March 2020

Passive versus Active Service Delivery: Comparing the Effects of Two Parenting Interventions on Early Cognitive Development in Rural China

Jingdong Zhong, Renfu Luo, Sean Sylvia, Sarah-Eve Dill, Alexis Medina, Scott Rozelle

Abstract

We present the results of a cluster-randomized controlled trial that evaluates the effects of a free, center-based parenting intervention on early cognitive development and parenting practices in 100 rural villages in China. We then compare these effects to a home-based intervention conducted in the same region, using the same parenting curriculum and public service system. We find that the center-based intervention significantly improved children's cognitive skills by 0.11 standard deviations, accompanied by increases in the material investments, time investments, and parenting skills of caregivers. The average impact of the center-based intervention, however, was approximately half that of the home-visiting intervention. Analysis of the possible mechanisms suggests that the difference in effects was driven primarily by different patterns of compliance. Although children with lower levels of initial skills at baseline benefited the most from the center-based intervention, they were less likely to participate in the program.

Keywords: Center-based parenting intervention; home-based parenting intervention; early cognitive development; randomized controlled trial; program participation

---

Working Paper

March 2020

reap.fsi.stanford.edu

REAP
Rural Education Action Program

Passive versus active service delivery: Comparing the effects of two parenting interventions on early cognitive development in rural China

Jingdong Zhong[a], Renfu Luo[b*], Sean Sylvia[c], Sarah-Eve Dill[d], Alexis Medina[e], Scott Rozelle[f]

[a] China Center for Agricultural Policy, School of Advanced Agricultural Science, 409 Wangkezhen Bldg., Peking University, No. 5 Yiheyuan, Haidian District Beijing, China 100871. Email: jdzhong@pku.edu.cn

[b] China Center for Agricultural Policy, School of Advanced Agricultural Science, 409 Wangkezhen Bldg., Peking University, No. 5 Yiheyuan, Haidian District Beijing, China 100871. Email: luorf.ccap@pku.edu.cn

[c] Gillings School of Global Public Health, 1101-D McGavran-Greenberg Bldg., University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 Chapel Hill, USA. Email: sean.sylvia@unc.edu

[d] Rural Education Action Program, Freeman Spogli Institute for International Studies, 616 Jane Stanford Way, Stanford University, Stanford, CA, USA 94305. Email: sedill@stanford.edu

[e] Rural Education Action Program, Freeman Spogli Institute for International Studies, 616 Jane Stanford Way, Stanford University, Stanford, CA, USA 94305. Email: amedina5@stanford.edu

[f] Rural Education Action Program, Freeman Spogli Institute for International Studies, 616 Jane Stanford Way, Stanford University, Stanford, CA, USA 94305 Email: rozelle@stanford.edu

[*]Corresponding Author

Acknowledgements:
We are grateful for funding from the 111 Project (Grant no: B16031), the National Natural Science Foundation of China (Grant no. 71703083, 71703084, 71803108), the International Initiative for Impact Evaluation (3ie), UBS Optimus Foundation, China Medical Board, Bank of East Asia, Huaqiao Foundation, and the Bill and Melinda Gates Foundation-HBGDki Initiative. We would also like to thank the National Health Commission and its local cadres for their unparalleled assistance in implementing this study.

Passive versus active service delivery: Comparing the effects of two parenting interventions on early cognitive development in rural China

Abstract

We present the results of a cluster-randomized controlled trial that evaluates the effects of a free, center-based parenting intervention on early cognitive development and parenting practices in 100 rural villages in China. We then compare these effects to a home-based intervention conducted in the same region, using the same parenting curriculum and public service system. We find that the center-based intervention significantly improved children's cognitive skills by 0.11 standard deviations, accompanied by increases in the material investments, time investments, and parenting skills of caregivers. The average impact of the center-based intervention, however, was approximately half that of the home-visiting intervention. Analysis of the possible mechanisms suggests that the difference in effects was driven primarily by different patterns of compliance. Although children with lower levels of initial skills at baseline benefited the most from the center-based intervention, they were less likely to participate in the program.

Keywords: center-based parenting intervention; home-based parenting intervention; early cognitive development; randomized controlled trial; program participation

1. Introduction

Early childhood development (ECD) is central to the future of low- and middle-income countries (LMICs). Although child mortality rates in many LMICs have decreased dramatically in recent decades, approximately 250 million children under five years old in LMICs remain at risk of not reaching their developmental potential (Black et al., 2017). Because developmental outcomes in early childhood are critical to adult outcomes, including labor market returns (Heckman et al., 2010; Huggett et al., 2011; Gertler et al., 2014), health status (Heckman, 2007; Campbell et al., 2014), and social mobility (Heckman and Mosso, 2014), early developmental delays can have significant negative effects on later quality of life. At a macro level, widespread developmental delays can inhibit LMICs from raising human capital, which has been shown to be critical for sustaining long-term economic growth and development (Li et al., 2017).

In light of this concern, a growing body of research has provided strong theoretical and empirical support for targeted ECD intervention programs that train caregivers in stimulating parenting practices. Cunha and Heckman (2007, 2008, 2009) and Cunha et al. (2010) have established a framework that shows that early parenting interventions can effectively boost the skills development of disadvantaged children and that parental investments in the earliest stages of life can effectively increase the impacts of later-stage investments. Empirical studies have further shown that parenting programs that target caregivers of young children can have meaningful effects on early skills. A recent meta-analysis of 21 randomized controlled trials (RCTs) of parenting interventions for children aged 0–3 conducted in LMICs since 2000 have found that parenting programs improved child cognitive development by 0.42 standard deviations (*SD*), on average (Aboud and Yousafzai, 2015).

