Preanalysis Plan: Improving School Preparedness and Child Outcomes through Integrated Child Development Services in Tamil Nadu, India

Alejandro Ganimian
New York University and JPAL

Karthik Muralidharan
UC San Diego, JPAL and NBER

Christopher R. Walters
UC Berkeley and NBER

March 22, 2018

1 Introduction

A large body of research demonstrates the importance of early childhood in molding the capabilities that determine children's life courses. Exposure to adverse conditions in the first few years of life hampers development across multiple domains (Alderman et al., 2006; Grantham-McGregor et al., 2007; Walker et al., 2007). Conversely, effective early childhood educational interventions can generate persistent benefits that translate into improved economic outcomes in adulthood (Heckman, 2008; Chetty et al., 2011; Heckman et al., 2013).

Despite this promising evidence, little is known about which government policies foster child development at scale in developing countries, a context in which early intervention seems likely to be especially important. Much of the research on effective early education programs comes from developed countries (Ludwig and Miller, 2007; Heckman et al., 2010; Puma et al., 2012; Peck and Bell, 2014; Walters, 2015; Elango et al., 2016). Some studies of early childhood interventions in developing countries show impressive gains, but these programs are often implemented at small scales outside of government education systems (Berlinski and Galiani, 2007; Berlinski et al., 2009; Attanasio et al., 2014; Gertler et al., 2014; Martinez et al., 2013). Further, estimates from several representative household surveys suggest that child health and nutrition outcomes in developing countries like India are extremely poor. For instance, the most recent National Family Health Survey reports that over 35% of children under 5 years old in India are underweight. Thus, it is a top policy priority for governments and international development agencies to identify cost-effective policies that can feasibly operate at scale.

This document details our planned analysis of data from a randomized experiment evaluating interventions within the Integrated Child Development Services (ICDS) system in Tamil Nadu, India. ICDS is the primary government program providing services related to early childhood development in India. The study is being conducted by J-PAL South Asia (J-PAL SA) in partnership with the government of Tamil Nadu (GoTN), with the goal of identifying effective interventions within a large-scale public early childhood care system. We will evaluate four interventions implemented in *anganwadi* centers (AWCs) operated by ICDS: (1) Hiring an additional staff member focused on early childhood education; (2) Supplemental nutrition; (3) A performance pay plan that links compensation for *anganwadi* workers (AWWs) to improvements in child health outcomes; and (4) An unconditional across the board pay increase for AWWs. The core objective of the study is to determine the impacts of these four interventions on child health, educational outcomes, and household behavior.

The sample for our experiment will include 800 AWCs selected to represent the diverse nutrition levels and geographic areas of Tamil Nadu. Analysis of this sample will be organized into two separate studies. The first study randomly assigns 640 AWCs in four rural districts to the four treatment arms or a control arm with no intervention. This rural study will focus on treatment effects and cost effectiveness of the four interventions. The second study randomly assigns 160 urban AWCs in Chennai to a treatment group that combines interventions 1-3 or a control group. In addition to treatment effects of the combined interventions, this urban study will feature results from a household survey measuring households' behavioral responses to improved public preschool services (including revealed preference valuation of these improvements as seen in switching behavior across public and private preschool services).

The next section describes the interventions, lays out the sampling and randomization protocols, and discusses the instruments used for measurement. Section 3 lays out our planned empirical analyses. Finally, Section 4 reports the results of some preliminary analyses based on baseline and midline data that have already been collected.

2 Study Design

2.1 Interventions

We next describe the four interventions to be evaluated in the experiment, including the motivation for each, details of implementation, and theories of change.

2.1.1 Supplemental Nutrition

A. Motivation

Tamil Nadu's ICDS offers a number of nutrition programs for children served by AWCs, including breast-feeding promotion, complementary feeding promotion, and take-home rations for mothers of children less than 2 years old, targeted supplementary nutrition in the form of porridge or raagi balls for 2- to 3-year-olds, and noon meals for 2- to 6-year-olds (including eggs three days per week). Nevertheless, our scoping study found a large share of moderately undernourished children in Tamil Nadu. Undernutrition was particularly prevalent among girls and in children aged 3 and above. This suggests scope for benefits associated with supplementary nutrition that could reduce the rate of undernourishment.

B. Intervention

The supplemental nutrition intervention will provide powdered milk to children aged 2-6 years attending selected AWCs. The supplemental nutrition will be provided each day that children attend the AWC.

C. Theory of change

The results of this intervention will determine whether supplemental nutrition can increase the weight of undernourished children, and reduce the prevalence of severe and moderate undernutrition. The improved nutrition and health may also result in greater enrollment at AWCs, generating benefits for children that do not currently attend AWCs. Finally, we will also measure learning outcomes and estimate the impact of providing supplemental nutrition on education outcomes.

2.1.2 Pay for Performance

A. Motivation

AWWs receive an average monthly honorarium of Rs. 7,142, which includes basic pay, grade pay, and a dearness allowance. Our scoping study found that AWWs are generally dissatisfied with this pay: 75 percent reported that they were either unsatisfied or very unsatisfied with their current salaries. This motivates an investigation of how best to increase salaries for AWW workers.

Research on incentive pay schemes demonstrates that linking compensation to performance can improve worker effort (Muralidharan and Sundararaman, 2011). Previous work by Singh (2015) shows the potential for positive effects of incentive pay in ICDS, but this study was conducted at a small scale and evaluated an incentive scheme that may not be sustainable in the long run. The impacts of incentive pay programs appear to be highly sensitive to elements of program design, and programs providing suboptimal incentives may fail to improve outcomes (Barrera-Osorio and Raju, 2017). Recent theoretical work emphasizes the desirable properties of incentive plans that reward teachers for growth in student outcomes relative to appropriate comparison groups (Barlevy and Neal, 2012). Our study will test the effectiveness of an optimal performance pay incentive plan operated at scale.

B. Intervention

The pay for performance (PFP) intervention will award bonuses to AWWs based on the magnitude of improvement in the weight-for-age ratio of each child aged 2-6 years attending the AWC, compared to improvements for a carefully-selected comparison group. Children will be grouped according to their relative standing in the distribution of baseline weight-for-age measurements. The comparison group for a particular child will consist of children at other AWCs in the same baseline nutrition category.

Every six months, a score will be assigned to each child based on his/her weight-for-age rank relative to others in his/her comparison group. The bonus paid to an AWW will be proportional to a weighted sum of scores for children in her AWC. This bonus structure implements the “pay for percentile” incentive scheme outlined by Barlevy and Neal (2012), with the additional modification that gains among the most malnourished children are given a higher weight in the bonus formula.

The scores used to calculate the bonus will be weighted according to a child's initial weight-for-age z-score, with weights equal to five for children with z-scores below -3, three for children with z-scores between -2 and -1, two for children with z-scores between -1 and 0, one for children with z-scores between 0 and 1, and zero for children with z-scores above 1. This weighting scheme is based on evidence regarding the benefits of early childhood nutrition provided in Alderman et al. (2006), and provides stronger incentives for improvement among malnourished children while avoiding incentives to contribute to obesity.

