(0.003)	(0.003)
F-Test : P-value	0.99	0.99	0.99	0.99	0.99	0.99	0.94	0.99
Observations	1,222,415	1,238,781	1,238,781	578,177	1,222,415	1,238,781	1,238,781	578,177
Adjusted R-squared	0.274	0.273	0.275	0.273	0.265	0.264	0.266	0.263

Table 5. Intra-household Redistribution *Notes.* This table presents estimates from an interaction of years of policy exposure with various indicators of the potential for intra-household redistribution. Exposure is measured in terms of linear years of potential program exposure. All specifications include state, birth year, and time fixed effects and a state-level linear time trend. MDM stands for midday meal. Reported P-values are from an F-test of whether the effect of Exposure is positive for the respective interaction variable. For household size, the effect is evaluated at the mean value of household size. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors in parentheses are clustered by state and year of birth.

5.3. **Intra-Household Redistribution.** Our implicit assumption throughout this paper has been that midday meals improve the nutritional status of children directly exposed to the program. But this is not self-evident. Although midday meals are targeted in-kind transfers, intrahousehold redistribution away from the targeted child towards other family members may temper the program's effect on learning; see, for example, Das et al. (2013). The extent to which this takes place depends, among other things, on parental preferences and family composition. Redistribution may be triggered by additional children in the household, traditional son-preference, or simply by having more mouths to feed.

Table 5 investigates these possibilities. The first row confirms our baseline results that each additional year of exposure to the program improves test scores significantly. The negative coefficient on the interaction term in columns 1 and 5 suggests that children living in larger households experience smaller improvements in reading scores for additional years of exposure to the program compared to children living in smaller households. However, the marginal effect of one additional member in the household is close to zero. It is, of course, conceivable that when redistribution occurs it takes place across children and not from children to adults. This possibility is explored in columns 2 and 6 by comparing children with and without any siblings. We do not find any evidence of redistribution to other children in the household.

The implicit assumption in columns 2 and 6 is that households do not distinguish between additional children based on their eligibility for midday meals in school. However, the need for redistribution across children is likely to arise only when the household budget constraint binds. Additional siblings lead to fewer resources per child in the household only when they are not eligible for midday meals. On the other hand, siblings eligible for free meals in school are likely to relax the household budget constraint. Columns 3 and 7 explore this by comparing children who have the same number of siblings but vary in terms of their siblings' eligibility for midday meals in school. We categorize siblings as non-eligible for school lunches either if they are too young to be in school or if they have already completed primary school. The coefficients indicates that children who have siblings receiving midday meals in school experience a 0.08 (0.05) point increase in reading (math) scores for each additional year of policy exposure. However, for a child with siblings who do not receive free meals in school, the reading or math score goes up only by half that amount. This suggests the presence of partial intra-household redistribution away from the child receiving midday meals in school to the child not eligible for midday meals in school.

Finally, we investigate intra-household redistribution against the backdrop of a well-known male bias in Indian households. Columns 4 and 8 compare the effect of program exposure on female children with and without male siblings. Our prior here is that in the presence of an inherent male bias, parents are more likely to spend the additional resources freed up on the male sibling. However, the coefficient estimate, although negative, is small relative to the level effect of the exposure and is statistically significant only in the case of reading. This points to negligible redistribution.

Overall, we find some evidence of substitution away from potential midday meal recipients. However, even in situations where substitution exists, redistribution is only partial so that, in general, the benefits from the school meals tend to stick to the targeted child.

6. Robustness

In this section we show that our main results from Section 4 are robust to a number of specification checks. Although the findings in this section generally hold for a quadratic specification as well, we keep things simple by estimating β using OLS in the following linear specification where exposure is measured in months:

$$y_{itcs} = \alpha + \beta \cdot \text{Exposure}_{i(tcs)} + \delta_t + \delta_c + \delta_s + \gamma_s t + \varepsilon_i \tag{4}$$

The estimates we present in this section are therefore directly comparable to those in columns 2 and 6 of Table 3 for reading and math, respectively. Section 6.1 deals with test score measurement issues. Section 6.2, demonstrates that a "treatment effect" is absent for a placebo group of secondary-school-aged children who were presumably not exposed to the program. Section 6.3 shows that the results are not driven by cohort-specific performance within states. In Section 6.4, we provide further evidence that our results are not driven by differences in the timing of implementation. Finally, in Section 6.6, we show that our results are robust to the inclusion of contemporaneous changes in other schooling inputs.

Table 7 presents estimates for β in equation (4) for these specification tests, with minor variations, as we will explain where relevant. The remainder of this section discusses these robustness checks and their results in more detail.

