The Impact of Expanding Access to Early Childhood Education Services in Rural Indonesia

Brinkman, S.A.; Hasan, A.; Jung, H.; Kinnell, A.; Pradhan, M.

DOI
10.1086/691278

Publication date
2017

Document Version
Final published version

Published in
Journal of labor economics

The Impact of Expanding Access to Early Childhood Education Services in Rural Indonesia

Sally Anne Brinkman, *Fraser Mustard Centre, Telethon Kids Institute, and University of Western Australia, Perth*

Amer Hasan, *World Bank, Education Global Practice*

Haeil Jung, *Korea University, Seoul*

Angela Kinnell, *Fraser Mustard Centre, Telethon Kids Institute, and University of Western Australia, Perth*

Menno Pradhan, *University of Amsterdam, Vrije Universiteit Amsterdam, and Amsterdam Institute for Global Health and Development*

This paper examines the effects of an intervention that expanded access to low-cost, government-sponsored, community-based playgroups in rural Indonesia. Instrumental variables and difference-in-differences models indicate that while the intervention raised enrollment rates and durations of enrollment for everyone, on average, there was little impact on child development. The two models correspond to different durations of project exposure. The difference-in-differences model captures greater exposure and shows that there are modest and sustained impacts on child development—especially for children from more disadvantaged backgrounds. There is also evidence that the intervention encouraged substitution away from other services, such as kindergartens.

Authors are listed in alphabetical order. We would like to thank Dedy Junaedi, Upik Sabainingrum, Anas Sutisna, Lulus Kusbudiharjo, and Mulyana for managing

[*Journal of Labor Economics*, 2017, vol. 35, no. S1]
 0734-306X/2017/35S1-0016$10.00
Submitted June 27, 2015; Accepted December 16, 2016

I. Introduction

Early childhood education and development (ECED) services can be a cost-effective intervention to overcome the developmental losses associated with growing up in a disadvantaged environment (Heckman et al. 2010). With around 250 million children under age 5 in the developing world failing to reach their full development potential, it is no wonder that, in recent years, governments and donors have scaled up their commitment to early childhood development interventions in developing countries (Sayre et al. 2015; Black et al. 2016). In support of such policies, it is important to have evidence on the effectiveness of these interventions—particularly from developing countries.

ECED services in developing countries are delivered through a variety of mechanisms. Education ministries often focus on kindergartens, which cater to children ages 5–6 and are primarily geared toward increasing children's preparedness for primary school. In addition to that, there are playgroups, which cater to younger children from age 3. Often these take a more community-based approach, where higher private contributions are expected through volunteer labor, tuition fees, or making infrastructure available. Both governments and nongovernmental organizations (NGOs) can support communities to provide these programs, through subsidizing start-up costs and/or providing technical inputs, for example, through teacher training. In addition, there are many private initiatives that provide ECED services without any involvement of governments or NGOs.

In this paper, we present an impact evaluation of a community-based ECED program in rural Indonesia implemented by the Government of Indonesia and financed through a World Bank loan. The Indonesia Early Childhood Education and Development project provided 3,000 villages

the fieldwork. Mayla Safuro Lestari Putri, Mulyana, and Nozomi Nakajima provided excellent research assistance. We would like to thank participants in the conference "Essays on the Economics of Education, Training and Labor Markets: A Festschrift Conference in Honor of Robert Lalonde," held on April 17–18, 2015, for their feedback. Luis Benveniste, Samer Al-Samarrai, and Harry Patrinos provided valuable feedback on an earlier version of the paper. This collaboration has been partially supported by an Australian AusAID Development Research Awards Scheme Grant (ADRA 0800261). Data collection was partially funded by the Government of the Kingdom of the Netherlands through the Dutch Education Support Program Trust Fund (TF057272), as well as by the Government of Indonesia. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. Contact the corresponding author, Haeil Jung, at hijk@korea.ac.kr. Information concerning access to the data used in this article is available as supplementary material online.

in 50 districts across the country with a package of interventions intended to improve children's school readiness. As part of this package, community facilitators (a term used locally to describe individuals hired to support projects at the community level) raised awareness in selected communities on the importance of child development. In addition, communities received block grants to allow them to establish preschool services of their choosing to cater to children between the ages of 0 and 6. The overwhelming majority of communities chose to establish playgroups—typically suited for children between the ages of 3 and 6. The block grants of about US$18,000 could be used to establish new preschool services or to strengthen existing services. Most communities (79%) established new services. The project also included 200 hours of training for individuals from the community who were selected to be teachers.

We employ two analyses. The first analysis takes advantage of the fact that the project was intended to roll out in three groups and that villages were assigned to these groups by lottery. Because of imperfect compliance with the results of the lottery, we compare villages that received block grants early (early treatment) with those that received block grants later (late treatment). We use the randomized rollout plan as an instrumental variable (IV). This analysis allows us to examine the impact of different durations of exposure to treatment. The second analysis makes use of a nonexperimental comparison group of villages that never participated in the project (Pradhan et al. 2013). This comparison allows us to examine the impact of receiving the project relative to never having received the project. The first analysis has the usual advantages of an experimental evaluation, but because of imperfect compliance with the rollout timetable, it has the drawback that the difference in duration of exposure between the control group and the treatment group is relatively short. The second comparison has the advantage that pre-intervention baseline data are available, thus making a difference-in-differences (DID) analysis feasible, and it allows for an examination of impacts in the medium term (3 years). We follow a randomly sampled cohort of children, who were age 4 at baseline, in two subsequent rounds. Existing instruments to measure child development, in various cognitive and noncognitive dimensions, were adapted for the Indonesian context (Hasan, Hyson, and Chang 2013) and used to assess these children's development outcomes.

While the number of papers on ECED interventions in developing countries is growing, there are only a few that apply a credible evaluation strategy (Nores and Barnett 2010; Rao et al. 2014). There are a number of studies on the impact of parenting (Gertler et al. 2014; Yousafzai et al. 2014) and kindergarten investments (Berlinksi, Galiani, and Manacorda 2008; Berlinksi, Galiani, and Gertler 2009). The number of studies that estimate the impact of playgroup interventions is still very limited. To our knowledge, while there are a number of ECED evaluations that correct for child background characteristics using matching, such as those in Bolivia (Behrman, Cheng

and Todd 2004) and the Philippines (Armecin et al. 2006), there are only two studies that have applied a Randomized Control Trial Design (RCT). Martinez, Naudeau, and Pereira (2012) evaluated the impact of a playgroup intervention supported by Save the Children in rural Mozambique. They found, in a setting where hardly any ECED services were available, that around half of the children living in the treatment villages had enrolled in the playgroups. They also found that the program resulted in better cognitive scores, higher enrollment in primary school, and a reduction in child labor. Work by Bouguen et al. (2013) focused on Cambodia; this research also used a RCT design. This exemplifies that these positive results can be hard to replicate for government-implemented programs. The authors found a negative effect of new ECED services on the cognitive development of children, probably because the services crowded out enrollment in primary school. The Cambodia study also faced problems in the implementation of the intervention and in the rollout of the study, which could also have dampened any positive effects of ECED.

Our paper contributes to the literature by showing how a government-sponsored, community-based program focused on playgroups affects enrollment and child development in a rural context where kindergartens and private programs were already operating. Results from both IV and DID analyses suggest that on balance enrollment rates increased 7–9 percentage points, which, with a baseline value of around 20% enrollment, constitutes an increase of 35%–45%. This is despite the fact that there was considerable substitution between services. In particular, enrollment in kindergartens fell as a result of the expansion of government-sponsored playgroups. Likewise, by the time they entered primary school, children in treatment villages reported being enrolled in project playgroups for about 4 months more than their control group counterparts. Once we account for substitution between services, we find that the intervention led to a net increase in duration of enrollment in ECED services of about 1.4 months on average in the treated villages, irrespective of the analytical approach followed.