C. Theory of change

The incentives provided in the PFP intervention may lead to improvements in children's nutritional status by encouraging workers to ensure adequate provision of supplemental nutrition to undernourished children, and by fostering communication between AWWs and mothers about adequate caloric intakes at home. The average bonus paid in this intervention will be set to match the average payout in the across-the-board pay intervention (see below). This will allow a direct comparison of the cost-effectiveness of the two compensation schemes, holding total cost constant. More generally, the conventional approach to improving ICDS quality is to increase AWW salaries across the board and/or provide supplemental nutrition programs. Our study will allow for a comparison of the cost-effectiveness of these "business as usual" reforms vs. a theoretically-grounded reform designed to alter AWW incentives.

2.1.3 Across-the-board Increase in Honorarium

A. Motivation

Since 1975, ICDS has increased AWW salaries via across-the-board increases in honoraria distributed to all AWWs regardless of performance. Salaries compose the largest source of ICDS expenditure, so across-the-board honorarium increases crowd out resources that might be used for other program improvements. To understand the cost-effectiveness of alternative uses of ICDS funds it is essential to compare the effects of these reforms to the effects of across-the-board honorarium changes.

B. Intervention

The across-the-board honorarium intervention will provide an unconditional monthly honorarium payment to all AWWs. The additional payment will be an honorarium of Rs. 1,500 per month, equivalent to about 20 percent of the monthly salary of an AWW. This amount will match the total cost of the pay for performance intervention.

C. Theory of change

The across-the-board honorarium intervention provides no direct incentives for increased effort, and previous studies have found little effect of interventions of this sort (de Ree et al., forthcoming). As discussed by de Ree et al. (forthcoming), however, there are several mechanisms through which an unconditional pay increase could theoretically increase worker effort. Workers may increase effort due to a reciprocity or gift exchange motive (Akerlof, 1982); workers may reduce hours in outside jobs, increasing capacity for effort on the main job (UNESCO, 2014); higher pay may increase community expectations for workers, inducing more effort through social sanctions or rewards (Cotlear, 2006; Webb and Valencia, 2006); or workers may reduce shirking to increase the likelihood of keeping a job paying an above-market "efficiency" wage (Shapiro and Stiglitz, 1984). In the longer run an increase in pay may also change the types of workers that select in to teaching at *anganwadi* centers, though our study will not be able to evaluate this mechanism.

2.1.4 Early Childhood Care and Education (ECCE) facilitator

A. Motivation

Each AWC in Tamil Nadu is staffed with one AWW and one *anganwadi* helper (AWH). The AWW is responsible for multiple tasks, including providing pre-school education to children enrolled at the AWC, while the AWH is responsible for bringing children to the AWC, cooking food for them, and assisting with maintenance (Programme Evaluation Organization, 2011). We conducted a preliminary scoping study to gather information on the functioning of AWCs in Tamil Nadu (Muralidharan et al., 2015). This study found that the large number of tasks required of AWWs may prevent them from devoting sufficient attention to educational instruction. Specifically, of the 105 minutes officially allotted to pre-school education, only 33 minutes are devoted to pre-school instruction. Together with global evidence demonstrating the value of instructional time and smaller child/worker ratios (Chetty et al., 2011), these findings suggest that hiring additional staff at AWCs may generate benefits for children.

B. Intervention

The ECCE facilitator intervention will offer AWCs a one-time grant to hire a facilitator to assist AWWs with instructional tasks. As requested by GoTN, the ECCE facilitator will focus exclusively on providing pre-school education from 9.45am to 12pm.

Recruitment of ECCE facilitators will follow standard procedure stipulated by the ICDS department, except for a relaxation of the ICDS' usual marriage requirements, the age limits (we will allow recruits as young as 18 years old), and education requirements (we will allow recruits who have passed the 10th standard board examination). ECCE facilitators will be hired on two-year contracts. GoTN will develop training manuals for ECCE facilitators based on its current materials for AWWs. Using these manuals, GoTN will train ECCE facilitators and AWWs on their expected division of labor before the start of each school year.

C. Theory of change

The presence of an ECCE facilitator is expected to improve the performance of AWCs in several ways, including: (a) increasing enrollment and attendance; (b) ensuring that AWCs open on time; and (c) increasing time available for pre-school education (the ECCE facilitator will focus exclusively on this) while allowing the AWW to potentially have more time to perform other activities (including administrative tasks and health and nutrition related tasks).

2.2 Sampling and Randomization

2.2.1 Sampling of AWCs

There are two primary components to the sample: centers from four rural districts (the "rural study"), and centers in Chennai (the "urban study"). The rural sample was selected to capture heterogeneity across child nutrition levels and geography in rural areas of Tamil Nadu. Specifically, we split districts into four geographic zones used for administrative purposes by ICDS, and into four quartiles of the fraction of underweight children based on ICDS statistics. From the universe of all possible four-district combinations, we constructed the set of combinations that include one district from each geographic zone and one district from each nutrition quartile. We then sampled one combination from this set with probability proportional to total population across the four districts.

To select AWCs within each of the five sampled districts (four rural districts and Chennai), we began by excluding centers with NGO interventions, those in shared buildings, and those with vacancies in both staff positions (*anganwadi* worker and *anganwadi* helper). Sampling of districts from the remaining population was stratified by staffing vacancy, project and sector. We selected 80 centers per district with no staffing vacancy and 80 with one position vacant. The number of sampled centers per project was chosen in proportion to a project's share of the district total, and the number of sampled centers per sector was chosen in proportion to a sector's share of the project total. AWCs were sampled at random within sectors. All estimates of treatment effects will be re-weighted to be representative of the population of rural Tamil Nadu.

2.2.2 Random assignment

Random assignment for the rural sample was stratified by district, staffing vacancy, and a measure of local demographic characteristics. The demographic measure equals the first component from a principal components analysis (PCA) of population, age distribution, language, occupation distribution, and family income. The PCA analysis was conducted using population data for each AWC cachment area provided by Anna University. We divided the rural sample into 40 strata defined by district, vacancy status, and quintiles of the demographic principal component. Within each stratum we randomly assigned 2 centers each to the supplemental nutrition, pay for performance, and across the board interventions, four centers to the ECCE facilitator intervention, and six centers to the control group. The final sample therefore includes 80 centers each in the former three groups, 160 in the ECCE facilitator group, and 240 in the control group, for a total of 640 AWCs.

The random assignment protocol for the urban study was designed to account for the high geographic density of AWCs in Chennai. We calculated the distance from each sample AWC to the next closest sample AWC, then divided the sample into "singleton" centers with no other center within 0.5 kilometers and "non-singleton" centers with another center nearby. We randomly assigned 20 centers in the singleton group to treatment. In the non-singleton group we clustered random assignment by ward and sequentially randomly assigned wards to treatment until at least 20 non-singleton centers were treated. The final result is 40 centers in the urban treatment group combining the ECCE facilitator, supplemental nutrition, and pay for performance interventions, and 120 centers in the urban control group.