6.1. **Test Score Measurement.** In this section, we address two measurement issues pertaining to the outcome variable. The first is that we have treated test scores as interval scales, when in fact they are ordinal variables. The second is that there was a change in the number of levels of attainment captured in the test between the first couple of survey years and later survey years. We deal with each of these in turn.

ASER's tests comprise 5 levels, which represent cumulative skill mastery. For both tests separately, the child is marked at the highest level he or she can do comfortably. The first measurement issue arises from the fact that we have assigned integer values (0,1,2,3 or 4) to these levels. In so doing, we are treating test scores as interval scales, when in fact, they have an ordinal scale. This is problematic for two reasons. First, comparisons across groups—in our case children with different lengths of program exposure—are sensitive to the choice of scale. Second, interpretation of treatment effects are tricky in this context: with an ordinal scale, it is hard to know which levels of learning achievement are driving the percentage improvement over the baseline test scores.[25]

Table 6 addresses the ordinality of the test score measure by estimating linear probability models in which the dependent variable is a dummy variable equal to 1 if the child has achieved at least a certain level of mastery—separately for levels 1, 2, 3 and 4—in reading (Panel A) and math (Panel B). It confirms our main result that increased exposure fosters learning achievement. Furthermore, the estimates provide a useful interpretation of the effect of midday meal exposure on learning. Children with the full five years of exposure are not significantly more likely to move from having no mastery at all to at least basic knowledge of either reading (word recognition or better) or math (one-digit

[25]See Bond and Lang (2013) for a clear exposition of these issues and Singh (2017) for a nice application in a developing country context.

Test Score	≥1	≥2	≥3	=4
	(1)	(2)	(3)	(4)
		Panel A: Reading		
Exposure	0.000	0.002***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.992***	0.921***	0.556***	0.341***
	(0.023)	(0.030)	(0.031)	(0.030)
Adjusted R-squared	0.118	0.210	0.221	0.145
		Panel B: Math		
Exposure	0.001	0.002***	0.002***	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.972***	0.972***	0.691***	0.299***
	(0.024)	(0.032)	(0.037)	(0.028)
Adjusted R-squared	0.144	0.211	0.209	0.122
Observations	1,238,781	1,238,781	1,238,781	1,238,781

Table 6. Accounting for Ordinality of Test Scores. *Notes.* This table reports linear probability model estimates for β in equation (4) for reading and math. The dependent variable is a dummy variable equal to 1 if the child has mastered *at least* the skill level indicated in the column heading, separately for reading (Panel A) and math (Panel B). Exposure is the months of potential exposure to the midday meal program. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Standard errors in parentheses are clustered by state and year of birth.

number recognition or better). This probably reflects the fact that over 60 percent of children at the baseline have already reached at least level 1 in both subjects.

Increased exposure does, however, result in progress to higher levels. In terms of reading, children with the full five years of exposure are 10 percentage points more likely to be able to read at least a word; they are 15 percent points more likely to be able to read at least a paragraph; and they are, again, 10 percentage points more likely to be able to read a short story. In the case of math, they are, similarly, 10 percent points more likely to at least recognize double-digit number and also 10 percent points more likely to be able to do two-digit subtraction with carry over. There is, however, no improvement in the mastery of division with increased program exposure.

	Standardized Test Score[a]	Angrist-Lavy Index [b]	Placebo Sample[c]	State-Cohort Fixed Effects[d]	Excluding Earliest Implementers[e]	Excluding Latest Implementers[f]	Including Pre-mandate Implementers[g]	Household FE[h]	Other School Inputs[i]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Reading									
Exposure	0.0020***	0.0037***	-0.001	0.066***	0.009***	0.009***	0.009***	0.0352***	0.040***
	(0.000)	(0.001)	(0.002)	(0.018)	(0.002)	(0.002)	(0.001)	(0.001)	(0.007)
Mean at 4 months	0.27	0.76	3.23	1.07	1.09	1.05	1.07	1.08	
Observations	1,238,781	1,238,781	430,261	1,238,781	1,102,704	933,275	1,399,003	1,112,821	305,233
Adjusted R-Squared	0.273	0.242	0.043	0.276	0.276	0.299	0.278	0.370	0.298
Panel B: Math									
Exposure	0.0011**	0.0027***	0.000	0.050***	0.004	0.006***	0.006***	0.0275***	0.028***
	(0.000)	(0.001)	(0.002)	(0.017)	(0.003)	(0.002)	(0.001)	(0.001)	(0.006)
Mean at 4 months	0.27	0.72	3.10	1.06	1.08	1.03	1.06	1.07	
Observations	1,238,781	1,238,781	430,261	1,238,781	1,102,704	933,275	1,399,003	1,112,821	305,233
Adjusted R-Squared	0.264	0.236	0.061	0.268	0.268	0.289	0.266	0.361	0.304