The two approaches do, however, provide markedly different estimated impacts on child development. We find no positive impacts on child development outcomes for our IV analyses. This is to be expected given that in this analysis both sets of villages are ultimately exposed to the project and that this difference in exposure is relatively small—11 months. For the DID analysis, we find greater evidence of an impact on child development. In this analysis, the treatment villages have longer exposure (1 year by midline and 3 years by endline) to the project, while the comparison villages are never treated. The results indicate some positive impacts on social competence, emotional maturity, and language and cognitive development, but these do not persist over time. For children from disadvantaged backgrounds, the impacts are larger and do persist over time. In particular for language and cognitive development and social competence, we find positive effects for dis-

advantaged children in both the midline and the endline. This suggests that disadvantaged children received a stimulus while attending playgroups that they lacked at home. Point estimates for other aspects of development, while positively signed, are not statistically distinguishable from zero.

The paper proceeds as follows. Section II provides more details on ECED in Indonesia and the intervention. Section III validates the evaluation design and presents the key variables used in the analyses. Section IV presents descriptive evidence and our empirical strategy. Our analysis and findings are presented in the Section V. Section VI discusses the implications of our findings and concludes.

II. Early Childhood Education and Development (ECED) in Indonesia

A. The Different Types of Services Available

The ECED services in Indonesia are intended to cater to children from birth to age 6. Different types of ECED services are intended to cater to children of a specific age. Playgroups are intended for children ages 3 and 4. Kindergartens are intended for children ages 5 and 6.[1] In practice, these age cut-offs are hard to enforce. Some children may continue in playgroups after the intended age of 4. Some may enroll in primary school early—either at age 6 or even age 5. The incentive to do so is strong because attending kindergarten is not yet mandatory and most kindergartens charge fees, while primary school is compulsory and free. Once children have reached the age of 7, they are expected to begin primary school. Almost all children are enrolled in primary school by age 7.

The ECED services in the country take a variety of forms and are overseen by various ministries. Toddler family groups (BKB) are overseen by the National Family Planning Board, and they provide parenting education services. The Ministry of Education and Culture (MoEC) regulates playgroups (KB) and daycare centers (TPA), although the latter are largely an urban phenomenon. Kindergartens are regulated either by MoEC or by the Ministry of Religious Affairs (MoRA), depending on whether they are regular (TK) or Islamic kindergartens (RA). This paper focuses on playgroups and kindergartens.

Historically, MoEC has drawn a distinction between nonformal (playgroup) and formal (kindergarten) early childhood services. Since 2010, this distinction has been abandoned—at least on paper. Now both kindergartens and playgroups are under the purview of the Directorate General of Early Childhood and Development at the MoEC. In practice, however, the distinction between formal and nonformal services continues, with dif-

[1] Children up to the age of 5 generally go to an integrated health service unit called a Posyandu for basic medical care and monthly weighing. Some Posyandus may provide a preschool program (Pos-PAUD) similar to playgroups.

ferent types of services and teachers eligible for different forms and levels of government support.

Not all ECED services are equally intensive.[2] Playgroups and kindergartens both operate from 8 a.m to 11 a.m. However, playgroups typically meet only 3 days per week, while kindergarten services are available 5–6 times per week.[3] Thus, when interpreting the findings, it is important to keep in mind that children attending playgroups receive fewer hours of intervention each week than children attending kindergartens.

Despite the fact that there are a variety of early childhood services in Indonesia, the provision of these services has historically been beset by several challenges: (i) low levels of coverage, (ii) largely private provision of services in the face of low levels of public investment, and (iii) volunteer teachers with little or no training since very few institutions provide training for early childhood teachers.

B. The Early Childhood Development and Education Project

In an effort to address some of these challenges, the Government of Indonesia, in partnership with the World Bank and the Government of the Kingdom of the Netherlands, developed an intervention to increase access to early childhood services and increase children's school readiness in 50 relatively poor districts with generally low ECED participation.[4] Within each of these districts, 60 priority villages were identified on the basis of their poverty rate, a sufficiently large population of children between the ages of 0 and 6, a sufficiently large overall population, and the village's willingness to contribute financially to the project. Consequently, project services were implemented in 3,000 villages—roughly 4% of all 69,000 villages in the country. Three hundred and ten of these villages participated in the evaluation that is reported in this paper.[5]

[2] Day care centers (TPA) are largely an urban phenomenon and are usually open from 8 a.m. to 4 p.m. to care for children of full-time working families.

[3] ECED posts (Pos-PAUD) also operate from 8 a.m. to 11 a.m., but these are not considered in this paper. Quranic kindergartens (TPQ) usually operate in the afternoons from 2 p.m. to 4 p.m. The exclusive focus of these services is to study the Quran. However, they are designated as an ECED institution. Many children are able to attend the TPQ after having attended another ECED center in the morning.

[4] Under the project, districts were selected on the basis of a composite score created using a weighted average of poverty rates, gross enrollment rates, Human Development Index (HDI) rankings, geographical remoteness, whether or not they were included in the government's list of 3T districts (a particularly lagging set of districts), and district assurance of being "committed" to early childhood services. The composite score was used to select 50 districts for inclusion in the project from the 422 districts in Indonesia at the time (12% of all districts).

[5] These villages are from nine districts.

Each participating district was required to set up a district early childhood services office. Each village that participated in the project received the following package:

1. The services of a community facilitator whose job was to raise community awareness on the importance of early childhood services and share information on the benefits available under the project. These facilitators also provided communities with training on how to prepare proposals for the block grants available through the project.
2. Block grants for 3 years, totaling US$18,000 per village, which was to be spent on establishing or supporting two centers. Thus villages received US$3,000 per center per year for 3 years. These funds came with the requirement that no more than 20% could be spent on building new infrastructure. This limit meant that most of the centers established under the project involved rehabilitating existing buildings rather than constructing new ones. The remaining 80% could be spent on learning activities, health and nutrition, and management and administration of the center (including teacher salaries).
3. Two hundred hours of teacher training for up to two teachers per center. Teacher training was delivered via a two-tiered cascade training model. One hundred and ninety-two master trainers were trained for 500 hours and went on to provide 200 hours of training to approximately 12,000 teachers. Teachers were predominantly women from the village, who often had children of their own. Some had prior work experience in health and education. Others had no such prior experience.[6]

This package (community facilitation, block grants, and teacher training) is effectively the intervention evaluated in this paper. Together the components of the package were designed to encourage bottom-up community services that would be sustainable and suited to the needs of each village. While the original intent of the project had been to offer services to all children ages 0–6, in practice, most communities chose to establish playgroups. As is typical in Indonesia, these services operate three times a week for 3 hours a day, and they allow children between the ages of 3 and 6 to enroll.

For this package, the cost per child was about US$30 per year. This is calculated by dividing total project costs for implementation of the community-based component (US$54 million over 3 years) by the actual number of children (673,162 children) reported to have enrolled in the 3,000 villages where

[6] We are unable to link teachers in our sample to their children. Doing so would allow us to assess whether children whose mothers became community playgroup teachers received any additional benefit from the intervention.

the program operated.[7] This estimate excludes any voluntary contributions from the villages to the project. Villages often made available the land on which playgroups were housed. In contrast, other early childhood programs range in cost from US$37 per child in India to US$52 per child in Mexico to US$66 per child in Brazil—suggesting that this package was slightly less costly (Barnett 1997; Evans, Myers and Ilfeld, 2000).[8]

III. Evaluation Design, Validation, and Key Variables

A. Evaluation Design

The study is based on data collected from 310 villages spread out over nine districts in Indonesia that participated in the Early Childhood Development and Education Project. The districts were selected on the basis of their willingness to cooperate with a randomized rollout of the program and their location, the latter to ensure that the study locations encompassed the regional variety of the project locations. The evaluation design, sampling procedures followed, and the measures taken to address noncompliance are described in the study protocol (Pradhan et al. 2013).