2.3 Instruments and Measurement

The project includes eight rounds of data collection: (1) a survey in September-October 2016, prior to the roll out of the interventions (which occurred in December 2016); (2) a round of measurement and intervention monitoring in April-May 2017; (3) a survey to update AWC rosters in July 2017; (4) a household survey in September 2017, conducted only in Chennai; (5) a round of measurement and intervention monitoring in November 2017; (6) an ECCE facility survey in January 2018, only in Chennai; (7) a classroom observation survey in February 2018; and (8) a final set of endline data collection in March through June 2018. We next describe the purposes and content of each round of data collection.

2.3.1 Round 1 (baseline, September-October 2016)

The round 1 survey was conducted to check balance in AWC characteristics across treatment arms, to document inputs, processes and outcomes of AWCs prior to intervention, and to collect baseline measurements that serve as inputs into the pay for performance bonus schedule. With these three goals in mind, we administered three instruments:

- A survey of AWCs that included questions on infrastructure, human capital and resources.
- A survey of child height and weight.
- A survey of children's math, language, and executive function skills.

2.3.2 Round 2 (April-May 2017)

The second round data collection aimed to capture changes in child weight-for-age in order to compute the first round of pay for performance bonuses. We therefore collected child weight and height in this round. We also administered an intervention monitoring module in this round to observe the implementation of the interventions.

2.3.3 Round 3 (July 2017)

The third round survey was added to the study to capture a new wave of children transitioning into *anganwadi* centers for the next school year and to address the problem of high student mobility out of AWCs. Children move both across centers within ICDS and in and out of the ICDS system. Mobility is especially high in Chennai, and during the summer months, when many families go back to their native villages. We therefore conducted this roster update survey to capture a new set of children to be included in the study (with the caveat that these measurements occurred after the start of the interventions). The data collection for this survey uses the weight/height instruments from the round 1 survey.

2.3.4 Round 4 (September 2017, Chennai only)

The household survey is designed to provide additional household characteristics for investigating effect heterogeneity and to analyze changes in household behavior resulting from improvements in AWC quality. The urban household sample therefore includes both students enrolled in AWCs in the urban study at baseline and students eligible to attend these AWCs who were not enrolled. The urban household survey sample was be drawn from a census of eligible children irrespective of baseline AWC enrollment. The instrument includes questions on household demographics and on preschool choices, a full account of past and current preschool attendance, tuition and other characteristics for chosen schools, perceptions of nearby preschool options, preferences for preschool characteristics and school attendance choices for siblings of eligible children.

2.3.5 Round 5 (November 2017)

As in round 2, the fifth round of data collection measured child weight and height to calculate performance pay bonuses, and also included an intervention monitoring survey.

2.3.6 Round 6 (January 2018, Chennai only)

The round 6 ECCE facility survey complements the household survey by measuring the characteristics of nearby preschool options for households in Chennai. The sample for this survey included all registered preschools located in the cachment areas of AWCs in the urban sample. The instrument collected information on enrollment, tuition, organizational form (public or private), and basic infrastructure.

2.3.7 Round 7 (February 2018)

The round 7 classroom observation survey aims to gather information on the day to day functioning of AWCs and measure changes in activities resulting from the ECCE facilitator intervention. It will include a combination of unannounced and announced visits in a subset of ECCE facilitator and control centers. This will allow us to capture staff and student attendance and observe preschool instruction and worker effort throughout the day, both in situations when observation is anticipated and when it is unanticipated.

2.3.8 Round 8 (endline, March-June 2018)

The final set of measurements is designed to study the impact of our AWC interventions on child health, learning, and household behavior. The endline round will take place in two parts. Between March and April 2018, we will implement the *anganwadi* worker survey, height/weight measurement, and academic skills measurement at the AWC, as in round 1. Between May and June 2018, we will conduct a survey of 7,680 households in the rural sample. The sample for this final survey will include households for 12 children at each AWC, drawn as a random sample stratified by the first round of data collection at which a child was observed.

The final household survey will allow us to measure changes in household resource allocation, worker effort to engage with households (including home visitations and quality of interactions between workers and mothers), and heterogeneous treatment effects in the rural sample. The household survey will also measure child height and weight. Both the final AWC measurement and the endline household survey will also measure mid-upper arm circumference (MUAC). As explained below, measuring MUAC and including anthropometrics in the final household survey will allow us to investigate whether workers in the pay for performance intervention took actions to temporarily increase weight before measurements occurring at the AWC, while also studying the persistence of any treatment effects after the end of the intervention. We will therefore be able to determine whether impacts of the interventions measured at the AWC reflect short-term gaming or long-term sustainable improvement. In addition, measuring outcomes in the household survey will help to insulate the analysis against the possibility of high attrition levels and/or differential attrition from treatment and control AWCs.

The household survey will also collect data on AWW interactions with households and the quality of these interactions and the advice provided.

3 Empirical Analysis

3.1 Rural Study

3.1.1 Study organization and anchor outcomes

Our analysis of the rural sample will be organized into two distinct papers. The first will focus on effects of the supplemental nutrition, pay for performance, and across the board interventions relative to the control group. The supplemental nutrition and pay for performance interventions are directly aimed at improving child nutrition; the across the board intervention represents a "business as usual" increase in *anganwadi* center funding, and improving nutrition is the primary focus of the *anganwadi* system. The analysis in this first paper will therefore be anchored by nutrition outcomes.

Functions of the weight-for-age z-score (WAZ) will serve as the main anchor outcomes for the first paper. Our core analysis will report effects on four WAZ outcomes: average WAZ among children with an initial WAZ below -2 (moderately underweight) the probability that WAZ falls below -2 (moderately underweight), the probability that WAZ falls below -3 (severely underweight), and average overall WAZ score. The first two outcomes are the main outcomes that we will focus on since around 30% of children were in this category at the baseline, and the main policy motivation for the interventions is to reduce the extent of moderate and severe malnutrition. The last outcome (average overall WAZ score) will be reported for completeness, but is a secondary outcome of interest (since it includes WAZ scores of non-malnourished children).

In addition, we will report estimates of the effects of the interventions on the quantiles of WAZ scores. We will specifically focus on effects on WAZ for children with baseline WAZ scores below -2, as this is the population that policy makers care the most about, though we will also report estimates for the full sample of children.

The pay for performance provides direct incentives for AWWs to increase WAZ scores. This creates the possibility of a multi-tasking problem in which AWWs target WAZ scores at the expense of other dimensions of nutrition (Holmstrom and Milgrom, 1991) and even raises the risk of short-term manipulation of the incentivized outcome (Figlio and Winicki, 2005). For example, AWWs might overfeed children in the days leading up to WAZ measurements, leading to transitory increases in WAZ that dissipate after the measurement period. To investigate this possibility we will also report estimates of effects on the mid-upper arm circumference (MUAC). MUAC is not directly targeted by the pay for performance intervention and is much less sensitive than WAZ to short-term overfeeding. We will therefore also report effects on MUAC for each intervention, and interpret it as a measure of nutrition that is not subject to multi-tasking distortions generated by high-powered incentives. However, it is possible that the MUAC is less responsive to our interventions, and so our anchor item of interest will still be WAZ (which is also what the policy makers care about most). We will therefore also collect a measure of WAZ in the household a few weeks after the center-level measurement (that bonuses will be paid on). Specifically, our round 8 household survey will also measure both WAZ and MUAC 4 to 6 weeks after final measurements at the AWC. This measure of WAZ is likely to provide the best measure of the "sustainable" impact of the interventions, and is likely to be unaffected by short-term strategies to boost WAZ before the high-stakes measurement. The household-level measurement will also facilitate a reduction in differential attrition across treatment and control groups.