Table 7. Robustness Checks *Notes.* With the exception of columns 3 and 8, this table presents ITT estimates for β in equation (4) using OLS. Exposure is the months of potential exposure to the midday meal program. Except for columns 1 and 2, the dependent variable is the raw reading test score in Panel A and the raw math test score in Panel B. Each column in Panel A and and each column in Panel B reports the estimate of β for a different regression. *a*. The dependent variable is the standardized test score. *b*. The dependent variable is the Angrist-Lavy math index. *c*. Sample comprises children aged 12-16 at the time of observation and also older than 10 at the time of program implementation. Exposure in this column captures placebo program exposure. *d*. This column presents estimates for β in equation (5), which includes state-cohort fixed effects instead of a state-level time trend. *e*. Sample excludes states which implemented midday meals in 2002. *f*. Sample excludes states which implemented midday meals in 2005 or 2006. *g*. Sample includes pre-2001 implementers Kerala, Gujarat, Puducherry and Tamil Nadu. *h*. This column presents estimates for β in equation (6), which includes household fixed effects. *i*. This column corrects for other schooling inputs, available only for survey years 2009-2012. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. Standard errors in parentheses are clustered by state and year of birth.

The second measurement issue arises from a change in the test instrument. From 2007-2012, test scores took 5 integer values: 0, 1, 2, 3 and 4. In 2005 and 2006, however, ASER aggregated single and double digit number recognition, so math test scores took on only 4 values. In our preferred specification, we use these raw test scores as outcome variables for the ease of interpreting our results. Due to the lack of disaggregation in the early rounds, we have coded single or double digit mastery in 2005 and 2006 as a math test score equal to 2, so our data contain no math test score values equal to 1 in those two years. This means that average test scores are mechanically higher in the first two survey years than in later survey years[26]. Survey year fixed effects will pick up these differences in part. Here we account for this change in measurement by standardizing the test scores in two ways: first, through z-scores, and second by constructing Angrist-Lavy Indices following Angrist and Lavy (1997). The z-scores are constructed in the usual manner, by standardizing the test score separately for each survey year. The Angrist-Lavy measure takes this standardized test score and assigns the index a value 0 if the standardized score is 0, a value 1 if it is less than or equal to one-half, and a value 2 if it is greater than one-half.

Table 7 presents estimates for β in equation (4) with the dependent variable being standardized test scores in column 1 and the Angrist-Lavy index in column 2. The results are qualitatively identical to our main results, indicating that the change in math assessment scoring is not material to our main result.

6.2. **Placebo.** In this section, we present estimates for equation (4) on the set of children aged 12-16 at the time of the survey; and who were also 11 or older—above the primary school going age—at the time of midday meals implementation. This constitutes a placebo group, which was potentially unexposed to the program.[27] In an exercise similar to the construction of potential years of exposure, we assign pseudo years of exposure to these older children. Specifically, pseudo years of exposure for children of age 12 at the time of the survey would be the same as the years of actual exposure assigned to children of age 6 according to the Online Appendix Table B1. The Exposure variable in the table reflects this placebo exposure.

We expect a null result: since these children were potentially not exposed to the program, our ITT estimate β on this placebo exposure should be statistically insignificant and close to zero. This is confirmed in column 3 of Table 7. The result also lends support to the underlying assumption of parallel-trends in our estimation strategy.

6.3. **State-cohort Specific Effects.** In this section, we tackle the concern that the timing of program implementation is correlated with cohort-specific performance within states, by introducing cohort-state fixed effects. Column 4 of Table 7 presents estimates for β in the following regression:

$$y_{itcs} = \alpha + \beta \cdot \text{Exposure}_{i(tcs)} + \delta_t + \delta_c + \delta_s + \gamma_{cs} + \varepsilon_i \tag{5}$$

[26]In an alternate measure, we aggregated the math scores in 2007-2012, recoding double digit mastery in 2007-2012 as a math test score equal to 1. This alternate measure has 4 overall categories and a mechanically lower average test score in the later years. The results using this alternate measure are very similar in spirit to the ones we obtain in Table3.

[27]We say "potentially unexposed" because it is possible that some schools in some states also introduced midday meals to this group. In some cases the states voluntarily extended the program to secondary schools prior to this; others did so later, or not at all. Clear documentation on implementation of the secondary school meals program does not exist. The upper bound on this age group is restricted by the data availability as ASER only tested children up to the age of 16.

As we explained in Section 3.2, variation in exposure is jointly determined by a child's age at the time of program implementation, and their age at the time of observation. Both depend on the child's birth cohort. For a given birth cohort, the latter depends on the survey year; the former depends on the timing of implementation in the child's state. With state-cohort fixed effects, variation in the exposure variable comes from observing the same cohort in a given state over consecutive survey years, as the child progresses through primary school grades. The admittedly strong identifying assumption in this context, is non-linearity in learning achievement within a state-cohort, as it progresses through consecutive grades.