In each of these districts, we sampled three groups of villages: 10 randomly sampled villages were assigned by lottery to receive the project in the first round, and 10 were assigned to receive the project later. In addition, we sampled a comparison group of 10 villages, which were recommended by local administrators as villages that were similar to the randomized villages but which were not going to receive the project.

Our design allows us to employ two comparisons. The first comparison takes advantage of the randomized allocation of villages to receive the project early or late. We address noncompliance by using an IV estimator, using the original assignment as an instrument.[9] The second comparison is between the villages that received the project late and the comparison group of villages. This allows us to employ a difference-in-differences analysis. The models are described in detail in Section IV.

Figure 1 overlays the timeline of project implementation with that of the surveys fielded for this study. We implemented three surveys between 2009 and 2013. Because of delays in the procurement of the survey firm, the baseline survey was fielded 6 months after early-treatment villages received their

[7] This information is drawn from the Implementation Status and Results (ISR) Report no. 11 of the project. This is available online at http://documents.worldbank.org/curated/en/684441468267567691/pdf/ISR-Disclosable-P089479-12-29-2013-1388324682405.pdf.

[8] All costs are per child per year and in 2014US$.

[9] Noncompliance with the randomized rollout was an issue, and this is addressed in the statistical analysis. However, one district (the 10th district that was originally included in the study) decided to break completely with their earlier commitment for a randomized rollout. As a consequence, we moved the sample from this district to a study district (Lombok) that had decided to randomize the over all of its project villages.

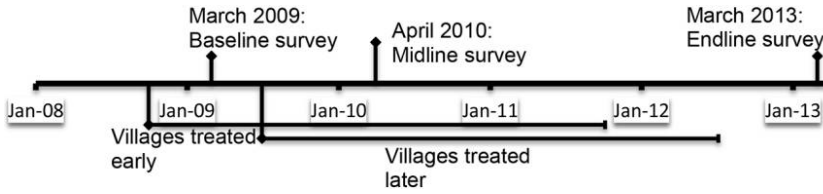

FIG. 1.—Timeline of project and surveys

block grants. Most of the late-treatment villages received their block grants about 5 months after the baseline survey. Thus, the average difference in exposure between early and late treatment villages is about 11 months.[10] The midline survey was fielded 9 months after the late-treatment villages received their block grants. Thus, the average difference in exposure between late-treatment villages and the comparison group is about 9 months at midline and 40 months at endline with no project exposure in the comparison group villages.

Our study sample consists of a random sample of children who were 4 years old at baseline. We restrict the analysis sample to children who were age 8.5 years or younger at endline because older children have outgrown some of our child development measures. After accounting for attrition (which is very low—only 15 children are lost between baseline and endline), we have 1,425 children in the IV analysis and 1,260 children in the DID analysis.

B. Validation

Since the evaluation was designed as a RCT, we expect balance in the villages originally assigned to receive the project early and late. To test this, we draw on two sources of data. The first is the 2008 round of a census of all Indonesian villages that takes place every 3 years—the Village Potential Statistics (PODES), which we merge with our data.[11] This data source allows us to assess balance on a number of village-level characteristics. We also use household and child characteristics from the baseline survey. As the baseline was conducted after the start of the projects for the villages treated early, we did not test for balance on child outcome variables between villages assigned to receive the project early and late.[12]

Table 1 presents a comparison of the villages that ended up receiving the project early and late and the comparison group of villages. Out of 310 villages,

[10] It takes 2 months between the disbursement of the first tranche and the start of the operation. In between, the community sets up the physical aspects of the center and waits for the teachers to be trained.

[11] Due to village mergers and splits over time, we are unable to match every single village in the evaluation sample to the PODES data.

[12] Appendix table A1 confirms that, as expected, the villages originally assigned to receive the project early and late are balanced on village and household characteristics.

Table 1
Balance between the Treatment and Comparison Groups at Baseline

	Early Treatment versus Late Treatment			Late Treatment versus Comparison Group		
	Early	Late	Difference	Late	Comparison	Difference
Using PODES 2008[a]:						
Village population	4,905	4,695	210	4,695	3,340	1,354***
			(749)			(414)
Proportion of families without electricity	.214	.189	.025	.189	.211	−.023
			(.038)			(.036)
Public kindergartens per 1,000 people	.005	.046	−.040**	.046	.039	.007
			(.020)			(.028)
Private kindergartens per 1,000 people	.408	.519	−.111**	.519	.543	−.024
			(.050)			(.064)
Public primary schools per 1,000 people	.771	.769	.002	.769	.773	−.004
			(.055)			(.056)
Private primary schools per 1,000 people	.328	.275	.052	.275	.216	.059
			(.069)			(.058)
Using the analysis sample[b]:						
Ever enrolled in any ECED service till the survey year (1=yes)				.181	.201	−.021
						(.035)
Cumulative months in any ECED service[c] till the survey year:				1.769	1.947	−.177
						(.343)
EDI: Physical health and well-being				−.012	−.008	−.004
						(.072)
EDI: social competence				−.053	−.015	−.038
						(.087)

EDI: emotional maturity				−.003	−.025	.022
						(.075)
EDI: language and cognitive development				−.130	−.066	−.064
						(.076)
EDI: communication and general knowledge				−.097	−.043	−.055
				(.086)		
Age of child (in years)	4.355	4.358	−.004	4.358	4.373	−.014
			(.018)			(.019)
Household size	4.637	4.701	−.064	4.701	4.769	−.068
			(.103)			(.105)
Wealth Z-score	.000	.041	−.042	−.023	.027	−.049
			(.094)			(.084)
Parenting scores Z-score	.064	−.006	.069	−.006	.046	−.051
			(.079)			(.079)
Mother's education primary school or less (1 = yes)	.511	.529	−.017	.529	.539	−.011
			(.036)			(.037)
Children's gender: male = 1	.476	.499	−.023	.499	.511	−.012
			(.025)			(.028)

NOTE.—Robust standard errors clustered at the village level are in parentheses.

[a] Unit of observation is villages. Sample size for villages varies because of missing data. For early treatment villages, *N* varies from 74 to 102 villages depending on the variable. For late-treatment villages, *N* varies from 63 to 105 villages. For comparison villages, *N* varies from 57 to 92 villages. In total, we have data from PODES on 299 out of the 310 villages in the sample.

[b] Unit of observation is the individual. The sample size for the column marked Early Treatment versus Late Treatment varies from 1,418 to 1,425. The sample size for the column marked Late Treatment versus Comparison Group varies from 1,256 to 1,260.

[c] Any ECED service includes playgroups and kindergartens.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

19 did not comply with the original assignment (see table A2). The numbers indicate that noncompliance results in villages that lacked kindergartens receiving treatment first. Otherwise, the villages seem rather similar. We also compare the villages that ended up receiving the project late with those in the comparison group. Late-treatment villages have larger populations than comparison group villages. This is expected given the targeting criteria of the project. Roughly 20% of families in the villages in our sample do not have access to electricity. Early-treatment villages have slightly fewer kindergartens per 1,000 people. All other aspects of the children and their households are balanced.

C. Key Variables

In 2013, we collected enrollment histories for each academic semester going back 5 years (from 2008—before the project started—to 2013). We use these enrollment histories to identify whether children were ever enrolled in playgroups or kindergartens and the duration for which they were enrolled in each type of service. The dependent variables in our analyses are percentage of children ever enrolled and months of enrollment.

The main instrument used for measuring children's development is the Early Development Instrument (EDI), which assesses children's readiness for primary school across five domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication and general knowledge (Brinkman et al. 2017).[13] For all aspects relating to the child's health and development, their primary caregiver was interviewed. At baseline and midline, the EDI data were collected using a short-form questionnaire (47 items), while the endline EDI data were from a long-form questionnaire (103 items). For the IV analysis, we standardize all measures using the standard deviation of the late-treatment group at baseline. For the DID analysis, we standardize all measures using the mean and standard deviation of the comparison group at baseline.