As secondary outcomes, the first paper will also report effects on height-for-age (HAZ) scores as well as academic outcomes, which include math skills, reading skills, and executive function. We will report effects on each of these educational outcomes separately, as well as effects on a composite formed by taking a principal component of the three measures. Since height responds much more slowly than weight to changes in nutrition, we expect impacts on HAZ to be more modest. We expect any impacts of the supplemental nutrition, pay for performance, and across the board interventions on academic outcomes *to operate through impacts on health.* We will therefore compute the implied impact of health on education by dividing the effects of each intervention on educational outcomes by the corresponding effect on WAZ scores. Since it is possible that there are direct impacts of these interventions on education that do not operate through nutrition, we will interpret these instrumental variables estimates as upper bounds on the impact of nutrition on education.

The second paper will focus on effects of the ECCE facilitator intervention relative to the control group. We expect the ECCE facilitator to generate more time spent on preschool instruction. This analysis will therefore be anchored equally by math skills, reading skills, and executive function. Definitions of these outcomes are provided in the appendix. WAZ, MUAC, and HAZ scores will serve as secondary outcomes for the ECCE facilitator intervention. Since provision of the ECCE facilitator will also free up the time of the AWW to focus more on her nutrition-related functions, we may also find positive effects here.

3.1.2 Impact estimation

Experimental impacts will be estimated using models of the form:

$$Y_{ic} = \alpha_{s(c)} + X'_{ic}\gamma + D'_{c}\beta + \epsilon_{ic}, \tag{1}$$

where Y_{ic} is an outcome for child i enrolled at *anganwadi* center c, $s(c)$ is the randomization stratum of center

c and $\alpha_{s(c)}$ is a stratum fixed effect, X_{ic} is a vector of baseline covariates that includes a baseline measure of the outcome variable for individual students and the mean baseline outcome for all students at the center, and D_c is a vector of intervention indicators. The parameter of interest is the vector β, which measures the effects of each intervention relative to the control group. The key null hypotheses to be tested are that the elements of β equal zero.

We will estimate equation (1) in two samples. The first will be restricted to children who were present at baseline. The second will include all children at the AWC, including those who enrolled after baseline. The baseline sample is unaffected by any extensive margin responses to treatment that might lead to selection bias; the post-baseline sample may be affected by such selection, but includes more observations. We will therefore deemphasize the post-baseline sample if we find significant extensive margin effects, and also report Lee (2009) bounds on treatment effects under the assumption that treatment moves the likelihood of attrition in the same direction for all children.

The baseline sample enters the interventions earlier and therefore has longer potential treatment exposure than children who enter later at each point in time. We will therefore calculate average potential exposure length for the treated group in the pooled specification (in months) and interpret the pooled estimate as an average effect over this duration of treatment. Both samples will exclude children who were five years old at baseline measurement, who we expect to have exited the center shortly after the beginning of the experiment.

As noted above, we expect to see larger impacts on children with lower baseline nutrition levels. For children who enter the sample after baseline, there is no clean pre-randomization measure of nutrition available. To deal with this issue we will a set of "value-added" style specifications that use the WAZ score from the first date a child is present in the sample as a measure of initial nutrition, and look at effects on changes in WAZ between this initial measurement and later outcome measurements. These specifications will come with the caveat that adjusting for a post-baseline WAZ measure has the potential to introduce bias if this measure is affected by the treatment.

We will weight each specification with sampling weights that make the sample representative of the population of AWCs in Tamil Nadu, and for the baseline sample we will report estimates with and without baseline control variables. Inference will be clustered by AWC to account for the fact that treatments are assigned at the AWC level. We will assess the extensive margin response to treatment by estimating equation (1) at the center level with log enrollment as the outcome variable and the log of baseline enrollment included as a control.

3.1.3 Balance and attrition checks

We will validate random assignment by estimating versions of (1) with baseline covariates on the left-hand side, excluding the baseline covariates on the right. These covariates include baseline measures of each of the primary outcome variables discussed in Section 3.1.1. This specification will be used to test the null hypothesis that treatment assignment is unrelated to pre-treatment characteristics.

Even with random assignment, differential attrition has the potential to introduce selection bias into treatment/control comparisons. We will investigate the extent of differential attrition by estimating versions of equation (1) using an indicator for observed followup data as the dependent variable. To characterize the nature of attrition, we will calculate attrition rates and estimates of differential attrition separately by quartile of the WAZ score. We will also describe the distribution of baseline WAZ for children enrolled after randomization to assess whether treatment induced differential selection of new children into the sample.

3.1.4 Effect heterogeneity

To investigate heterogeneity in treatment effects for each primary outcome, we will split the sample into groups above and below the median of a baseline measure of the dependent variable (and by quartile), and conduct the main analyses separately in these groups. As noted in Section 3.1.1, for WAZ scores we will also report estimates for children with baseline values below and above -2, the conventional threshold for malnutrition. We will test the null hypotheses that each intervention has no effect in any category and that effects on all categories are equal. Similarly, we will split the sample by sex and report results separately for boys and girls, and test the hypothesis that these effects are equal. We expect the impact of the ECCE facilitator intervention to differ for centers with and without staffing vacancies, so we will also report separate estimates for this intervention by initial vacancy status.

3.1.5 Cost/Benefit Analysis

A key component of the first paper on the rural study will be a cost/benefit analysis focused on comparing the cost-effectiveness of business as usual reforms (supplemental nutrition and across the board pay increases) and the pay for performance incentive program. We will compute the improvements per dollar of additional government spending generated by each intervention. As a rough calibration we will assign monetary values to improvements in WAZ, HAZ, and academic skills based on the most recent available research linking short- and long-run outcomes, then calculate a summary measure of cost effectiveness for each intervention.

3.2 Urban Study

3.2.1 Treatment effects

Analysis of the Chennai sample will include three components. The first will parallel the rural analysis described above, including basic impact estimates, balance and attrition checks, and an investigation of effect heterogeneity. The urban investigation of heterogeneity will look at variation in effects by caste and family socioeconomic status, measured from the household survey. The Chennai intervention combines three treatment components from the rural study, so we expect it to have effects on nutrition as well as academic outcomes. We will therefore anchor this analysis equally on WAZ and academic skills (math skills, language, and executive function).