Although past research has demonstrated the positive and significant impacts of ECD interventions, the question remains as to how to effectively deliver ECD intervention programs at scale. Building new infrastructure and employing new workers to deliver ECD interventions can be costly, particularly for LMICs that face more stringent resource constraints (Richter et al., 2017). As an alternative, international organizations (e.g., World Bank, United Nations, World Health Organization) have proposed that LMICs integrate ECD interventions into existing public infrastructure and public service systems (Chan, 2013; Richter et al., 2017).

Although few ECD intervention programs in LMICs have been upscaled for widespread delivery, studies of potentially scalable ECD interventions have relied mainly on one of two delivery models: home-based delivery or center-based delivery (Abdul Latif Jameel Poverty Action Lab [J-PAL], 2019). Home-based ECD interventions typically involve regular visits from parenting trainers who conduct one-to-one parenting lessons in the home, whereas center-based interventions require caregivers to bring their children to a central location to participate in parenting training sessions. Several mechanisms may drive differences in the effectiveness of the two models, and it remains unclear which model may be more effective and cost-effective in different contexts (J-PAL, 2019). To date, there have been no randomized trials that directly compare the two. Moreover, although the two models have been evaluated independently, past studies have been conducted by different research teams, using different curriculums, intervention protocols, and outcome measures in different regions, complicating comparison.

To address this gap, the goal of this study is to evaluate the effects of a free, center-based parenting intervention on ECD outcomes and parenting practices in rural China. We then compare the effects of the center-based delivery model against a home-based intervention conducted in the same region of rural China, using the same parenting curriculum and public service system. To implement the center-based delivery model, we worked with China's National Health Commission (NHC) to conduct a large-scale, cluster-RCT of a center-based ECD parenting intervention in 100 villages in an underdeveloped rural area in northwestern China.

This center-based program was implemented by the same research team in the same target area as a home-based program evaluated by Sylvia et al. (2018), used the same curriculum and public service system (NHC) to deliver the intervention, and measured the same outcomes, using the same instruments. Our study compares the program treatment effects on child skill development and secondary parenting outcomes (caregiver investment and parenting skills) and examines potential sources of differential impacts between the two interventions.

Both delivery models are viable in rural China, where there is a need for scalable ECD interventions. Like other LMICs, China is facing widespread early childhood developmental delays, with nearly half (49%) of rural children aged 0–3 years as exhibiting cognitive delay (Luo et al., 2017; Yue et al., 2017; Bai et al., 2019; Wang et al., 2019). At the same time, China has a large public infrastructure and abundant human resources, which can be leveraged to implement ECD interventions at scale should the government decide that this is a priority. Due to large-scale rural-to-urban labor migration and other demographic trends, there is a large number of disused schoolhouses, cultural centers, and office spaces in rural areas that can be repurposed for ECD centers. China also has one of the largest health bureaucracies in the world, NHC, available to implement ECD interventions. Since the national government relaxed the One Child Policy in 2016, the Family Planning Commission, now part of NHC, has shifted its focus from managing the population quantity to improving the quality of human capital, including improving investments in ECD (Wu et al., 2012).

We find that the center-based parenting intervention significantly improved the cognitive skills of treatment children by 0.11 *SD* after 12 months. The effects were accompanied by increases in caregiver investments in children, including increased material investments in toys and picture books, increased time investments in stimulating parenting activities, and improved parenting skills. Moreover, children with low cognitive skills at baseline showed significant improvements as a result of the intervention, while children with higher baseline cognitive skills did not, indicating that the intervention provided the greatest benefit to children with the greatest need.

The center-based intervention was, however, less effective than was the home-based intervention, producing smaller average impacts on child cognitive development, caregiver investments, and parenting skills. Further analysis of the two interventions indicates that this difference may be due to the differing nature of compliance, or uptake, of the two interventions. Whereas parents of children with low baseline cognitive skills tended to select out of the center-based program (i.e., their levels of participation were lower than were the parents of children with higher baseline cognitive skills), the home-visiting program effectively provided parental training to children with both low and high baseline cognitive skills. Because the two programs had larger impacts on children with lower levels of skills before intervention, as has been found with parenting programs in other countries, our findings suggest that the greater compliance of more vulnerable children in the home-visiting program may have led to larger average impacts than in the center-based program.

This study contributes to the growing literature on the effects of ECD interventions delivered through public resources in LMICs. To the best of our knowledge, this study is the first to compare two popular delivery models for ECD interventions, using the same target region, curriculum, and measurements. The results emphasize the importance of program participation in ensuring program effectiveness and cost effectiveness. Due to selection in program compliance, the active home-based delivery model may be more effective and cost-effective than is the more passive form of service delivery through parenting centers in rural areas.

The remainder of this paper is organized as follows. Section 2 presents the sample, experimental design, and empirical approach used to analyze the center-based intervention.

Section 3 contains the findings for the center-based intervention. Section 4 provides a comparison of the differences in design and effects between the center-based and home-based interventions, and Section 5 concludes.