IV. Descriptive Evidence and Empirical Strategy

A. Descriptive Evidence

Before turning to our empirical strategy, in figure 2 we plot the average months of enrollment for each year and type of service for children in the three types of villages used in our analyses. Contrasting the villages treated early with the villages treated late provides descriptive evidence of the experimental comparison (IV) we report below. Likewise, a contrast of the villages treated late and the comparison group of villages provides descriptive evidence of the DID analysis we report below.

[13] A detailed discussion on these instruments and how they were adapted for use in Indonesia is provided in Alatas et al. (2013).

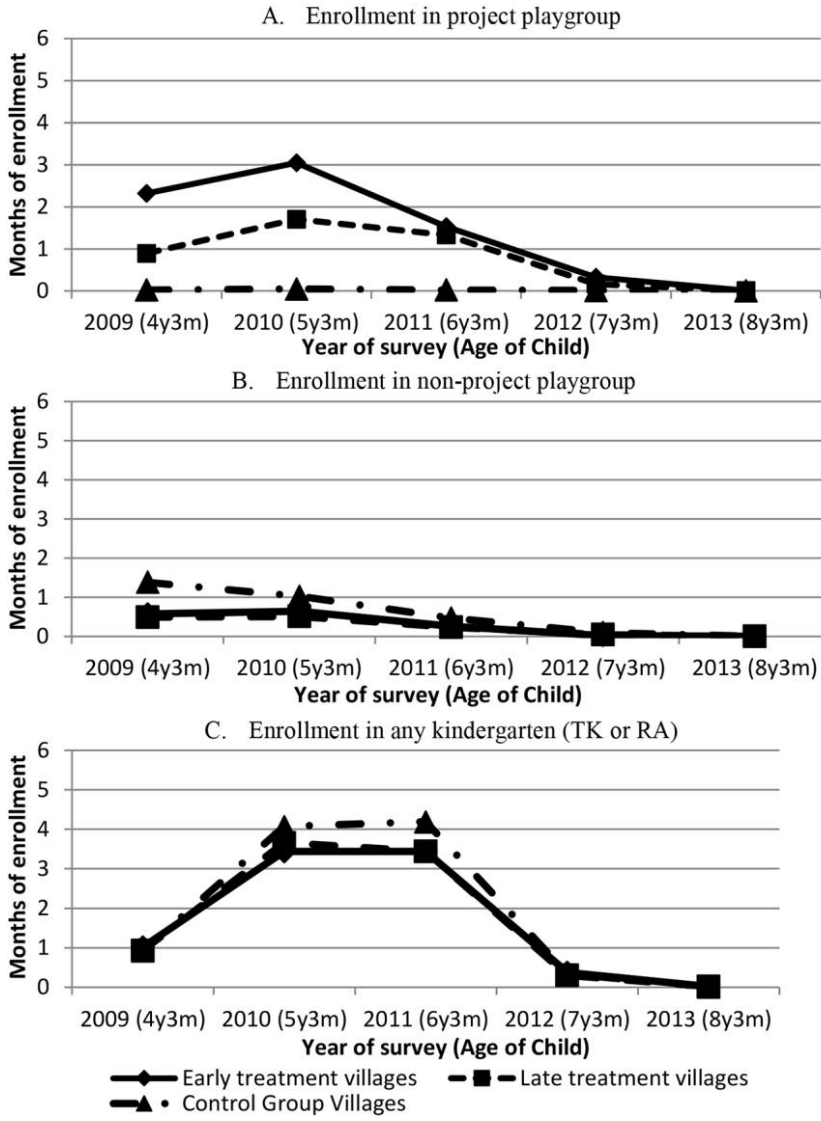

FIG. 2.—Enrollment patterns by age and year (villages treated later and the comparison group).

Each panel of figure 2 displays how children from different types of villages accumulate months of enrollment over time. The panels for project and nonproject playgroups show that in practice these services mostly cater to children ages 4 and 5. We found that children from the late-treatment group reported an average of 1 month of enrollment in project playgroups

because some villages used the grants to strengthen the existing centers that were run before the program started. Enrollment in playgroups rapidly declines at age 6 and is negligible at age 7. Children ages 4–5 in villages treated early enroll on average 2–3 months per year in project services. In contrast, these children only report about half a month of enrollment in nonproject playgroups. Children in villages treated late have lower average enrollment durations in project playgroups.[14]

In figure 2, there is also some evidence of project playgroups crowding out existing services—most noticeably kindergartens. Kindergarten primarily serves children ages 5 and 6. Whereas children in comparison villages enroll in kindergartens for about 3.7 months in a year on average, children in project villages do so for around 3 months. This suggests that to some extent the project playgroups crowded out kindergarten services.

B. Empirical Strategy

To test whether the descriptive evidence holds up to further scrutiny, we conduct two analyses—each of which compares two different sets of villages to each other to estimate the effect of the project. Our first analysis uses an IV method. It compares villages that were treated early to those villages that were treated late. At the time of the baseline survey, this comparison provides experimental evidence of the impact of 6 months of project exposure versus no exposure.[15] The estimation is therefore based on outcomes observed at the time of the survey, correcting for time-invariant characteristics (mother's education and child's gender) and time-variant characteristics that are not influenced by the project: child's age, household size, household wealth *Z*-scores,[16] and caregiver's parenting practices *Z*-scores.[17]

[14] Note that although the project was not yet present in villages treated later in the 2008–9 academic year, we do observe enrollment in project centers. This is because the project could have invested in existing playgroups. A center that eventually received support from the project is categorized as a project playgroup in this figure.

[15] Since children in villages treated early were already exposed to the project by the time of the baseline survey, we cannot use outcomes observed at baseline as a control when analyzing the midline and endline data.

[16] Households were asked if they owned any of the following: radio, television, refrigerator, bicycle, motor cycle, car, boat, mobile phone, livestock including chickens, pigs, cows, and goats. They were also asked about the materials used in the construction of the roof, walls, and floor of their homes, whether or not they had access to electricity in the home, and whether or not they had received social assistance (in cash or in kind). Responses were combined into a single index using principal components analysis. The score of the first principal is then standardized with the resulting variable having a mean of zero and a standard deviation of one.

[17] Drawing on the Longitudinal Study of Australian Children, we collected self-reports from the primary caregiver in our sample on how often they used each of a number of parenting practices that relate to their warmth, consistency, and hostility. Higher total parenting scores indicate higher levels of warmth and consistency

More formally, our IV estimate uses a two-stage least squares model (2SLS) as follows:

First-stage regression:

$$\text{Early Treatment}_j = \rho_0 + \text{Middle Assignment}_j\gamma_t + \text{Late Assignment}_j\pi_t + X_{ijt}\theta_t + \varepsilon_{ijt}. \quad (1)$$

Second-stage regression:

$$Y_{ijt} = \rho_1 + \widehat{\text{Early Treatment}_j}\alpha_t + X_{ijt}\beta_t + \varepsilon_{ijt}. \quad (2)$$

Here Y_{ijt} is the outcome variable of child i in village j at time t (takes values of 0, 1, or 2 for baseline, midline, and endline, respectively), and X_{ijt} is a vector of child and household characteristics observed at time t. Early Treatment$_j$ is a dummy variable that takes on a value of 1 for villages treated early and a value of 0 for villages treated late. From the original assignment strata, Middle Assignment$_j$ and Late Assignment$_j$ are the dummy variables indicating the middle and late random assignments, with the early assignment as a reference group; $\widehat{\text{Early Treatment}_j}$ is the predicted value from the first-stage equation in (1).

We run this 2SLS model for each time period, t. Thus, α_t indicates the impact of project implementation at time t. In order to obtain a consistent IV estimator in this analysis, we assume that the original random assignment of villages affects children's outcomes only through the assignment to early or late treatment status. All standard errors are clustered at the village level.