3.2.2 Willingness to pay

The second component of the Chennai analysis will use data from the household survey to analyze the demand for preschool services. The goal of this analysis is to recover an estimate of households' willingness to pay for the improved services provided by the AWC intervention. We will begin this analysis with a simple investigation of substitution between AWC enrollment and other preschool alternatives. Specifically, we will estimate versions of equation (1) with AWC enrollment, private preschool enrollment, other government preschool enrollment, and home care as dependent variables. These equations will be estimated in the household survey sample, which includes a representative set of children eligible for sampled AWCs irrespective of enrollment. The ratio of the intervention's impact on enrollment in an alternative preschool type to its impact on AWC enrollment identifies the share of children induced to attend the AWC that would otherwise enroll in the relevant alternative option (Kline and Walters, 2016).

Next, we will use a structural model of household choice to estimate willingness to pay parameters. Following Kremer et al. (2011), our analysis of willingness to pay will be based on a random utility discrete choice model of the form:

$$U_{ij} = W_j'\gamma - \delta P_j - \psi d_{ij} + \lambda AWC_j + \beta AWC_j \times D_{c(i)} + \xi_{ij}, \tag{2}$$

where U_{ij} is the utility of household i for preschool alternative j, W_j is a vector of preschool characteristics, P_j is the price of alternative j, d_{ij} is distance from i's home to alternative j, AWC_j is an indicator equal to one if alternative j is an AWC, $D_{c(i)}$ is an indicator equal to one if i's AWC is in the treatment group, and η_{ij} is an unobserved preference for alternative j. Households are assumed to choose the preschool alternative that maximizes utility.

In this framework the ratio of the treatment effect on household preferences to the coefficient on price, β/δ, serves as a measure of households' valuation of improved AWC quality. We will estimate this ratio assuming that ξ_{ij} follows an *iid* extreme value type I distribution, which makes (2) a standard multinomial logit model.

It is possible that price variation may be limited, in which case this ratio will be imprecisely estimated. Another concern is that the estimated price coefficient may be biased due to unobserved differences between more- and less-expensive alternatives, a standard concern in demand estimation. We will therefore also report estimates of β/ψ from models that exclude price. This ratio measures households' valuation of AWC improvements in terms of travel distance. In addition, we plan to estimate models that allow heterogeneity in the parameters $(\gamma, \delta, \psi, \lambda, \beta)$ with respect to caste and family socioeconomic status, as well as heterogeneity with respect to unobserved characteristics as in Berry et al. (1995).[1]

3.2.3 Parent expectations and investments

The third part of the urban study will study the effects of improved preschool quality on parents' human capital investment decisions and their expectations regarding their children's educational attainment. Our AWC intervention may improve child development and raise parents' perceptions of child ability, possibly inducing increases in parent investments. Evidence from the US suggests that parents respond to preschool interventions by increasing their interactions with children (Gelber and Isen, 2013). On the other hand, interventions may "crowd out" parental investments or lead parents to redirect resources to other household members.

[1]We also plan to estimate the model on the control group and test its predictions on the treatment group in order to probe the validity of the model's behavioral assumptions.

Motivated by these questions, we will use the household survey to investigate (a) whether parents underestimate their children's cognitive ability; (b) whether parents' estimations of cognitive ability are related to their investment choices (both across and within families); (c) whether the accuracy of parents' perceptions are impacted by the AWC intervention; and (d) whether parents' interactions with children and intrahousehold resource allocations respond to the intervention.

3.3 Statistical Power

3.3.1 Framework for power analysis

We next report some rough calculations to project statistical power for the study. These calculations are based on the following simplified model:

$$Y_{ic} = \alpha + \beta D_c + \mu_c + \eta_{ic}, \tag{3}$$

where Y_{ic} is an outcome for child i at center c, D_c is a dummy for assignment to an intervention group, and μ_c is an unobserved component of outcomes common to children at center c. Here we abstract from the fact that there are multiple interventions and from controls for baseline covariates; the latter will tend to boost the precision of the estimates.

Equation (3) implies that center-average outcomes are given by

$$\bar{Y}_c = \alpha + \beta D_c + \mu_c + \bar{\eta}_c, \tag{4}$$

where $\bar{Y}_c = n^{-1} \sum_i Y_{ic}$ and n is the within-center sample size (assumed to be the same across centers for simplicity). Under the null hypothesis that $\beta = 0$ and assuming homoskedasticity in the error η_{ic}, standard calculations imply that the variance of the ordinary least squares (OLS) estimator of β is given by

$$Var\left(\hat{\beta}_{OLS}\right) = \frac{Var(\mu_c + \eta_{ic}) \times \left[\tau + n^{-1}(1-\tau)\right]}{Cp(1-p)},$$

where C is the number of centers, $p = Pr[D_i = 1]$, and $\tau = Var(\mu_c)/Var(\mu_c + \eta_{ic})$ is the intraclass correlation coefficient (ICC).

The minimum detectable effect size (MDE) for a significance level α (probability of type I error) and power level κ (one minus probability of type II error) is given by

$$MDE = (z_{\alpha/2} + z_{1-\kappa})\sqrt{Var(\hat{\beta}_{OLS})},$$

where z_x is the value that puts mass x in the upper tail of a standard normal distribution. Putting the outcome in standard deviation units of the control distribution and using conventional values of $\alpha = 0.05$ and $\kappa = 0.8$, this is

$$MDE = (1.96 + 0.84)\sqrt{Var(\hat{\beta}_{OLS})}.$$

$$= 2.8\sqrt{\frac{\tau + n^{-1}(1-\tau)}{Cp(1-p)}},$$

where the last equality follows from the fact that $Var(\mu_c + \eta_{ic}) = 1$ when the outcome is in standard deviation units.

3.3.2 Power for the rural study

Our rural sample will include 240 control centers, 80 treatment centers for three of the interventions, and 160 for the ECCE facilitator intervention. Preliminary data suggests roughly 16 children per center. The ICC will likely vary by outcome. The solid line in Figure 1 displays the MDE for the three interventions with 80 centers as a function of the ICC for the rural study. For low values of τ the MDE is on the order of 0.1 standard deviations, a moderate effect size. This suggests that our design will be well-powered unless within-center outcomes are very highly correlated across children. In Section 4 we provide a preliminary empirical analysis using midline data that corroborates this theoretical calculation.

3.3.3 Power for the urban study

The urban study will include 40 treatment centers and 120 control centers. The dashed line in Figure 1 displays the MDE for treatment effect estimation in the urban sample. The urban MDE exceeds the rural MDE at every value of τ, reflecting the smaller urban sample size. We expect to be able to detect treatment effects on the order of 0.2 standard deviations for moderate values of the urban ICC. Since the urban study combines all interventions, we expect larger effects than in the rural study and effects of this magnitude seem reasonable.

It is less straightforward to assess power for estimation of households' willingness to pay for AWC improvements. Statistical precision for this parameter will depend on the distributions of prices for preschool alternatives and distances to AWCs across households, which are unknown. One necessary input for the willingness to pay calculation is an estimate of the change in the share of households attending the AWC in response to the intervention. To get a rough sense of power for the structural analysis we can therefore explore power for detecting this extensive margin response.