The estimates in column 4 are bigger than our main results in columns 2 and 6 of Table 3. This suggests that failing to account for state-cohort effects, if anything, may result in underestimation of the true effects of midday meal exposure.

6.4. **Timing.** Columns 5-7 of Table 7 explore to what extent the timing of implementation influences our result by considering alternative samples based on the date of program implementation. Column 5 excludes states who were the earliest (2002) implementers. Column 6 excludes the late (2005-2006) implementers. Column 7 includes Kerala, Gujarat, Puducherry and Tamil Nadu, which had midday meals in place prior to the 2001 Supreme Court Mandate. In one specification the estimate for math scores is imprecise, but in general the estimates are statistically significant and their magnitudes are very similar to those in our main results. Finally, dropping states one-by-one does not affect the results either, indicating that our findings do not hinge on any one state (results not shown.)

6.5. **Household Fixed Effects.** The main results in this paper neglect the possibility that there may be unobserved heterogeneity at the family level. More specifically, it is plausible that children from better-off families have higher test scores, and do not suffer nutritional disadvantage, resulting in a downward bias in the treatment effect of exposure. We account for this by estimating the following household fixed effects model:

$$y_{ihtcs} = \beta \cdot \text{Exposure}_{i(tcs)} + \delta_{h(i)} + \delta_c + \gamma_s t + \varepsilon_i \tag{6}$$

where $\delta_{h(i)}$ is a household fixed effect of child i living in household h, and the remaining parameters are defined as before. Identification of β in this specification exploits variation in exposure across different children in the same household. Column 8 of Table 7 presents the results. They suggest that, indeed, the estimates in Table 3 may be downward biased estimates of the true treatment effect. The treatment effects estimated with household fixed effects remain statistically significant at the 1 percent level and are 4-6 times as large as the analogous OLS estimates.

6.6. **Other Schooling Inputs.** One concern with our quasi-experiment design is the possibility of simultaneous changes in schooling inputs other than midday meals. We account for this in Column 9 of Table 7, by controlling for a range of schooling inputs that could be spuriously correlated with midday meal exposure. These controls include teacher attendance; the availability of usable blackboards and the availability of supplementary learning material in school; whether different grades are taught in different classrooms; the total number of usable classrooms; and access to a functional tap for drinking water. As explained in Section 5.1, including these school-level controls entails a considerable loss in sample size since comprehensive school surveys were only conducted from 2009-2012.

As a result, the coefficient estimates in Column 9 are not directly comparable to those in Table 3. Table A4 in the Online Appendix provides estimates for each of the controls and elaborates on the ramifications of sample loss for the (much larger) point estimates. Here we simply note that the coefficient on program exposure remains positive and statistically significance, even after controlling for other schooling inputs.

7. CONCLUSION

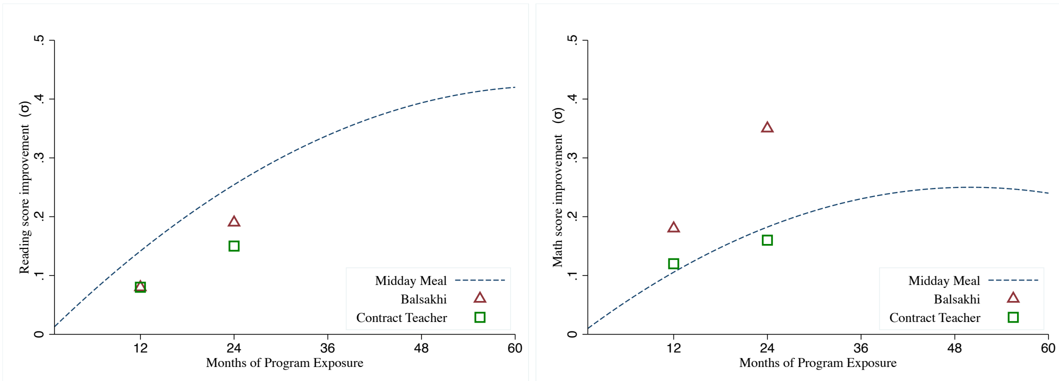
This paper has explored the effect of school feeding programs on children's learning achievement. What sets it apart from previous papers is that it studies the effect of long-term program exposure, to a large-scale program, using a large data set. The results indicate that exposure to midday meals for the nearly five-year duration of primary school increases test scores by 18% (0.17σ) for reading and 9% (0.09σ) for math relative to children with less than a year of exposure.

We show that there are complementarities between teaching- and learning-related classroom inputs, though not with more general schooling infrastructure. Relatively disadvantaged children show no differential treatment response, probably because the children in this rural Indian sample may have low baseline nutrition to begin with. And there is some evidence of partial redistribution away from program recipients but, in general, this in-kind transfer tends to "stick" to its intended beneficiaries.