2. Methods

2.1 Sampling and Randomization

Our trial of the center-based parenting intervention was conducted in 22 nationally designated poverty counties[1] in a northwestern province of China. According to statistics from the National Bureau of Statistics of China (NBSC), for 2016, the per capita income of rural residents in the sample province was 9,396 yuan ($1,337 USD), around the national median of 12,363 yuan ($1,759 USD) for China's rural areas.

Within the sample region, we followed a three-step protocol to select the study sample. First, in each sample county, we obtained a list of all townships from local NHC officials. We excluded the township in each county that housed the county seat (which tend to be wealthier and more urban than the average rural township) as well as townships that did not have any villages with a population of 800 or more. From the remaining townships, we then randomly selected 100 for inclusion in the sample. Second, within each township, we randomly selected one village to participate in the study, totaling 100 villages. To ensure that all sample villages would have the potential space to conduct the center-based parenting intervention, villages that could not supply a 60–80 m^2 space for the intervention site were excluded. If a village did not have the available space, it was replaced with a randomly selected village from within the same township. Finally, a list of all registered births over the past 24 months was obtained from the

[1] To focus effort on its rural poverty alleviation program, the central government in China used per capita net income to designate certain rural counties as national poverty counties (The State Council of the People's Republic of China, 2006).

local NHC official in each sample village, and all children in the desired age range (6–24 months) and their caregivers were enrolled in the interventional study. In total, 1,720 children in 100 villages were sampled at baseline.

After the baseline survey, the research team randomly allocated 50 sample villages to the treatment arm of the study and 50 villages to the control arm. In total, 881 sample children and their caregivers were assigned to the treatment arm. The remaining 839 children and caregivers were assigned to the control arm (Figure 1).

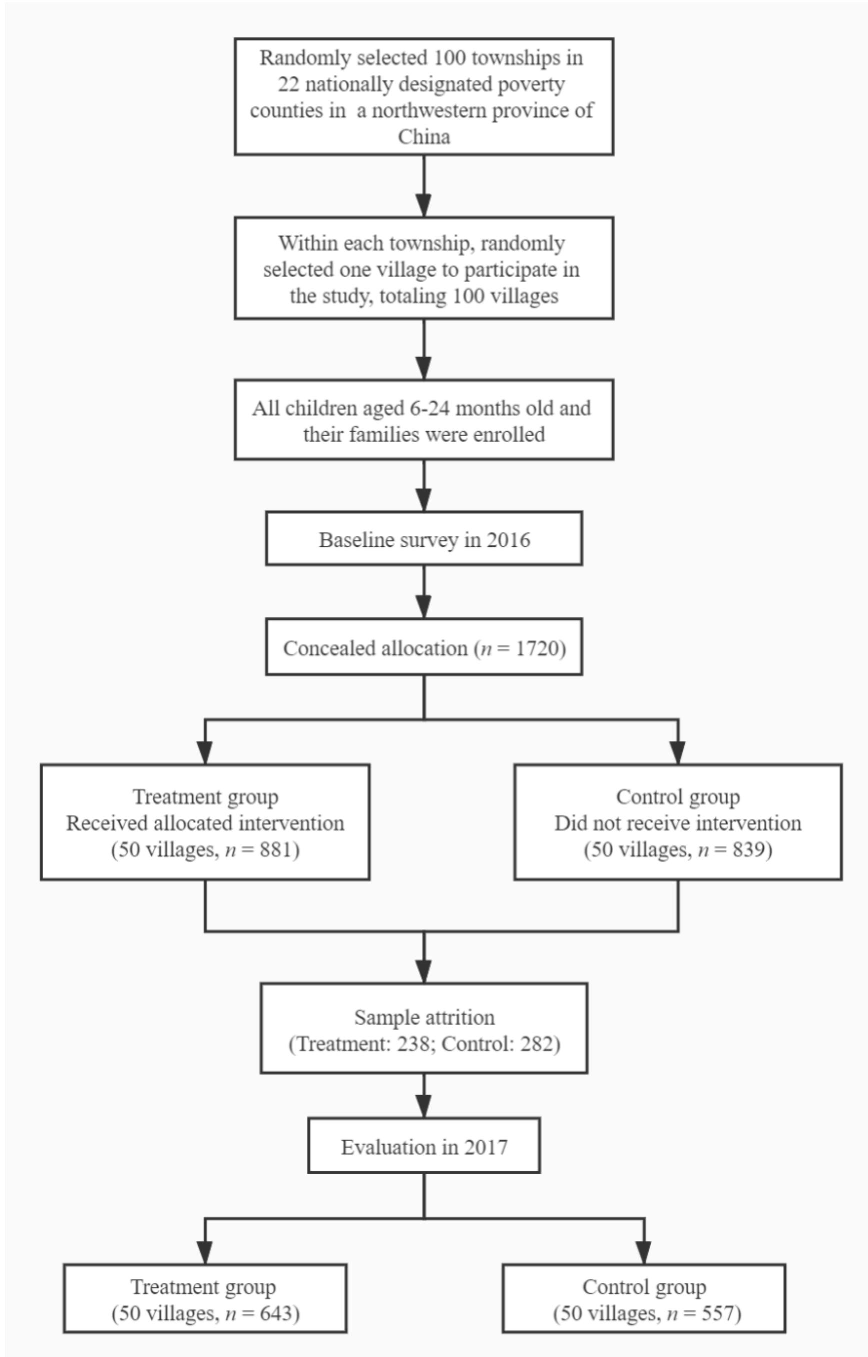

Figure 1. Flowchart of the center-based parenting intervention.