We also rerun this same specification for several subsamples in order to investigate whether the impact of the program is heterogeneous across different subgroups of children, such as poor and nonpoor children (above and below the sample average of the household wealth at baseline) and children with low- and high-parenting-score parents (above and below the average of the parenting scores at baseline).[18]

Our second analysis uses a difference-in-differences approach to compare villages treated late with a comparison group of villages over time. This comparison is quasi-experimental, but it has the advantage that neither group had been exposed to the project at baseline. Moreover, the compar-

and lower levels of hostility. The scores are standardized with the resulting variable having a mean of zero and a standard deviation of one. See Hasan, Hyson, and Chang (2013, 69–70) for a fuller description of this measure.

[18] We examined a balance table for each subgroup (available upon request) and found that most of characteristics and outcomes are balanced, except for social competence of poor and nonpoor children. At baseline, it seems that poor children in the late treatment group were worse than those in the comparison group, while nonpoor children in the late treatment group were better than those in the comparison group.

ison group never received the project, making it an ideal comparison group in the classical sense.

The DID model controlling for child fixed effects is estimated as follows:

$$Y_{ijt} = \varphi_i + \sum_{t=1}^{2} D_t \tau_t + \text{Late Treatment}_j \alpha + \sum_{t=1}^{2} \text{Late Treatment}_j D_t \delta_t + X_{ijt}\beta + u_{ijt}. \quad (3)$$

Here Late Treatment$_j$ is a dummy variable that takes on a value of 1 for villages treated late and a value of 0 for comparison villages. Thus, α indicates the baseline ($t = 0$) difference between the two groups; φ_i is a child fixed effect; and D_t is the time dummy for $t = 1$ (midline) or $t = 2$ (endline), which controls for age and time effects in the model. The coefficient δ_t at $t = 1$ or $t = 2$ are the DID estimators at midline and endline, respectively.

First, we investigate the effect of the project on enrollment in ECED services by analyzing both whether the child ever enrolled in a given type of service and the months of enrollment. The program could result in substitution away from existing services to the playgroups established under the project, and this could, in turn, affect child development outcomes because of the differences in dose and quality of the programs. Second, we analyze the effect on child development outcomes using the EDI.[19]

V. Empirical Results

A. Assessing the Validity of the Instrumental Variables Approach

The appendix documents the validity of the instrumental variables approach. Appendix table A3 reports the outcome of the first-stage regression for our IV estimates. Each round is reported separately. For example, the first three rows report the first-stage regression for all children at baseline, midline, and endline. In the first-stage regression, we instrument actual treatment with the original random assignment strata. There are two such strata—middle assignment and late assignment. Thus, the reference group for these two instrumental variables are villages in the early assignment stratum. We also control for the following child and household characteristics: child's age, gender, household size, household wealth, and maternal education. We repeat this exercise for all the subsamples for which we report results. The results across all of these first-stage regression results are virtually identical. This is as expected based on the data reported in appendix table A2—during the implementation process there was very little movement of villages across original assignment strata. To measure the strength of two instrumental variables, we re-

[19] For all survey rounds, we standardize these outcomes using means and standard deviations observed in the entire sample.

port the *F*-test statistics under the null hypothesis that the coefficients of two instrumental variables are 0. The *F*-test statistics for these two instrumental variables range from 135 to 191, which easily exceeds the normal threshold of 10 for a strong IV (Stock, Wright, and Yogo 2002). Thus, the first-stage regression results indicate that our IVs are strong enough to run the second-stage regression. Likewise, appendix table A1 captures the fact that randomization was successful and that there is balance between the original assignment strata.

B. IV Estimates of the Impact on Enrollment

Under the assumption that the randomization of villages is valid, the estimated difference in enrollment and child development outcomes between children in villages that were treated early or late yields an unbiased estimate of the impact of the project.[20] These results are reported in table 2.[21]

Each row displays the results of a separate regression that is intended to capture the impacts at baseline, midline, and endline. Each column reports the results for a different subsample of children: children from poor and nonpoor households (defined on the basis of a household wealth as described above),[22] as well as children whose parents report low- or high-parenting practices (defined on the basis of self-reported data described above).[23] These subgroup analyses allow us to assess whether disadvantaged children benefit more from exposure to the ECED project.

The column marked All Children in table 2 shows that, on average, children in villages treated early were more likely to be enrolled in project playgroups than those in villages treated later at all three survey periods. Specifically, the results show that they were 15, 20, and 23.6 percentage points more likely to be enrolled in baseline, midline, and endline, respectively. The duration of enrollment in project playgroups is also longer for children in villages treated early than for those in villages treated later at all three survey periods. Enrollment duration is 1.5, 3, and 4 months longer, respectively. There is no significant difference in enrollment rates or duration of enrollment in nonproject playgroups. There is, however, strong evidence of substitution away from kindergartens. Children from villages treated early are 10 percentage points less likely to be enrolled in kindergartens by midline and 15 percentage points less likely to be enrolled in kindergartens by endline than children from villages treated late. Their enrollment duration is also shorter—by about 1.1 months at midline and 2.5 months at endline.

[20] As shown in table 1, there are no considerable differences in child backgrounds between children in these two groups. This implies that the early-treatment and late-treatment villages are comparable.

[21] The OLS results are very close to the IV results. OLS estimates are available upon request.

[22] All children in the sample are from poor rural areas. Thus, the definition of poor and nonpoor children is relative, not absolute.

[23] Tables of balance between subgroups are available upon request.

Table 2
Instrumental Variables Estimated Impact on Enrollment Outcomes

	All Children	Poor	Nonpoor	Low Parenting Score	High Parenting Score
Ever enrolled in project playgroups until the survey year (1 = yes):					
Baseline	.150***	.151***	.143***	.156***	.140***
	(.031)	(.034)	(.052)	(.037)	(.043)
Midline	.200***	.216***	.187***	.197***	.194***
	(.047)	(.054)	(.065)	(.059)	(.055)
Endline	.236***	.216***	.270***	.249***	.218***
	(.053)	(.061)	(.067)	(.066)	(.060)
Ever enrolled in nonproject playgroups until the survey year (1 = yes):					
Baseline	.005	−.011	.023	.002	.006
	(.018)	(.014)	(.030)	(.017)	(.029)
Midline	.002	−.024	.039	−.021	.023
	(.025)	(.026)	(.038)	(.030)	(.035)
Endline	−.005	−.038	.039	−.023	.012
	(.029)	(.032)	(.042)	(.031)	(.039)
Ever enrolled in any kindergarten until the survey year (1 = yes):					
Baseline	−.017	.018	.019	−.018	−.021
	(.017)	(.020)	(.028)	(.023)	(.023)
Midline	−.100***	−.069	−.172***	−.090*	−.110**
	(.038)	(.046)	(.055)	(.048)	(.046)
Endline	−.147***	−.129**	−.171***	−.181***	−.113**
	(.051)	(.063)	(.055)	(.061)	(.056)

Cumulative months in project playgroups until the survey year:					
Baseline	1.482***	1.429***	1.491***	1.474***	1.430***
	(.309)	(.336)	(.518)	(.365)	(.425)
Midline	2.979***	3.246***	2.732***	3.101***	2.698***
	(.630)	(.706)	(.893)	(.781)	(.770)
Endline	4.022***	4.300***	3.707***	4.242***	3.682***
	(.888)	(1.051)	(1.075)	(1.073)	(1.050)
Cumulative months in nonproject playgroups until the survey year:					
Baseline	.042	−.111	.210	−.000	.058
	(.177)	(.140)	(.294)	(.162)	(.292)
Midline	.063	−.326	.618	−.097	.190
	(.308)	(.320)	(.456)	(.338)	(.443)
Endline	−.192	−.744	.539	−.374	−.049
	(.392)	(.482)	(.526)	(.441)	(.532)
Cumulative months in any kindergarten until the survey year:					
Baseline	−.184	−.179	−.210	−.233	−.180
	(.169)	(.202)	(.277)	(.226)	(.225)
Midline	−1.188**	−.827	2.006***	−1.114*	−1.293**
	(.476)	(.585)	(.697)	(.617)	(.556)
Endline	−2.453***	−1.917*	−3.479***	−2.624**	−2.323***
	(.811)	(1.022)	(1.003)	(1.062)	(.855)

NOTE.—The IVs are two dummy variables indicating the groups of villages assigned to middle and late treatment. The group of villages assigned to be treated early is the reference group. Each row is the result of a separate regression. Models are estimated separately for baseline, midline, and endline. Controls include child's age, gender, household size, household wealth, mother's education level, and district dummies. Robust standard errors clustered at the village level are in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

In terms of subgroup analyses, the columns marked Poor and Nonpoor reveal no striking differences in impact for enrollment rates in project playgroups. Likewise, there is evidence of substitution away from kindergartens for both groups of children—both in terms of enrollment rates (13 percentage points and 17 percentage points, respectively) and in terms of duration of enrollment (1.9 fewer months and 3.5 fewer months, respectively). This is to be expected given that kindergartens tend to charge fees while playgroups do not.