Suppose we have a representative sample of households in each AWC catchment area, and use this sample to fit a version of equation (4) in which $\bar{Y}_c$ is the sample average AWC attendance rate in the catchment area of center c. Calculations based on census data imply that the overall AWC attendance rate is roughly 0.3, implying that $Var(\eta_{ic})$ for this outcome is approximately $0.3 \times (1 - 0.3) = 0.21$. Applying the same formulas above yields an MDE of roughly $0.23 \times \sqrt{\tau + n^{-1}(1-\tau)}$. At $\tau = 0$ and $n = 20$ we could therefore detect an extensive margin effect size of roughly 0.05, with somewhat larger MDEs at higher values of τ.

4 Preliminary Analysis

We next report some preliminary results through the second-round survey (we refer to this as midline) for the rural sample, as this data was collected prior to filing our preanalysis plan. These results reflect impacts after four months of intervention, before the main rounds of post-treatment data collection.

4.1 Descriptive Statistics and Covariate Balance

Children enrolled in AWCs in Tamil Nadu begin with low nutrition status. This can be seen in Table 1, which shows descriptive statistics from the round 1 survey separately for each of the four districts in the rural study. The average age in the sample is around 3.5 years, and about half of children enrolled in AWCs are male. Average WAZ scores range from -1.75 in Thiruchirappalli to -1.45 in Kancheepuram, and statistics for HAZ scores are similar. This indicates that the average child is well below a normal weight and height in all districts, demonstrating the need for improvements in child health and nutrition in Tamil Nadu. Similarly, the fraction of malnourished children ranges from 0.29 to 0.40 across districts.

Table 2 investigates experimental balance by reporting estimates of equation (1) with round 1 (baseline) measures of WAZ and HAZ on the left-hand side. Averages of these variables are similar across intervention groups, and joint p-values give no cause for concern ($p = 0.50$ and 0.62). This suggests that randomization successfully produced comparable treatment and control groups.

4.2 Attrition

Information on attrition for the midline survey appears in Table 3. This table reports estimates of equation (1) in the baseline rural sample, using an indicator for an observed midline WAZ score as the dependent variable. Column (1) shows that 64 percent of the baseline sample is observed at midline, and differences between the four treatment groups and the control group are modest and statistically insignificant. We cannot reject the null hypothesis that attrition is balanced across all treatment arms ($p = 0.15$).

Columns (1) through (4) investigate heterogeneity in followup rates by carrying out this analysis separately by quartile of baseline WAZ score. The followup rate is highest in the lowest quartile (72 percent). Differences in attrition rates between the intervention and control groups are statistically insignificant in all quartiles. This pattern suggest that though there is a fair amount of attrition from the data, the extent of differential attrition is small and largely unrelated to baseline nutrition status.

4.3 Impact Estimates

The small time interval between our baseline and midline surveys provided little time for the interventions to generate impacts. It is therefore unsurprising that we find modest evidence of treatment effects at midline. Table 4 reports estimates of midline experimental treatment effects on WAZ scores, HAZ scores, and log AWC enrollment. Columns (1) and (2) show estimates of equation (1) for midline WAZ and HAZ scores, controlling for baseline measures of the dependent variable. We cannot reject the null hypothesis of no effect on outcomes in either specification ($p \geq 0.27$). Column (3) displays results for log AWC enrollment, estimated at the center level. These estimates indicate lower enrollment for the across the board and supplemental nutrition interventions, but we cannot reject the hypothesis that enrollment effects for all interventions equal zero.

These midline estimates offer an opportunity to assess the power of our research design to detect treatment effects. Standard errors for individual interventions in models with baseline controls are roughly $0.02 - 0.03$

standard deviations for WAZ/HAZ scores and 6-9 percentage points for enrollment. This suggests we have the power to detect effect sizes below 0.1 standard deviations, an estimate that is consistent with the more optimistic power calculations in Section 3.3.

Table 5 reports estimates for subsamples above and below the median of WAZ and HAZ at baseline. As in the full sample, most the estimates here are statistically insignificant. An exception is the effect of the supplemental nutrition intervention below the median, which is around 0.09 standard deviations and highly statistically significant. Figure 2 presents an additional look at impact heterogeneity for the supplemental nutrition intervention by splitting the sample into quartiles of the baseline WAZ score. The estimates here show declining effects across quartiles, with positive and statistically significant effects in the lowest two quartiles. These findings are consistent with our expectation that the supplemental nutrition intervention will have the largest impact on the most malnourished children, which motivates us to anchor this intervention based on effects on WAZ for children with low WAZ scores at baseline.

Figure 1: Power analysis

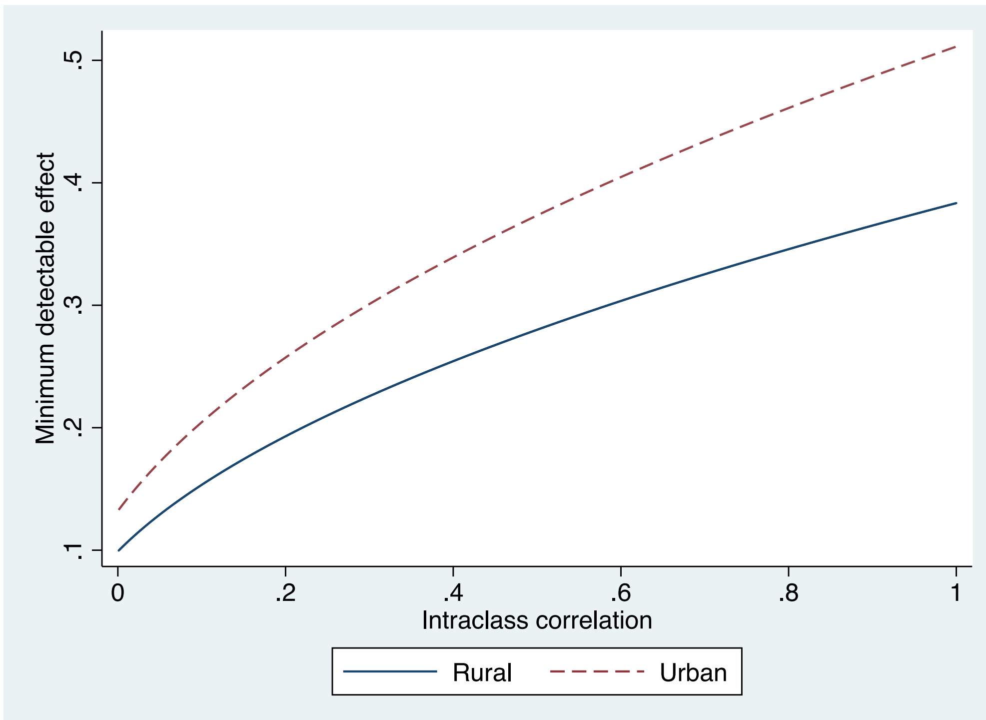
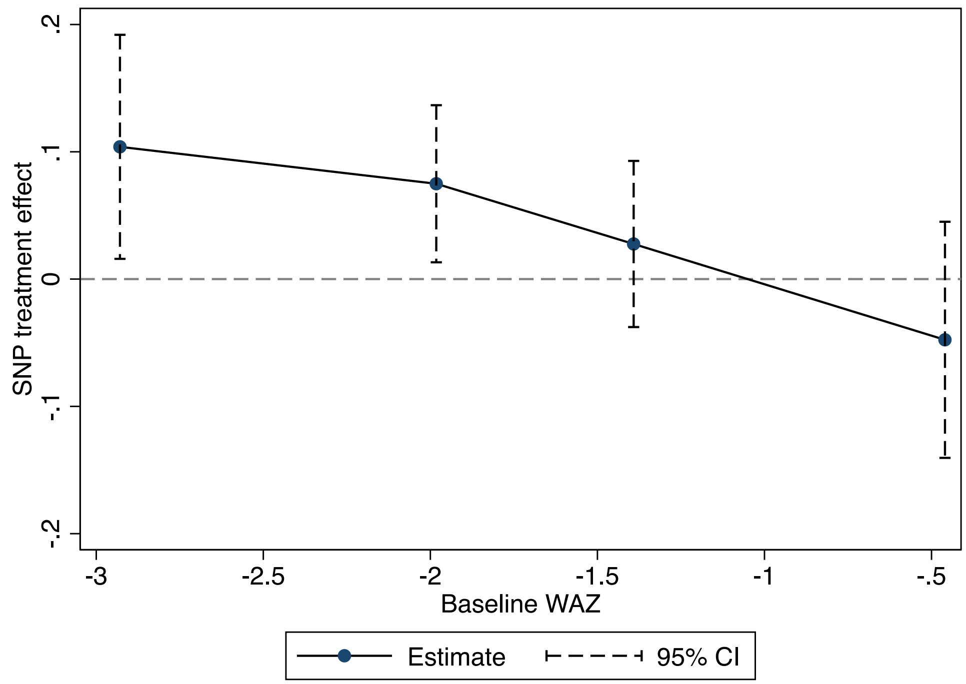