2.2 Intervention: A center-based parenting program

In each of the 50 treatment villages, one parenting center was established in a centrally located building (e.g., a repurposed schoolhouse, cultural center, office space) provided by the village committee (Figure 2). Each parenting center was renovated to be child friendly, with colorful walls, non-lead-based paint, and soft floors. All parenting centers contained a large play area, as well as a standard set of toys, baby books, and decorations provided by the research team. Each parenting center also contained a smaller room for one-on-one parenting training sessions. The parenting centers were designed to be open 5 hours a day, 6 days a week. According to monitoring data collected by the research team, the parenting centers were open for an average of 279 days during the first year of operation. Caregivers were encouraged to bring their children to the parenting centers during open hours but were not allowed to leave their children alone in the parenting centers.

Figure 2. Images of parenting centers established in repurposed buildings in rural villages of China.

Clockwise from left: public village building used for parenting center in a sample village; established parenting center in a sample village; caregiver and child reading in a parenting center; children and caregivers playing in parenting center.

In addition to parents' having access to the center whenever it was open, each parenting center was staffed by two parenting trainers from the local branch of the NHC, who conducted weekly one-to-one lessons for caregivers on interactive parenting practices to stimulate child development. Before the start of the intervention, all parenting trainers underwent comprehensive training in child development and the structured week-by-week parenting curriculum used in this intervention.

The parenting curriculum is adapted from Reach Up and Learn, a curriculum developed and evaluated in Jamaica by Walker et al. (2011) and used in ECD intervention studies in multiple LMICs (J-PAL, 2019). The curriculum aims to teach caregivers how to interact with their children through age-appropriate, stage-based stimulating activities. It consists of weekly interactive training sessions that target caregivers of children aged 6 to 36 months, and each session consists of two age-appropriate activities that involve both caregivers and children. The curriculum was adapted by child development experts in China to fit the context of rural China, and it has previously undergone field testing and evaluation in rural China by members of the research team (Sylvia et al., 2018).

In addition to the two parenting trainers, the research team also hired one local center manager for each parenting center. The center manager was responsible for managing all center activities, including recording program participation. For each visit to the parenting center, the center manager recorded the caregiver’s name, the date, and the relationship between the caregiver and the child.

2.3 Data collection

We collected data in two surveys rounds, which we refer to as the baseline and endline survey, respectively. The baseline survey was conducted in August 2016, after which we began the intervention in treatment villages. The endline survey was conducted one year later, in August 2017. Both surveys were identical and collected data on child development outcomes, parental investment and parenting skills, and demographic characteristics of sample children and households.

Child developmental outcomes. The primary outcome of interest in this study are measures of child skill development. In each survey round, children were administered the third edition of the Bayley Scales of Infant Development (BSID-III; Bayley, 2006). The BSID-III includes four scales that assess cognitive, language, motor, and social-emotional skills, respectively. The BSID-III has been formally adapted to the Chinese language and environment and used in multiple studies across rural China (Wang et al., 2019).

The BSID-III test was administered by trained enumerators, using a standardized toy kit and a detailed scoring sheet. A child's scores on the BSID-III are determined by the child's performance on a series of tasks, adjusted for age in months and premature birth. The caregiver of each child was present but was not allowed to assist the child during the test. All enumerators attended a one-week intensive training course on BSID-III administration, including 2.5 days of experiential learning in the field, before the survey.

To combine the four BSID-III scales (cognitive, language, motor, and social-emotional) into a single index, we follow Cunha and Heckman (2008) and Cunha et al. (2010) to construct a latent factor measure for child skills, using a dedicated measurement system. We estimated the measurement system separately at baseline and endline as:

$$m_{ik}^{h} = \mu_{k}^{h} + \alpha_{k}^{h} h_i + \delta_{ik}^{h} \tag{1}$$

where m_{ik}^{h} is the k-th measure of child i's skills, μ_{k}^{h} is the mean of the k-th measure of child skills, α_{k}^{h} is the factor loading of the k-th measure, and δ_{ik}^{h} is mean zero measurement error term, which is assumed independent of the latent factor, h_i. This measurement system is assumed invariant to the treatment assignment. That is, any differences in the observed measures of child skills between the control group and treatment group result only from a change in the latent child skill factor. After estimating the measurement system, we use the Bartlett (1937) scoring method to predict the factor score of latent factors for each child, based on the estimated means and factor loadings. The predicted child skill factor is standardized by the distribution of the control group. Further details about the measurement system are described in Appendix A.

Parental investments and parenting skills. The parenting curriculum was designed to affect child development by increasing the parental investments and parenting skills of caregivers. To assess the program effects on these secondary outcomes, we collected data on the material investments and time investments as well as parenting skills of the child's primary caregiver (defined as the individual most responsible for child's daily care, typically either the mother or paternal grandmother). The primary caregiver was administered a detailed questionnaire adapted from the Family Care Indicators (FCI), which was developed by UNICEF to measure the home environment of young children in developing countries (Frongillo et al., 2003). Previous studies have demonstrated that the FCI is a reliable measure of parenting and the home environment in developing settings (Hamadani et al., 2010). The FCI has been adapted to the Chinese language and context and used in previous studies across rural China (Wang & Yue, 2019; Wang & Zheng, 2019). A full list of items included in the FCI are reported in Table B1 of Appendix B. As seen in the table, the Cronbach's alpha coefficient for all items is larger than 0.8, indicating that it has high internal consistency in our sample (Nunnally, 1978).