The columns marked Low Parenting Score and High Parenting Score corroborate the findings of the poor and nonpoor subgroups. While the point estimates are slightly different, they are not statistically significant.

This result seems in line with Peltzman (1973): in-kind government subsidies can have the anomalous result of decreasing total consumption even though they are intended to increase individual consumption.

C. IV Estimates of the Impact on Child Development

Table 3 presents the estimates of the impacts on child development outcomes measured using the Early Development Instrument. By the endline, there is no domain where there is a statistically significant difference between the villages treated early and the villages treated late.[24] Given that the difference in exposure to the project between villages in this analysis is only 11 months, this result is not very surprising.

Unpacking these results by wealth shows that there are no statistically significant estimates for either poor children or for nonpoor children. Similarly unpacking these results by self-reported parenting scores also shows no heterogeneity of impact among two sets of villages that were both treated albeit 11 months apart.

D. Difference-in-Differences Estimates of the Impact on Enrollment

In table 4, we present the DID estimates, as specified in equation (2).[25] These estimates are very different from the IV estimates just reported above because they compare a group of villages that received the project later to a group of villages that never received the project.

For each regression model, there are DID estimates at midline and endline.[26] The midline estimates indicate the impact of the project for children

[24] While at midline there is a statistically significant negative effect in the communications and general knowledge domain, we believe that this is an artifact of ceiling effects. In 2009, 89% of respondents reported that their children were at the maximum score and by 2010, 95% did so.

[25] This is the fixed effects model that controls for both observed and unobserved time-invariant individual characteristics.

[26] The OLS results are very close to the IV results. OLS estimates are available upon request.

Table 3
Instrumental Variables Estimated Impact on Early Development Instrument (EDI) Outcomes

	All	Poor	Nonpoor	Low Parenting Score	High Parenting Score
Physical health and well-being:					
Baseline	.103	.078	.120	−.086	.254**
	(.076)	(.096)	(.125)	(.103)	(.108)
Midline	.063	.099	.030	.043	.082
	(.072)	(.089)	(.101)	(.099)	(.091)
Endline	−.023	−.034	−.052	−.013	−.052
	(.060)	(.073)	(.088)	(.090)	(.070)
Social competence:					
Baseline	.051	.188	−.100	.071	−.010
	(.095)	(.131)	(.096)	(.123)	(.092)
Midline	−.089	−.102	−.078	−.130	−.067
	(.068)	(.076)	(.106)	(.098)	(.081)
Endline	−.111	−.073	−.157	−.161*	−.092
	(.070)	(.094)	(.097)	(.096)	(.093)
Emotional maturity:					
Baseline	.008	.056	−.069	−.028	.011
	(.082)	(.100)	(.115)	(.117)	(.096)
Midline	.015	.060	−.014	.078	−.059
	(.059)	(.082)	(.089)	(.083)	(.090)
Endline	−.020	.073	−.137	−.091	.026
	(.060)	(.082)	(.084)	(.089)	(.077)
Language and cognitive development:					
Baseline	.090	.095	.110	.100	.065
	(.066)	(.080)	(.100)	(.095)	(.088)
Midline	.051	−.008	.055	−.015	.088
	(.097)	(.120)	(.139)	(.119)	(.125)
Endline	−.049	−.071	−.003	−.053	−.054
	(.047)	(.066)	(.060)	(.070)	(.057)
Communication and general knowledge:					
Baseline	−.032	.050	−.097	.038	−.102
	(.090)	(.126)	(.095)	(.150)	(.072)
Midline	−.162**	−.128	−.192**	−.145	−.173**
	(.063)	(.079)	(.086)	(.090)	(.074)
Endline	−.026	−.085	.007	−.046	−.043
	(.128)	(.159)	(.173)	(.170)	(.173)

NOTE.—The IVs are two dummy variables indicating two groups of villages being assigned to be treated late. The group of villages being assigned to be treated early is the reference group. Models are estimated separately for baseline, midline, and endline. Controls include child's age, gender, household size, household wealth, mother's education level, and district dummies. Robust standard errors clustered at the village level are in parentheses.

* $p < .10$.
** $p < .05$.
*** $p < .01$.

Table 4
Impact on Enrollment Outcomes Using the Fixed Effects Difference-in-Differences Method

	All	Poor	Nonpoor	Low Parental Score	High Parental Score
Ever enrolled in project playgroups until the survey year (1 = yes):					
Midline	.153***	.151***	.161***	.155***	.151***
	(.015)	(.020)	(.022)	(.021)	(.021)
Endline	.269***	.293***	.240***	.269***	.270***
	(.018)	(.025)	(.026)	(.026)	(.027)
Ever enrolled in nonproject playgroups until the survey year (1 = yes):					
Midline	−.044***	−.055***	−.031*	−.018	−.072***
	(.014)	(.020)	(.018)	(.018)	(.021)
Endline	−.069***	−.089***	−.046**	−.041*	−.098***
	(.017)	(.024)	(.023)	(.022)	(.026)
Ever enrolled in any kindergarten until the survey year (1 = yes):					
Midline	−.058**	−.060*	−.057	−.088**	−.027
	(.026)	(.033)	(.042)	(.036)	(.039)
Endline	−.130***	−.155***	−.085**	−.164***	−.090**
	(.028)	(.038)	(.042)	(.039)	(.041)
Cumulative months in project playgroups until the survey year:					
Midline	1.914***	1.791***	2.145***	1.879***	1.966***
	(.158)	(.209)	(.249)	(.219)	(.234)
Endline	3.855***	4.094***	3.620***	3.785***	3.935***
	(.254)	(.352)	(.375)	(.347)	(.379)
Cumulative months in nonproject playgroups until the survey year:					
Midline	−.647***	−.687***	−.604***	−.343*	−.973***
	(.156)	(.216)	(.225)	(.200)	(.247)
Endline	−.943***	−1.198***	−.649**	−.561*	−1.354***
	(.238)	(.360)	(.287)	(.316)	(.359)
Cumulative months in any kindergarten until the survey year:					
Midline	−.590**	−.574*	−.649	−.930**	−.203
	(.270)	(.344)	(.425)	(.370)	(.402)
Endline	−1.546***	−1.860***	−1.011	−1.987***	−.998
	(.444)	(.602)	(.656)	(.611)	(.649)

NOTE.—The baseline year is used as the reference year to estimate the difference-in-differences estimates at midline and endline, controlling for child fixed effects. Robust standard errors clustered at the village level are in parentheses.