Figure 2: SNP treatment effects by baseline WAZ quartile



Notes: This figure plots treatment effects of the supplemental nutrition intervention on WAZ scores relative to the control group separately by quartile of the baseline WAZ score.

Table 1. Descriptive statistics

	Kancheepuram (1)	Karur (2)	Thiruchirappalli (3)	Virudhunagar (4)
Age	3.42	3.51	3.50	3.58
Male	0.484	0.505	0.497	0.495
Weight-for-age z-score (WAZ)	-1.445	-1.562	-1.751	-1.722
Fraction with WAZ below -2	0.292	0.316	0.397	0.382
Fraction with WAZ below -3	0.058	0.079	0.108	0.099
Height-for-age z-score (HAZ)	-1.404	-1.503	-1.654	-1.608
Fraction with HAZ below -2	0.314	0.311	0.374	0.360
Fraction with HAZ below -3	0.084	0.090	0.113	0.104
N	2308	2221	2687	2569

Notes: This table shows means of baseline variables by district in the rural sample.

Table 2. Baseline balance

	WAZ	HAZ
	(1)	(2)
Pay for performance	0.008	0.034
	(0.046)	(0.060)
Across the board	0.032	-0.004
	(0.049)	(0.064)
Supplemental nutrition	0.047	0.059
	(0.043)	(0.054)
Extra worker	0.040	0.059
	(0.037)	(0.046)
Control mean	-1.652	-1.577
F-stat: All coefficients zero	0.502	0.615
P-value	0.734	0.652
N	9879	9785

Notes: This table shows coefficients from regressions of baseline measures of weight-for-age (WAZ) and height-for-age (HAZ) on intervention indicators. All models include strata effects. Inference is clustered by AWC. *F*-statistics and *p*-values come from tests of the null hypotheses that differences between the five treatment groups and the control group equal zero.

***significant at 1 percent; **significant at 5 percent; *significant at 10 percent

Table 3. Attrition

		Differentials by baseline WAZ quartile			
	Differential	Lowest	Second	Third	Highest
	(1)	(1)	(2)	(3)	(4)
Pay for performance	0.034*	0.056	0.012	0.019	0.072**
	(0.020)	(0.038)	(0.034)	(0.035)	(0.036)
Across the board	-0.017	-0.018	0.006	-0.054	0.002
	(0.024)	(0.035)	(0.035)	(0.035)	(0.038)
Supplemental nutrition	-0.026	-0.014	-0.037	-0.036	-0.016
	(0.021)	(0.036)	(0.030)	(0.033)	(0.035)
Extra worker	0.006	0.039	0.036	-0.036	-0.005
	(0.018)	(0.029)	(0.027)	(0.028)	(0.030)
Control mean	0.636	0.724	0.520	0.568	0.552
F-stat: All coefficients zero	1.69	1.32	1.20	1.25	1.36
P-value	0.150	0.260	0.308	0.289	0.245
N	9879	2486	2484	2446	2463

Notes: This table reports coefficients from regressions of an indicator equal to one if a child has an observed midline WAZ score on intervention indicators. All models include strata effects. Inference is clustered by AWC. Calculations are restricted to children with baseline WAZ scores. Columns (2)-(5) show results of separate regressions by quartile of baseline WAZ score.

***significant at 1 percent; **significant at 5 percent; *significant at 10 percent

Table 4. Midline treatment effects on WAZ and HAZ

	WAZ	HAZ	Log enrollment
	(1)	(2)	(3)
Pay for performance	0.037*	-0.013	0.010
	(0.022)	(0.030)	(0.042)
Across the board	0.010	0.034	-0.033
	(0.026)	(0.038)	(0.051)
Supplemental nutrition	0.041*	-0.001	-0.047
	(0.022)	(0.034)	(0.046)
Extra worker	0.020	0.004	-0.017
	(0.017)	(0.028)	(0.034)
F-stat: All coefficients zero	1.29	0.31	0.42
P-value	0.271	0.869	0.796
N	6286	6176	639

Notes: This table displays coefficients from regressions of midline outcomes on intervention indicators. All models include strata effects. Samples in columns (1) and (2) are restricted to children present at baseline, and these models control for baseline measures of the dependent variable. Inference in these columns is clustered by AWC. Column (3) is estimated at the center level and includes a control for baseline log enrollment.