Using caregiver responses to the FCI, we created two measures of parental outcomes: material investments and time investments. We use six variables to measure caregiver material investments, including sources of play materials, varieties of play materials, total number of play materials, number of picture books, number of books for adults, and number of magazines and newspapers in the home. Time investments were calculated based on whether caregivers had participated in each of five at-home play activities with their child in the past three days: reading books or looking at picture books, telling stories, singing songs, playing with toys, and spending time in naming things, counting, or drawing.

We also collected information on the parenting skills of caregivers. To measure parenting skills, we asked each child's primary caregiver a series of questions about his or her beliefs and attitudes toward parenting, including whether the caregiver feels a duty to help the baby understand the world, whether the caregiver thinks it is important to play or read with the baby, and whether the caregiver knows how to play or read with the baby.

As we did for child skills, we developed a dedicated measurement system that related all observed measures of caregiver material investments, time investments, and parenting skills to their corresponding latent factors, using equation (1). We estimated the measurement system at baseline and at endline for caregiver material investments, time investments, and parenting skills. The predicted material investment factor, time investment factor, and parenting skill factor are all standardized by the distribution of the control group.

Demographic characteristics. Finally, we collected demographic information on child and household characteristics. Child characteristics include the child's gender, age in months, whether the child had a low birth weight, whether the child was born through a natural birth (as opposed to caesarean section), and whether the child was premature. The child's age and premature birth status were taken from his or her birth certificate. Household characteristics include whether the mother is the primary caregiver, the primary caregiver's age and level of education, whether the child has older siblings, and whether the household receives social security support through China's minimum living standard guarantee program.

Program participation. In addition to the baseline and endline surveys, we collected information on program compliance for all children and caregivers in the treatment group throughout the one-year intervention. As noted above, the manager of each parenting center recorded each visit to the center, including the child's name, the date, and the relationship between the caregiver and the child, using a registration form designed by the research team. The research team also made phone calls to randomly chosen households to double-check the accuracy of the records. Based on these records, we calculated the average number of center visits per month for each treatment household.

2.4 Baseline characteristics, balance, and attrition of center-based sample

Table 1 presents the descriptive statistics and balance test of baseline characteristics between the control and treatment groups. All *p*-values account for clustering within villages, and the differences in child and household characteristics are all insignificant across the two groups. We also ran a joint significance test for balance by regressing the treatment status on all baseline characteristics reported in the table and tested that the coefficients of all covariates are jointly zero. The *p*-value of this test is 0.898.

Table 1. Descriptive statistics and balance test of demographic characteristics

Variable	Control (n = 839)	Treatment (n = 881)	p-value
Panel A: Child Characteristics			
Male (yes = 1)	0.52	0.51	0.562
	(0.02)	(0.02)	
Age in months	14.49	14.24	0.198
	(0.22)	(0.21)	
Low birth weight (yes = 1)	0.05	0.04	0.547
	(0.01)	(0.01)	
Natural birth (yes = 1)	0.65	0.63	0.539
	(0.02)	(0.02)	
Premature (yes = 1)	0.05	0.04	0.738
	(0.01)	(0.01)	
Panel B: Household Characteristics			
Caregiver age (years)	35.37	35.51	0.794
	(0.62)	(0.46)	
Caregiver years of schooling	8.18	8.11	0.966
	(0.13)	(0.19)	
Mother is primary caregiver (yes = 1)	0.70	0.70	0.708
	(0.02)	(0.02)	
Child has elder siblings (yes = 1)	0.51	0.51	0.999
	(0.02)	(0.02)	
Household receives social security support (yes = 1)	0.10	0.12	0.599
	(0.01)	(0.02)	

Note. Standard errors presented in parentheses. The p-values account for clustering at the village level. An omnibus balance test, conducted by regressing treatment status on all listed covariates and conducting an F-test, which cannot reject that the coefficients are jointly zero, yields a p-value of 0.898.

Panel A provides the baseline statistics for child characteristics. In our sample, children were just over 14 months old, on average, at baseline. Slightly over half (51%) of the children were male. About 4% of the children were born with low birth weight, and 64% of children were born naturally (the balance, 36%, were born by caesarean section). Fewer than 5% of sample children were premature at birth.

Panel B provides baseline statistics for caregivers and households. The mother was the primary caregiver in 70% of households.[2] The average age of primary caregivers (mothers and others) is around 35 years at baseline, and caregivers had slightly over 8 years of schooling on average. Nearly 50% of sample children had older siblings, and approximately 11% of households were receiving social security support at the time of the baseline survey.

Table 2 presents the descriptive statistics and balance tests for measures of child skills at baseline of the center-based program. In our sample, at baseline, the mean scores of the cognitive, language, motor, and social-emotional scales were 96.15, 92.72, 97.42, and 85.92, respectively. For cognitive, language, and social-emotional development, the mean scores in our sample are about 1 *SD* lower than the expected means of healthy population.[3] In addition, at baseline, 53% of children exhibited cognitive delay, 60% exhibited language delay, 36% exhibited motor delay, and 43% exhibited social-emotional delay. The differences in child development outcomes are all insignificant across the two groups. The *p*-value of the joint significance test is 0.706, which means that all coefficients are jointly zero.

[2] In rural areas, it is common that caregivers for children are grandmothers while their mothers out-migrate to urban areas for work (Yue et al., forthcoming).