* $p < .10$.
** $p < .05$.
*** $p < .01$.

in villages that received the project later versus children in comparison villages between baseline and midline. In essence these are short-term impacts—within 1 year of project implementation. The endline estimates indicate the impact of the project after it had been implemented for about 3 years. These estimates reflect impacts in the medium term. The key identification assumption is that at the time of midline (and endline), differences in enrollment and child development outcomes between project villages and comparison villages would have been the same as those at baseline if there had not been an intervention or if the intervention had had no effect (Heckman, LaLonde, and Smith 1999).[27]

Compared with children living in comparison villages, children living in project villages had higher rates of enrollment in project playgroups but lower rates of enrollment in nonproject playgroups and kindergarten programs at midline and endline. These estimates are in line with the IV estimates reported earlier. Overall, the project led to an increase in the enrollment of children in preschool services (playgroups and kindergarten programs combined) by 9 percentage points and 7 percentage points at midline and endline, respectively. The DID estimates from the fixed effects model also suggest that the impact on the enrollment duration in preschool services overall (i.e., playgroups and kindergarten programs combined) was about 0.7 and 1.4 months at midline and endline, respectively. As in the case of the IV estimates, when the impacts are disaggregated by baseline levels of household wealth and self-reported parenting scores, the results suggest that the project's impact on enrollment rates and duration of enrollment are very similar.

E. Difference-in-Differences Estimates of the Impact on Child Development

In contrast to the IV analysis, the DID estimates in table 5 imply statistically significant impacts of the project at the end of the first year in the social competence (0.22 standard deviations) and language and cognitive development domains (0.13 standard deviations). We also find evidence of an impact on the emotional maturity domain by the time of endline (0.16 SDs).

Unpacking these results by household wealth, however, reveals a strong degree of heterogeneity. Despite having comparable impacts in terms of enrollment, poor children clearly benefited more from the project than nonpoor children. Early indications of this are visible in the midline estimates, which capture impacts after the first year. These show that poor children in project villages made greater progress than poor children in nonproject villages on the following EDI domains after the first year: social competence

[27] We have the statistical tests on children in villages treated later and the comparison group of villages. Results show there are no differences in child backgrounds between children from these two groups of villages.

(0.43 SD) and language and cognitive development (0.28 SD). By endline, there is evidence of positive impacts on poor children living in project villages in social competence (0.27 SD), emotional maturity (0.22 SD), and language and cognitive development (0.20 SD). In contrast, we find no evidence of an impact of the project on nonpoor children. Instead, we see some evidence of worse outcomes in terms of social competence (0.29 SD) and language and cognitive development (0.15 SD). However, these point estimates cannot be interpreted as negative impacts per se. This has to do with the fact that nonpoor children in treatment villages are closer to the maximum at the baseline than nonpoor children in comparison villages. Thus catch-up by the comparison group in these domains is to be expected.

When the results are disaggregated by self-reports of parenting scores, the same pattern repeats: we find statistically significant impacts on child development outcomes of children from parents with below-average parenting scores and no or negative effects among children whose parents report high parenting scores. Children whose parents had low parenting scores in project villages show better development outcomes on the EDI in terms of social competence (0.28 SD) and language and cognitive development (0.19 SD) domains than their counterparts in nonproject villages even after only 1 year of project implementation. By the time these children are reinterviewed 3 years later (at endline), there is evidence of impacts on social competence (0.2 SD), emotional maturity (0.2 SD), and language and cognitive development (0.15 SD).

In summary, our subgroup analyses clearly indicate that the impacts of the project are larger for disadvantaged children—children from poorer households and those whose parents reported below-average parenting practices. Such findings are encouraging because the services provided through the project seem to supplement the limited household resources of disadvantaged children and seem to help such children reach their potential in many important domains of child development.

VI. Discussion and Conclusion

This paper adds to the literature on early childhood interventions by evaluating a low-cost, government-sponsored, community-based early-childhood education project in rural areas of a large middle-income country using both experimental and quasi-experimental techniques. We are able to use IV estimates to identify the impact of the project by comparing the outcomes of children in two groups of villages that received the project about 11 months apart—a short difference in exposure. We are also able to use difference-in-differences estimates to identify the impact of the project by comparing the outcomes of children in a group of villages that received the project later to the outcomes of children in a group of villages that never

Table 5
Impact on Early Development Instrument (EDI) Outcomes Using the Fixed Effects Difference-in-Differences Method

	All	Poor	Nonpoor	Low Parental Score	High Parental Score
Physical health and well-being:					
Midline	−.026	.012	−.029	.011	−.052
	(.076)	(.102)	(.114)	(.105)	(.109)
Endline	.104	.143	.089	.079	.151
	(.074)	(.099)	(.112)	(.103)	(.110)
Social competence:					
Midline	.223***	.428***	−.048	.284***	.118
	(.076)	(.107)	(.105)	(.102)	(.110)
Endline	.024	.267**	−.294***	.212**	−.219**
	(.075)	(.104)	(.109)	(.105)	(.108)
Emotional maturity:					
Midline	.014	.032	−.019	−.036	.043
	(.071)	(.096)	(.106)	(.098)	(.101)
Endline	.158**	.223**	.069	.202**	.109
	(.068)	(.094)	(.099)	(.095)	(.097)
Language and cognitive development:					
Midline	.128*	.284***	−.033	.192*	.036
	(.070)	(.095)	(.101)	(.099)	(.100)
Endline	.056	.199**	−.147*	.150*	−.083
	(.060)	(.081)	(.088)	(.082)	(.085)
Communication and general knowledge:					
Midline	.075	.121	.027	.133	−.032
	(.079)	(.111)	(.110)	(.118)	(.105)
Endline	.014	.089	−.087	−.016	.010
	(.132)	(.181)	(.191)	(.187)	(.186)

NOTE.—We used the baseline year as the reference year to estimate the difference-in-differences estimates at midline and endline, controlling for child fixed effects. Robust standard errors clustered at the village level are in parentheses.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

received the project—a longer difference in exposure. A number of interesting patterns emerge: (i) The project was able to increase enrollment even though there was some substitution away from existing types of preschool services. (ii) Children from disadvantaged backgrounds benefited from exposure to the project. (iii) The impacts we observed were not only detected after the first year but on many domains were also detected when children were reinterviewed 3 years later.

The estimates from the IV model indicate that the project had a measurable impact on enrollment. However, there is no evidence of a similarly measurable impact on child development. This may be due to a confluence

of factors. First, the IV approach has increased the size of our standard errors—rendering estimated impacts statistically indistinguishable from zero. Second, and in our case perhaps more relevant, the differences in exposure this analysis is able to capture is between 6 and 11 months. Analysis from other similar interventions would suggest that this is too short a time period for measurably large impacts to be observed.

Our preferred set of estimates—those obtained from the difference-in-differences model—allow us to compare up to 40 months of project exposure against no such exposure. These estimates reveal enrollment impacts that are virtually identical to those seen in the IV analysis. They also allow us to investigate whether the project led to any improvements in children's development outcomes. On average, by endline, we only find a statistically significant impact on one domain—emotional maturity. When we unpack these estimates using either household wealth or self-reports of parenting scores, we see that the impacts were larger in magnitude and affected more aspects of development for children from disadvantaged households. For these disadvantaged subgroups of children, we see that the project leads to significant improvements in language and cognitive development, social competence, and emotional maturity.

Indonesia has rapidly moved from a developing to a middle-income country in recent years. As the country has grown economically, so too have concerns about inequality, and in particular, concerns for children and families living in poor rural communities. The results from this study would indicate that a low-cost, government-sponsored, community-based early childhood program can have modest and sustained impacts on child development—particularly for those from disadvantaged backgrounds. At the same time, however, the evidence of substitution away from existing services would suggest that future programs should be designed with such possibilities in mind so as to ensure that all children benefit from access to appropriate levels of quality services.