While the focus of this paper has been on the learning effect of midday meals it is worth remembering that this is, if anything, a side benefit of the program. Its primary aim was to improve child nutrition, and available evidence suggests that it has been quite successful in so doing. For example, in a careful study, Afridi (2010) finds that the program reduced primary school children's daily deficiencies in protein, caloric intake and iron by roughly 100%, 30% and 10%, respectively. In keeping with the theme of this paper, the remainder of this section will discuss the cost effectiveness of this program when it comes to learning outcomes. This means that while the costs we describe cover the marginal cost of what is effectively a child nutrition program, the benefits we will describe pertain solely to the learning improvements; as such, the latter may be considered a lower bound on the "true" benefits of the program, which are likely to include both nutrition and health, in addition to learning.

One way to understand the cost effectiveness of the midday meals program with respect to learning outcomes is to compare it with alternative education interventions that target test score improvements. In principle, improving test scores can be accomplished by enhancing any number of schooling inputs including teacher quality, teacher numbers, teacher attendance or teaching aids. A few recent papers in the Indian context have studied the effect on test scores of improvements in the pupil-teacher-ratio (Muralidharan and Sundararaman, 2013), additional instruction through private tutoring (Banerjee et al., 2007; Berry and Mukherjee, 2016) and additional instruction through technology-aided methods (Muralidharan et al., 2016; Banerjee et al., 2007). Of these, Banerjee et al. (2007) and Muralidharan and Sundararaman (2013) provide contexts which are most closely comparable to ours in that their focus is also on primary school children. Hence, in what follows we use these two studies as benchmarks against which to compare the costs and learning benefits of the midday meal program.

Banerjee et al. (2007) study the effect of tutoring outside the classroom in the form of two alternate interventions. In the first—the "Balsakhi" program—weaker children in third and fourth grades were taught for two hours during school time by women from the local community. Effectively, the



(a) Reading **(b)** Math

Figure 5. Benefits of Midday Meals, Balsakhi, and Contract Teachers. *Notes:* This graph plots the effect of midday meal exposure on standardized test scores by month. The y-axis plots the improvement, measured in standard deviations (σ), in reading scores (panel a) and math scores (panel b). The midday meal program is compared to two comparable programs aimed at improving test scores, both of which lasted for two years: the Balsakhi program (Banerjee et al., 2007) and contract-teacher program (Muralidharan and Sundararaman, 2013).

program replaced regular classroom teaching by tutoring that involved more personalized attention. The second intervention included two hours of computer assisted learning (CAL) that involved math solving games during school time. Muralidharan and Sundararaman (2013) study the effect on test scores of reducing pupil-teacher-ratios in public schools by appointing contract teachers.

Figure 5 provides a graphical comparison of the standard deviation increases in reading scores (Panel a) and math scores (Panel b) from midday meal exposure to those of Banerjee et al. (2007)'s Balsakhis (squares) and Muralidharan and Sundararaman (2013)'s contract teachers (triangles).[28] The dashed blue line plots the standard deviation changes in test scores for each additional month of midday meal exposure, based on the quadratic "Exposure" estimates in columns 4 and 8 for reading and math, respectively; see Online Appendix Table A6. At the two-year mark the improvement in reading scores for midday meals compares favorably to that resulting from Balsakhi and contract teacher interventions. For math, improvement in scores from midday meals is substantially lower than that from Balsakhis, but slightly outperforms contract teachers.

Table 8 furnishes a more detailed cost-benefit comparison of these three programs. Columns 2 and 4 present the improvement in reading and math test scores, respectively, after two years of program exposure. Columns 3 and 5 present the improvement in test scores per dollar spent after two years of exposure, based on the annual cost estimates per child per year, presented in column 1. According to our calculations, discussed in Section 2, the cost of midday meal provision is 10 USD per child per

[28]Muralidharan and Sundararaman (2013) also provides estimates of the effect of reducing pupil-teacher-ratio by hiring permanent teachers on student achievement. However, these estimates are not comparable to our study since they do not capture the direct program effect of hiring an additional permanent teacher or the marginal cost of so doing. Hence we restrict our comparison to the contract teacher estimates. Similarly, the training, as part of the CAL program, in Banerjee et al. (2007), is restricted to numerical skills and effectively do provide comparable estimates of reading scores. We therefore do not incorporate CAL estimates in our cost-benefit comparisons either.