[3] In a healthy population, the mean BSID-III score (*SD*) is expected to be 105 (9.6), 109 (12.3), 107 (14), and 100 (15) for the cognitive scale, language scale, motor scale, and social-emotional scale, respectively (Wang et al., 2019).

Table 2. Descriptive statistics for child skills at baseline

Variable	Control (n = 839)	Treatment (n = 881)	p-value
Panel A: BSID-III score			
Cognitive score	96.07	96.22	0.442
	(0.73)	(0.83)	
Language score	93.08	92.37	0.503
	(0.75)	(0.79)	
Motor score	97.90	96.95	0.447
	(0.93)	(0.99)	
Social-emotional score	85.99	85.84	0.994
	(1.02)	(0.82)	
Panel B: Developmental delay			
Cognitive delay (score < 95.4)	0.53	0.54	0.816
	(0.03)	(0.03)	
Language delay (score < 96.7)	0.59	0.60	0.913
	(0.02)	(0.03)	
Motor delay (score < 93)	0.35	0.38	0.337
	(0.02)	(0.02)	
Social-emotional delay (score < 85)	0.43	0.43	0.982
	(0.03)	(0.03)	

Note. The statistics are the sample mean, and the standard error is presented in parentheses. We regressed the treatment status on all baseline child skill scores and skill delays reported. The p-value on each coefficient accounts for clustering at the village level. We conducted an F-test, which cannot reject that all coefficients are jointly zero, for which the p-value is 0.706.

Table 3 presents the descriptive statistics and balance tests for measures of caregiver material investments, time investments, and parenting skills at the baseline of the center-based program. At baseline, each household had a mean of 2.26 books, and only around 20% of caregivers reported reading books to their child. Only 17% of caregivers reported telling stories to their child, and 41% of caregivers reported singing songs to their child. Only two out of the 16 measures of parental investments and skills were unbalanced (number of magazines and newspapers in the home and whether caregiver knows how to play with the baby), with slightly higher scores in the control group at baseline. The p-value of the joint significance test is 0.281, which cannot reject that the coefficients of all baseline measures are jointly zero.

Table 3. Descriptive statistics for material investments, time investments, and parenting skills at baseline

Variable	Control (n = 839)	Treatment (n = 881)	p-value
Material investments			
Number of play material sources	2.25 (0.07)	2.31 (0.06)	0.327
Number of play material varieties	3.63 (0.14)	3.68 (0.12)	0.900
Number of picture books	1.68 (0.06)	1.69 (0.05)	0.690
Number of play materials	3.57 (0.09)	3.59 (0.08)	0.278
Number of books (except picture books)	2.29 (0.08)	2.27 (0.07)	0.913
Number of magazines and newspapers	1.62 (0.06)	1.50 (0.05)	0.056*
Time investments			
Read books or looked at picture books with child in last 3 days	0.19 (0.02)	0.20 (0.02)	0.379
Told stories to child in last 3 days	0.18 (0.02)	0.17 (0.02)	0.933
Sang songs with child in last 3 days	0.42 (0.03)	0.41 (0.02)	0.617
Played with the child with toys in last 3 days	0.67 (0.03)	0.69 (0.02)	0.321
Spent time with child in naming things, counting, or drawing in last 3 days	0.41 (0.03)	0.42 (0.02)	0.978
Parenting skills			
Caregiver feels duty to help baby understand the world	6.29 (0.12)	6.13 (0.11)	0.134
Caregiver finds it important to play with baby	5.08 (0.11)	5.00 (0.10)	0.843
Caregiver knows how to play with baby	4.83 (0.11)	4.55 (0.10)	0.036**
Caregiver finds it important to read stories to baby	4.33 (0.10)	4.36 (0.09)	0.152
Caregiver knows how to read stories to baby	4.20 (0.11)	3.94 (0.12)	0.105

Note. The statistics are the sample mean, and the standard error is presented in parentheses. We regressed the treatment status on all baseline material investments, time investments, and parenting skills. The p-value for each coefficient accounts for clustering at the village level. We conducted an F-test, which cannot reject that all coefficients are jointly zero, for which the p-value is 0.281.
$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Attrition was relatively high in both the treatment and control groups in our sample. As shown in Figure 1, of the 881 children in the treatment group at baseline, only 643 children participated the endline survey, a 27% attrition rate. In the control group, 557 of the original 839 children participated in the endline survey, an attrition rate of 33%. Most of the attrition in our sample is due to family out-migration to other parts of the prefecture, province, or nation. Importantly, as shown in Table B2, the attrition rate is balanced between the treatment and control groups. In Columns 3 and 4, the *p*-value of the Chow test is 0.36, which cannot reject that the correlates of attrition are similar in the control and treatment groups.

2.5 Estimation strategy

Average treatment effects. In a randomized controlled trial, comparisons of the mean(s) of the outcome variable(s) between the treatment control groups provide unbiased estimates of the program effects on outcomes due to random treatment assignment. In this study, following the methods of Bruhn and McKenzie (2009) and Sylvia et al. (2018), we controlled for randomization strata (county) and the baseline value of the dependent variable to increase power. We estimated the intent-to-treat (ITT) effects of the parenting intervention by ordinary least squares (OLS) using the following ANCOVA specification:

$$h_{ijt} = \alpha_1 + \beta_1 D_{ij} + \gamma_1 h_{ij(t-1)} + \tau_s + \xi_{ij} \tag{2}$$

where h_{ijt} is the skills of child *i* in village *j* at endline, D_{ij} is a dummy that indicates the treatment assignment of child *i* in village j, $h_{ij(t-1)}$ is the skills of child *i* in village *j* at baseline, and τ_s is the strata (county) fixed effects. The coefficient β_1 captures the ITT effects of the center-based parenting intervention on child skill development. We adjusted robust standard errors for clustering at the village level.