***significant at 1 percent; **significant at 5 percent; *significant at 10 percent

Table 5. Midline treatment effects by baseline category

	WAZ		HAZ	
	Below median	Above median	Below median	Above median
	(1)	(2)	(3)	(4)
Pay for performance	0.041	0.030	-0.037	0.029
	(0.027)	(0.033)	(0.037)	(0.043)
Across the board	0.032	-0.015	0.063	0.021
	(0.033)	(0.042)	(0.039)	(0.066)
Supplemental nutrition	0.089***	-0.009	0.004	-0.010
	(0.030)	(0.034)	(0.038)	(0.040)
Extra worker	0.023	0.018	-0.024	0.051
	(0.023)	(0.024)	(0.030)	(0.040)
F-stat: All coefficients zero	2.43	0.42	1.47	0.57
P-value	0.047	0.793	0.210	0.682
N	3297	2989	3223	2952

Notes: This table reports coefficients from regressions of midline WAZ and HAZ scores on intervention indicators separately by baseline quartile of the outcome. All models include strata effects and baseline outcome measures. Inference is clustered by AWC.
***significant at 1 percent; **significant at 5 percent; *significant at 10 percent

Appendix: Outcome Definitions

The academic outcomes for the child-level analyses will be defined as follows:

- Math skills will be measured as the average score across three assessments: verbal counting (which measures children's ability to count abstract numbers), producing a set (which measures children's ability to count physical objects), and number identification (which measures children's knowledge of numbers). Each assessment will first be scored in proportion-correct terms and these three scores will be averaged to obtain a single score per child. To the extent possible, we will also compute IRT scores for each child (though this may require excluding questions with partial credit).

- Language skills will be measured as the average score between two assessments: ordered object recognition (which measures children's receptive vocabulary) and letter name knowledge (which measures children's knowledge of words). Each assessment will first be scored in proportion-correct terms and these two scores will be averaged to obtain a single score per child.

- Executive function skills will be measured as the average score across three assessments: shape and color card sort (which measures children's flexibility to respond to changing stimuli), forwards digit span (which measures children's short-term memory), and Stroop white-black test (which also measures children's cognitive inhibition). Each assessment will first be scored in proportion-correct terms and these two scores will be averaged to obtain a single score per child.

References

Akerlof, G. A. (1982): "Labor contracts as partial gift exchange," *Quarterly Journal of Economics*, 97, 543–569.

Alderman, H., J. Hoddinott, and B. Kinsey (2006): "Long term consequences of early childhood malnutrition," *Oxford Economic Papers*, 58, 450–474.

Attanasio, O. P., C. Fernandez, E. Fitzsimmons, S. Grantham-McGregor, C. Meghir, and M. Rubio-Codina (2014): "Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial," *British Medical Journal*, 349:g5785, 1–12.

Barlevy, G. and D. Neal (2012): "Pay for percentile," *American Economic Review*, 102, 1805–1831.

Barrera-Osorio, F. and D. Raju (2017): "Teacher performance pay: experimental evidence from Pakistan," *Journal of Public Economics*, 148, 75–91.

Berlinski, S. and S. Galiani (2007): "The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment," *Labour Economics*, 14, 665–680.

Berlinski, S., S. Galiani, and P. Gertler (2009): "The effect of pre-primary education on primary school performance," *Journal of Public Economics*, 93, 219–234.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011): "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 126, 1593–1660.

Cotlear, D. (2006): "Improving Education, Health Care, and Social Assistance for the Poor," in *A New Social Contract for Peru: An Agenda for Improving Education, Health Care, and the Social Safety Net*, ed. by D. Cotlear, World Bank.

de Ree, J., K. Muralidharan, M. Pradhan, and H. Rogers (forthcoming): "Double for nothing? Experimental evidence on the impact of an unconditional teacher salary increase on student performance in Indonesia," *Quarterly Journal of Economics*.

Elango, S., J. L. Garcia, J. J. Heckman, and A. Hojman (2016): "Early childhood education," in *Economics of Means-Tested Transfer Programs in the United States*, ed. by R. A. Moffitt, Chicago, IL: University of Chicago Press, vol. II.

Figlio, D. N. and J. Winicki (2005): "Food for thought: the effects of school accountability," *Journal of Public Economics*, 89.

Gelber, A. and A. Isen (2013): "Children's schooling and parents' behavior: evidence from the Head Start Impact Study," *Journal of Public Economics*, 101, 25–38.

Gertler, P., J. J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014): "Labor market returns to an early childhood stimulation intervention in Jamaica," *Science*, 344, 998–1001.

Grantham-McGregor, S., Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, and B. Strupp (2007): "Development potential in the first 5 years for children in developing countries," *The Lancet*, 369, 60–70.

Heckman, J. J. (2008): "Schools, skills, and synapses," *Economic Inquiry*, 46, 289–324.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010): "The rate of return to the HighScope Perry Preschool program," *Journal of Public Economics*, 94, 114–128.

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013): "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 103, 2052–2086.

Holmstrom, B. and P. Milgrom (1991): "Multitask principal-agent analyses: incentive contracts, asset ownership, and job design," *Journal of Law, Economics, and Organization*, 7, 24–52.

Kline, P. and C. R. Walters (2016): "Evaluating public programs with close substitutes: the case of Head Start," *Quarterly Journal of Economics*, 131, 1795–1848.

Kremer, M., J. Leino, E. Miguel, and A. P. Zwane (2011): "Spring cleaning: rural water impacts, valuation, and property rights institutions," *Quarterly Journal of Economics*, 126, 145–205.

Lee, D. S. (2009): "Training, wages and sample selection: Estimating sharp bounds on treatment effects," *Review of Economic Studies*, 76, 1071–1102.

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

Martinez, S., S. Naudeu, and V. Pereira (2013): "The promise of preschool in Africa: a randomized impact evaluation of early childhood development in rural Mozambique," World Bank Impact Evaluation Database.

Muralidharan, K., A. Ganimian, and P. Singh (2015): "Improving school preparedness and child health outcomes through ICDS," J-PAL South Asia.

Muralidharan, K. and V. Sundararaman (2011): "Teacher performance pay: experimental evidence from India," *Journal of Political Economy*, 119, 39–77.

Peck, L. R. and S. Bell (2014): "The role of program quality in determining Head Start's impact on child development," S Department of Health and Human Services, Administration for Children and Families.

Programme Evaluation Organization (2011): "Evaluation report on Integrated Child Development Services," PEO Planning Comission, Government of India.

Puma, M., S. Bell, R. Cook, and C. Heid (2012): "Head Start Impact Study: Final Report," US Department of Health and Human Services, Administration for Children and Families.

Shapiro, C. and J. E. Stiglitz (1984): "Equilibrium unemployment as a worker discipline device," *American Economic Review*, 74, 433–444.

Singh, P. (2015): "Performance pay and information: reducing child undernutrition in India," *Journal of Economic Behavior and Organization*, 112, 141–163.

UNESCO (2014): "Teaching and learning: achieving quality for all," EFA Global Monitoring Report 2013/14.

Walker, S. P., T. D. Wachs, J. M. Gardner, B. Lozoff, and G. A. Wasserman (2007): "Child development: risk factors for adverse outcomes in developing countries," *The Lancet*, 369, 145–157.

Walters, C. R. (2015): "Inputs in the production of early childhood human capital: evidence from Head Start," *American Economic Journal: Applied Economics*, 7, 76–102.

Webb, R. and S. Valencia (2006): "Human Resources in Public Health and Education in Peru," in *A New Social Contract for Peru: An Agenda for Improving Education, Health Care, and the Social Safety Net*, ed. by D. Cotlear, World Bank.