Program	Annual Cost[a] (USD)	Reading Score: Benefit (σ)	Reading Score: Benefit per USD spent	Math Score: Benefit (σ)	Math Score: Benefit per USD spent
	(1)	(2)	(3)	(4)	(5)
Midday Meals[b]	10	0.25	0.013	0.22	0.011
Balsakhi[c]	3.5	0.19	0.027	0.35	0.050
Contract Teachers[d]	3.4	0.15	0.022	0.16	0.024

Table 8. Cost Benefit Comparison after Two Years of Program Exposure. *Notes.* This table reports the annual cost of the program in 2009 USD (column 1); standard deviation improvements (benefit) in test scores (columns 2 & 4); and the corresponding improvement in test scores per dollar spent (columns 3 & 5). Note that the latter accounts for 2 years of costs after two years of program exposure. *a.* USD costs are calculated by converting nominal Indian Rupee amounts to 2009 Indian Rupees using the general CPI for industrial workers and then converting this Rupee amount to USD using 2009 exchange rates. *b.* Benefits calculated based on the estimates in columns 4 and 8 of the Online Appendix Table A6; authors' cost calculations. *c.* Balsakhi program estimates from Banerjee et al. (2007). The Balsakhi program cost Rs. 107 per child per year in 2002. *d.* Contract teacher estimates from Muralidharan and Sundararaman (2013); a contract teacher cost of Rs. 1500 per month in 2006-2007. The per child cost was calculated assuming an average class size of 30 in grades one through four according to the sample in the paper.

year.[29] The marginal cost of midday meals is thus almost three times higher than that of Balsakhis or contract teachers. This likely reflects the fact that meals are an individual-level consumption activity with commensurate individual-level costs as opposed to teachers, from whom all the children in a given class jointly benefit.

The numbers in columns 3 and 5 indicate that, at the two year mark, reading scores increase by 0.013σ and math scores by 0.011σ for each additional dollar spent on midday meals. The corresponding returns to the Balsakhi and contract teacher programs are at least twice as large due to their considerably lower unit costs. In terms of bang for buck in learning outcomes, therefore, it would seem that those programs are a better investment. Then again, the aim of the midday meal program was to improve nutrition and not learning. Viewed in this light, the program's learning benefits seem rather remarkable.

[29] Kristjansson et al. (2016) independently comes to the same estimate, and remarks that this makes midday meals one of the cheapest school feeding programs in the world. Note that 10 USD per child per year is the marginal cost of the program which, in the Indian context, rides on (fixed costs of) a well developed Public Distribution System. Marginal costs are likely to be substantially higher in countries which do not have such a system.

REFERENCES

Adelman, S., H. Alderman, D. O. Gilligan, and K. Lehrer (2008). The Impact of Alternative Food for Education Programs on Learning Achievement and Cognitive Development in Northern Uganda. University of Maryland, College Park Working Paper.

Afridi, F. (2010). Child Welfare Programs and Child Nutrition: Evidence from a Mandated School Meal Program in India. *Journal of Development Economics 92*(2), 152–165.

Afridi, F. (2011). The Impact of School Meals on School Participation: Evidence from Rural India. *Journal of Development Studies 47*(11), 1636–1656.

Afridi, F., B. Barooah, and R. Somanathan (2014). School Meals and Classroom Attention: Evidence from India. Mimeo. Indian Statistical Institute.

Ahmed, A. U. (2004). Impact of Feeding Children in School: Evidence from Bangladesh. International Food Policy Research Institute.

Alderman, H. and D. Bundy (2012). School Feeding Programs and Development: Are We Framing the Question Correctly? *The World Bank Research Observer 27*(2), 204–221.

Angrist, J. D. and V. Lavy (1997). The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco. *Journal of Labor Economics 15*(1, Part 2), S48–S76.

ASER (2005). Annual Status of Education Report. Technical report, ASER.

Banerjee, A. V., S. Cole, E. Duflo, L. Linden, et al. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics 122*(3), 1235–1264.

Behrman, J. R., S. W. Parker, and P. E. Todd (2013). *Incentives for Students and Parents*, Chapter 5. University of Chicago Press Chicago, IL.

Belot, M. and J. James (2011). Healthy School Meals and Educational Outcomes. *Journal of Health Economics 30*(3), 489–504.

Berry, J. and P. Mukherjee (2016). Pricing of Private Education in Urban India. Preliminary Results.

Bhattacharya, J., J. Currie, and S. J. Haider (2006). Breakfast of Champions? The School Breakfast Program and the Nutrition of Children and Families. *Journal of Human Resources 41*(3), 445–466.

Bond, T. N. and K. Lang (2013). The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results. *Review of Economics and Statistics 95*(5), 1468–1479.

Bundy, D., C. Burbano, M. E. Grosh, A. Gelli, M. Juke, and D. Lesley (2009). *Rethinking School Feeding: Social Safety Nets, Child Development, and the Education Sector*. The World Bank.

Case, A. and C. Paxson (2008a). Height, Health, and Cognitive Function at Older Ages. *American Economic Review 98*(2), 463–467.