To explore the mechanisms through which the parenting intervention may have affected children's skill development, we considered the general human capital production function of Cunha et al. (2006) and Cunha and Heckman (2007):

$$h_{t+1} = f_{t+1}(h_t, M_{t+1}, T_{t+1}, P_{t+1}, X_t, \eta_{t+1}) \tag{3}$$

where h_{t+1} and h_t are child skills at endline and at baseline, respectively; M_{t+1}, T_{t+1}, and P_{t+1} are caregiver material investments, time investments, and parenting skills during the intervention period, respectively; X_t is baseline household characteristics; and η_{t+1} represents random shocks to child skills development.

This production function indicates three possible channels[4] through which the parenting intervention might affect the child's skills. These channels are (1) changes in material investments, (2) changes in time investments, and (3) changes in parenting skills. Hence, we estimated the ITT effects of the parenting intervention on these three channels using the same specification as that for child skills (equation (2)).

Dose response. After estimating the ITT effects of the parenting intervention on child skills and parenting outcomes, we explore the effects of the program, given the variation in compliance. We estimate the dose-response relationship between the average number of center visits per month and our outcomes of interest (child skills, material investments, time investments, and parenting skills), using the control function method to account for the endogeneity of the compliance decision (Wooldridge, 2015). We use three variables as instruments: random treatment assignment, distance between the household and the village office, and the interaction of these two measures. The treatment status is a valid instrument, as it was randomly assigned and significantly explains the number of visits. Distance between the

4 Another potential channel is that the intervention could change production technology. In a study of an ECD intervention in Colombia, however, Attanasio et al. (2014) do not find support for this channel. We therefore assume this channel to be negligible and focus on the reduced-form impacts on inputs to the production function.

household and village office is assumed to be related to treatment intensity (i.e., the average number of center visits per month) but not to directly affect child skills or parental investment decisions.[5] We estimate both the linear relationship between visits and outcomes as well as the non-linear relationship by including a quadratic function of visits. Further details on the dose-response analysis, as well as the first-stage regression results, are provided in Appendix C.

3. Results

3.1 Program effects on child developmental outcomes

Table 4 presents the estimated ITT effects of the center-based program on child skills. We find that the 12-month center-based parenting intervention increased child cognitive skills by 0.11 *SD*, significant at the 10% level. The estimated treatment effects on children's language, motor, and social-emotional skills, as well as the composite latent skill factor that combines the four domains, however, are not statistically significant.

[5] Linear estimates of the dose-response relationship between number of center visits and outcomes of interest are similar when we instrument compliance with only treatment assignment.

Table 4. Intention-to-treat (ITT) effects on child's skills

Skill	Point Estimate	Standard Error	*p*-value
Cognitive skill (*n* = 1200)	0.112*	0.059	0.061
Language skill *n* = 1200)	0.011	0.059	0.850
Motor skill (*n* = 1200)	-0.047	0.070	0.506
Social-emotional skill (*n* = 1200)	-0.106	0.074	0.155
Total child skill factor (*N* = 1200)	0.028	0.061	0.652

Notes. Child's skills are all standardized by the distribution of the control group. Each row corresponds to an independent regression, and all regressions control county fixed effects and corresponding baseline skills. OLS estimates are reported, and robust standard errors, clustered at the village level, are presented in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

3.2 Mechanisms: Program effects on caregiver material investments, time investments, and parenting skills

Table 5 presents the program's ITT effects on caregiver material investments, time investments, and parenting skills. As shown in Panel A, the program had a small but positive impact on the material investments of caregivers. Among the specific components of material investment, the program significantly increased the number of play material sources and the number of picture books for children in the treatment households, but it had no effect on the total number of play materials. This is most likely because the parenting center allowed households to borrow play materials from the center to bring home, which increased sources of play materials for children, but the absolute number of play materials at home did not change when the play materials were returned to the centers. Such a finding is consistent with a systematic review of 21 parenting interventions in LMICs by Aboud and Yousafzai (2015), who found that, when households receive free play materials as part of an ECD intervention, they are less likely to invest in additional play materials for their children.

Table 5. Intention-to-treat (ITT) effects on material investments, time investments, and parenting skills

ITT Effects	Point Estimate	Std. Error	*p*-value
Panel A: Parental material investments (*n* = 1200)			
Number of play material sources	0.112**	0.052	0.033
Number of play material varieties	0.076	0.049	0.123
Number of picture books	0.128**	0.052	0.016
Number of play materials	0.027	0.056	0.628
Number of books (except picture books)	0.045	0.047	0.343
Number of magazines and newspapers	0.045	0.051	0.380
Material investment factor	0.089*	0.046	0.056
Panel B: Parental time investments (*n* =1200)			
Read or look at picture books with child	0.303***	0.061	0.000
Tell stories to child	0.214***	0.053	0.000
Sing songs with child	0.172***	0.063	0.008
Play with child with toys	0.061	0.062	0.329
Spend time with child naming things, counting, or drawing	0.047	0.051	0.355
Time investment factor	0.246***	0.062	0.000
Panel C: Parenting skills (*n* = 1200)			
Caregiver feels duty to help baby understand the world	0.007	0.047	0.876
Caregiver finds it important to play with baby	0.158***	0.057	0.007
Caregiver knows how to play with baby	0.149**	0.058	0.012
Caregiver finds it important to read stories to baby	0.166***	0.057	0.004
Caregiver knows how to read stories to baby	0.182***	0.056	0.002
Parenting skill factor	0.220***	0.055	0.000