Appendix

Assessing the Validity of the IV Approach

Table A1
Balance between Original Assignment Strata

	Difference between Early and Middle Assignment	Difference between Early and Late Assignment
Using PODES data:		
Village population	−1,224.082	−678.613
	(1,967.947)	(693.850)
Proportion of families without electricity	−.087	−.011
	(.072)	(.032)
Public kindergartens per 1,000 people	.025	.051**
	(.015)	(.022)
Private kindergartens per 1,000 people	.026	.120**
	(.062)	(.049)
Public primary schools per 1,000 people	.033	.009
	(.048)	(.054)
Private primary schools per 1,000 people	.005	−.016
	(.062)	(.059)
Using analysis sample:		
Age of child (in years)	.044	.016
	(.029)	(.014)
Household size	.254	.113
	(.164)	(.092)
Wealth *Z*-score	.108	.075
	(.144)	(.081)
Parenting scores *Z*-score	−.218	−.023
	(.148)	(.073)
Mother's education = 1 for primary school or less	.052	.016
	(.058)	(.033)
Child's gender: male = 1	.010	.032
	(.049)	(.027)

NOTE.—*N* for analysis sample = 1,436. *N* for Village Potential Statistics (PODES) = 210 villages. Tests of differences between original assignment strata are reported. All rows are results from separate regressions. Dummies for districts are included. Robust standard errors clustered at the village level are in parentheses.

** Significant difference at the .05 level.

Table A2
Original and Actual Allocation of Villages (Number of Villages)

	Original Random Assignment				
Actual Group Allocation	Group 1	Group 2	Group 3	Comparison Group	Total
Early treatment	89	10	6	0	105
Late treatment	9	10	93	1	113
Comparison group	2	0	1	89	92
Total	100	20	100	90	310

NOTE.—Each cell shows the number of villages.

Table A3
First-Stage Regression Results for Instrumental Variables Estimation

	Middle Assignment Group (= 1)	Late Assignment Group (= 1)	Control Variables	R^2	Number of Observations
All:					
Baseline	−.393***	−.844***			
	(.119)	(.033)	Yes	.744	1,317
Midline	−.393***	−.844***			
	(.119)	(.033)	Yes	.744	1,317
Endline	−.391***	−.842***			
	(.118)	(.033)	Yes	.745	1,317
Poor children:					
Baseline	−.370***	−.864***			
	(.118)	(.033)	Yes	.745	764
Midline	−.369***	−.863***			
	(.118)	(.033)	Yes	.745	764
Endline	−.369***	−.860***			
	(.117)	(.033)	Yes	.746	764
Nonpoor children:					
Baseline	−.503***	−.831***			
	(.154)	(.040)	Yes	.763	553
Midline	−.503***	−.831***			
	(.154)	(.040)	Yes	.763	553
Endline	−.495***	−.833***			
	(.155)	(.040)	Yes	.763	553
Below-average parenting score:					
Baseline	−.394***	−.827***			
	(.128)	(.038)	Yes	.716	675
Midline	−.396***	−.828***			
	(.128)	(.037)	Yes	.716	675
Endline	−.393***	−.825***			
	(.126)	(.038)	Yes	.718	675
Above-average parenting score:					
Baseline	−.386***	−.860***			
	(.135)	(.037)	Yes	.779	641

Table A3 (*Continued*)

	Middle Assignment Group (= 1)	Late Assignment Group (= 1)	Control Variables	R^2	Number of Observations
Midline	−.384***	−.860***			
	(.135)	(.037)	Yes	.778	641
Endline	−.385***	−.860***			
	(.135)	(.037)	Yes	.779	641

NOTE.—Child characteristics and household background variables are included. These are child's age, gender, household size, household wealth, and mother's education level. All standard errors (in parentheses) are robust and clustered at the village level. Separate first-stage results are reported because of variations in time-varying control variables. Regression results are only for villages included in the randomization and exclude comparison villages. Sample size for each year is smaller than in table 2 and varies from 1,311 to 1,317. *F*-statistics for the two IVs are larger than 100 in all first-stage regressions.

*** Significant difference at the .01 level.

References

Alatas, Hafid, Sally Brinkman, Mae Chu Chang, Titie Hadiyati, Djoko Hartono, Amer Hasan, Marilou Hyson, Haeil Jung, Angela Kinnell, Menno Pradhan, et al. 2013. Early childhood education and development services in Indonesia. In *Education in Indonesia*, ed. Daniel Suryadarma and Gavin Jones. Singapore: Institute of Southeast Asian Studies.

Armecin, Graeme, Jere Behrman, Paulita Duazo, Sharon Ghuman, Socorro Gultiano, Elizabeth King, and Nannette Lee. 2006. Early childhood development through an integrated program: Evidence from the Philippines. Working Paper no. 3922, World Bank Policy Research, Washington, DC.

Barnett, W. Steven. 1997. Costs and financing of early child development programs. In *Early child development: Investing in our children's future*, ed. Mary E. Young. Amsterdam: Elsevier.

Behrman, Jere R., Yingmei Cheng, and Petra E. Todd. 2004. Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *Review of Economics and Statistics* 86, no. 1:108–32.

Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. The effect of pre-primary education on primary school performance. *Journal of Public Economics* 93, no. 1:219–34.

Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics* 92, no. 5:1416–40.

Black, Maureen M., Susan P. Walker, Lia C. H. Fernald, Christopher T. Andersen, Ann M. DiGirolamo, Chunling Lu, Dana C. McCoy, Günther Fink, Yusra R. Shawar, Jeremy Shiffman, et al. 2016. Early childhood development coming of age: Science through the life course. *Lancet* 389:77–90.

Bouguen, Adrien, Deon Filmer, Karen Macours, and Sophie Naudeau. 2013. Impact evaluation of three types of early childhood development interventions in Cambodia. Working Paper no. 6540, World Bank Policy Research, Washington, DC.

Brinkman Sally A., Angela Kinnell, Amelia Maika, Amer Hasan, Haeil Jung, and Menno Pradhan. 2017. Validity and reliability of the early development instrument in Indonesia. *Child Indicators Research*. Forthcoming.

Evans, Judith L., Robert G. Meyers, and Ellen Ilfeld. 2000. *Early childhood counts: A programming guide on early childhood care for development*. Washington, DC: World Bank Publications.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* 344, no. 6187:998–1001.

Hasan, Amer, Marilou Hyson, and Mae Chu Chang. 2013. *Early childhood education and development in poor villages of Indonesia: Strong foundations, later success*. Washington, DC: World Bank.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. The economics and econometrics of active labor market policy. In *Handbook of labor economics*, vol. 3, ed. Orley Ashenfelter, Robert Layard, and David Card, 1–865. Amsterdam: Elsevier.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94, no. 1:114–28.

Martinez, Sebastian, Sophie Naudeau, and Vitor Pereira. 2012. The promise of preschool in Africa: A randomized impact evaluation of early childhood development in rural Mozambique. World Bank, Washington, DC.

Nores, Milagros, and W. Steven Barnett. 2010. Benefits of early childhood interventions across the world: (Under)Investing in the very young. *Economics of Education Review* 29, no. 2:271–82.

Peltzman, Sam. 1973. The effect of government subsidies-in-kind on private expenditures: The case of higher education. *Journal of Political Economy* 81, no. 1:1–27.

Pradhan, Menno, Sally A. Brinkman, Amanda Beatty, Amelia Maika, Elan Satriawan, Joppe de Ree, and Amer Hasan. 2013. Evaluating a community-based early childhood education and development program in Indonesia: Study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group. *Trials* 14, no. 1:1–16.

Rao, Nirmala, Jin Sun, Jessie M. S. Wong, Brendan Weekes, Patrick Ip, Sheldon Shaeffer, Mary Young, Mark Bray, Eva Chen, and Diana Lee. 2014. *Early childhood development and cognitive development in developing countries: A rigorous literature review*. London: Department for International Development.

Sayre, Rebecca K., Amanda E. Devercelli, Michelle J. Neuman, and Quentin Wodon. 2015. *Investing in early childhood development: Review of the World Bank's recent experience*. Washington, DC: World Bank Publications.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20, no. 4:518–29.

Yousafzai, Aisha K., Muneera A. Rasheed, Arjumand Rizvi, Robert Armstrong, and Zulfiqar A. Bhutta. 2014. Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: A cluster-randomised factorial effectiveness trial. *Lancet* 384, no. 9950:1282–93.