Case, A. and C. Paxson (2008b). Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy 116*(3), 499–532.

Centre for Policy Research (2013). Midday Meals Scheme, GOI, 2013-14. Technical report.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2013). School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics 5*(2), 29–57.

Deaton, A. and J. Drèze (2009). Food and Nutrition in India: Facts and Interpretations. *Economic and Political Weekly Vol. 44*(Issue No. 07), 42–65.

Desai, A. H. and S. Muralidhar (2000). Public Interest Litigation: Potential and Problems. Technical report. International Environmental Law Research Centre.

Drèze, J. and A. Goyal (2003). Future of Mid-day Meals. *Economic and Political Weekly 38*(44), 4673–4683.

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.

Jacoby, E. R., S. Cueto, and E. Pollitt (1998). When Science and Politics Listen to Each Other: Good Prospects from a New School Breakfast Program in Peru. *The American Journal of Clinical Nutrition 67*(4), 795S–797S.

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

Jayaraman, R. and D. Simroth (2015). The Impact of School Lunches on Primary School Enrollment: Evidence from India's Midday Meal Scheme. *The Scandinavian Journal of Economics 17*(4), 1176–1203.

Jomaa, L. H., E. McDonnell, and C. Probart (2011). School Feeding Programs in Developing Countries: Impacts on Children's Health and Educational Outcomes. *Nutrition Reviews 69*(2), 83–98.

Karp, R., R. Martin, T. Sewell, J. Manni, and A. Heller (1992). Growth and Academic Achievement in Inner-City Kindergarten Children: The Relationship of Height, Weight, Cognitive Ability, and Neurodevelopmental Level. *Clinical Pediatrics 31*(6), 336–340.

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.

King, E. M. and J. R. Behrman (2009). Timing and Duration of Exposure in Evaluations of Social Programs. *The World Bank Research Observer 24*(1), 55–82.

Kingdon, G. G. (2002). The Gender Gap in Educational Attainment in India: How Much Can Be Explained? *Journal of Development Studies 39*(2), 25–53.

Kingdon, G. G. (2007). The Progress of School Education in India. *Oxford Review of Economic Policy 23*(2), 168–195.

Kremer, M. and C. Vermeersch (2005). *School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation*. The World Bank.

Kristjansson, E., A. Gelli, V. Welch, T. Greenhalgh, S. Liberato, D. Francis, and F. Espejo (2016). Costs, and Cost-outcome of School Feeding Programmes and Feeding Programmes for Young Children. Evidence and Recommendations. *International Journal of Educational Development 48*, 79–83.

Kristjansson, E., V. Robinson, M. Petticrew, B. MacDonald, J. Krasevec, L. Janzen, T. Greenhalgh, G. Wells, J. MacGowan, A. Farmer, et al. (2007). *School Feeding for Improving the Physical and Psychosocial Health of Disadvantaged Elementary School Children*. Number 1. Cochrane Collaboration.

Lawson, T. M. (2012). Impact of School Feeding Programs on Educational, Nutritional, and Agricultural Development Goals: A Systematic Review of Literature. Master's Thesis.

McEwan, P. J. (2013). The Impact of Chile's School Feeding Program on Education Outcomes. *Economics of Education Review 32*, 122–139.

McEwan, P. J. (2015). Improving Learning in Primary Schools of Developing Countries A Meta-Analysis of Randomized Experiments. *Review of Educational Research 85*(3), 353–394.

MHRD, G. o. I. Documentation on Meal Provision of Midday Meal Scheme. Technical report, Ministry of Human Resource Development, Department of School Education and Literacy, Government of India. `http://mdm.nic.in`, accessed on 24 April, 2016.

Muralidharan, K. (2006). Public-Private Partnerships for Quality Education in India. Mimeo.

Muralidharan, K., A. Singh, and A. J. Ganimian (2016). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. Technical report, National Bureau of Economic

Research Working Paper. No. 22923.

Muralidharan, K. and V. Sundararaman (2013). Contract Teachers: Experimental Evidence from India. Technical report, National Bureau of Economic Research Working Paper. No. 19440.

Neumann, C. G., N. O. Bwibo, S. P. Murphy, M. Sigman, S. Whaley, L. H. Allen, D. Guthrie, R. E. Weiss, and M. W. Demment (2003). Animal Source Foods Improve Dietary Quality, Micronutrient Status, Growth and Cognitive Function in Kenyan School Children: Background, Study Design and Baseline Findings. *The Journal of Nutrition 133*(11), 3941S–3949S.

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 Rrial of the Effects of Breakfast in Rural Primary School Children. *The American Journal of Clinical Nutrition 68*(4), 873–879.

Probe Team (1999). *Public Report on Basic Education in India*. Oxford University, New Delhi, IN.

Schick, A. and R. H. Steckel (2010). Height as a Proxy for Cognitive and Non-Cognitive Ability. Technical report, National Bureau of Economic Research. No. 16570.

Secretariat of the Right to Food Campaign (2005). Midday Meals: A Primer. Right to Food Campaign Materials.

Singh, A. (2008). Do School Meals Work? Treatment Evaluation of the Midday Meal Scheme in India. Young Lives Mimeo.

Singh, A. (2017). Learning More With Every Year: School Year Productivity and International Learning Divergence. Draft Paper.

Singh, A., A. Park, and S. Dercon (2014). School Meals as a Safety Net: An Evaluation of the Midday Meal Scheme in India. *Economic Development and Cultural Change 62*(2), 275–306.

Spears, D. (2012). Height and Cognitive Achievement among Indian Children. *Economics & Human Biology 10*(2), 210–219.

Strauss, J. (1986). Does Better Nutrition Raise Farm Productivity? *The Journal of Political Economy 94*(2), 297–320.

Strauss, J. and D. Thomas (1998). Health, Nutrition, and Economic Development. *Journal of Economic Literature 36*(2), 766–817.

Van Stuijvenberg, M. E., J. D. Kvalsvig, M. Faber, M. Kruger, D. G. Kenoyer, and A. S. Benadé (1999). Effect of Iron, Iodine, and β-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.

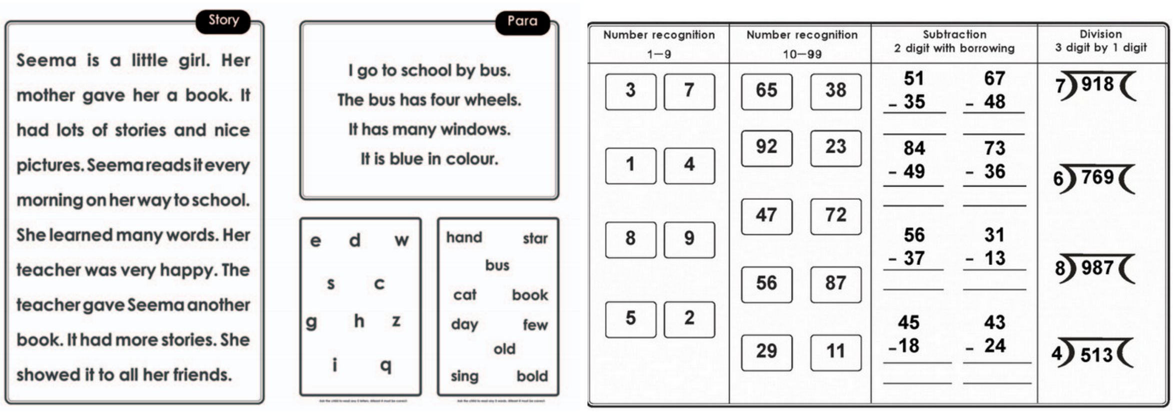
WFP (2013). State of School Feeding Worldwide. Technical report, World Food Program. World Food Program annual report.

Whaley, S. E., M. Sigman, C. Neumann, N. Bwibo, D. Guthrie, R. E. Weiss, S. Alber, and S. P. Murphy (2003). The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children. *The Journal of Nutrition 133*(11), 3965S–3971S.

APPENDIX: SUPPLEMENTARY TABLES AND FIGURES

State	Implementation Month	Year
Andhra Pradesh	January	2003
Arunachal Pradesh	July	2004
Assam	January	2005
Bihar	January	2005
Chhattisgarh	April	2002
Dadar & Nagar Haveli	February	2002
Daman & Diu	June	2003
Haryana	August	2004
Himachal Pradesh	September	2004
Jammu & Kashmir	April	2005
Karnataka	July	2003
Madhya Pradesh	January	2004
Maharashtra	January	2003
Manipur	November	2004
Meghalaya	January	2003
Mizoram	February	2006
Orissa	September	2004
Punjab	September	2004
Rajasthan	July	2002
Sikkim	October	2002
Tripura	April	2003
Uttar Pradesh	September	2004
Uttranchal	July	2003
West Bengal	March	2005

Appendix Table 1. Timing of States' Midday Meal Implementation. *Notes.* The states listed in this table are all included in the main sample. States available in ASER but excluded from the main sample due to lack of information regarding when the scheme was introduced: Jharkhand and Nagaland. States or union territories excluded from the main sample due to implementation prior to the mandate under study: Kerala, Gujarat, Kerala, Puducherry and Tamil Nadu. The month and year of midday meal policy implementation were collected from state midday meal scheme audit and budget reports.



(a) Reading Test

(b) Math Test

Appendix Figure 1. ASER Tests in the English Language