Note. Each row corresponds to an independent regression, and all regressions control county fixed effects and corresponding baseline material investments, time investments, or parenting skills measures. All outcomes are standardized by the distribution of the control group. OLS coefficient estimates are reported, with standard errors clustered at the village level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

As also seen in Table 5, we find that the intervention produced positive effects on caregiver time investment (Panel B) and parenting skills (Panel C). In the treatment households, caregivers more actively engaged in positive parenting activities, such as reading books, telling stories, and singing songs with their children, all of which have been shown to benefit early cognitive development (Yue et al., 2017; Wang et al., 2019). Moreover, the intervention increased the skills and confidence of caregivers to engage in these activities. At endline, caregivers of children in the treatment group reported not only stronger beliefs on the importance of playing and reading with children but also being more knowledgeable as to how to play and read with their children.

3.3 Compliance and dose-response

Next, we consider program compliance. Based on administrative records from the parenting centers, treatment households completed a mean of 6.3 center visits per month during the study, which is somewhat less than two visits per week. More than half (55.62%) of treatment households completed no more than four visits per month, or an average of one visit per week. Less than one quarter (22.70%) completed at least 12 visits per month (three visits per week).

Table 6 presents the control function estimates[6] of the dose-response relationship between the number of center visits per month and our outcomes of interest (child cognitive skills, caregiver material investments, time investments, and parenting skills). In Columns (1), (3), (5), and (7), we assume a linear dose-response relationship. We find that each completed center visit per month led to an increase in child cognitive skills by 0.02 *SD* and increased caregiver material investments, time investments, and parenting skills by 0.02 *SD*, 0.04 *SD*, and 0.04 *SD*, respectively (all significant at the 1% level). The results in Columns (2), (4), (6), and (8) show that the coefficients of the squared number of center visits on all outcomes are not statistically significant at the 10% level, suggesting that the dose-response relationships are linear rather than concave.

[6] Estimates of first stage regression are reported in Table C1. *F*-tests of joint significance of all excluded instruments have *p*-values smaller than 0.001.

Table 6. Control function estimates of dose-response

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Visit	Cognitive skills		Material investments		Time investments		Parenting skills	
Visits	0.022***	0.042***	0.019***	0.024*	0.041***	0.059***	0.038***	0.042***
	(0.008)	(0.015)	(0.006)	(0.014)	(0.008)	(0.016)	(0.007)	(0.012)
Visits^2		-0.002		-0.00001		-0.001		0.001
		(0.001)		(0.001)		(0.001)		(0.001)
N	1200	1200	1200	1200	1200	1200	1200	1200

Note.(i) The control function estimates use random treatment assignment, distance from home to the village office, and interaction of both treatment and distance as instruments. Estimates of the first stage regression are reported in Table C1. F-tests of joint significance of the excluded instruments have p-values < 0.001. All regressions control county fixed effects and corresponding baseline outcomes as in the primary ATE specification. Robust standard errors, clustered at the village level, are presented in parentheses.
$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

3.4 Heterogeneous effects

Table 7 presents the heterogeneous effects on child cognitive skills of the center-based program by child age and initial levels of child cognitive skills and parental investment. The intervention had substantially larger effects on younger children (6 to 17 months at baseline) compared to older children (18 to 24 months at baseline).[7] We estimate that the program increased cognitive skills by 0.21 *SD* in the younger age group but had no effect on older toddlers.

[7] We split the sample at 18 months of age to facilitate comparison with the home-visiting program (Section 4). The results are similar when comparing effects on children 6–15 months and 16–24 months at baseline (younger group: 0.21 *SD*s, significant at the 1% level; older group: no effect; difference: 0.21 *SD*s, significant at the 5% level).

Table 7. Heterogeneous effects of the center-based intervention

Outcome: Cognitive skills at endline	(1)	(2)	(3)	(4)
Baseline characteristics	Age < 18 Months	Low Cognitive Skills	Low Material Investments	Low Time Investments
Treatment (a)	-0.090 (0.100)	0.063 (0.078)	0.215*** (0.077)	0.147* (0.077)
Baseline characteristics	-0.366*** (0.076)	-0.420*** (0.058)	-0.092 (0.059)	-0.172** (0.072)
Treatment * baseline characteristics (b)	0.296*** (0.112)	0.149 (0.118)	-0.202** (0.084)	-0.100 (0.108)
County FE	Yes	Yes	Yes	Yes
Baseline cognitive skills	Yes	--	Yes	Yes
Treatment effect on those with baseline characteristics (a + b)	0.206*** (0.068)	0.212** (0.089)	0.013 (0.067)	0.047 (0.083)
Observations	1200	1200	1200	1200

Note. OLS estimates are reported, and robust standard errors, clustered at the village level, are presented in parentheses. Low cognitive skills (1 = baseline cognitive skills lower than the median value, 0 = no), low material investments (1 = baseline material investments lower than the median value, 0 = no), low time investments (1 = baseline time investments lower than the median value, 0 = no), and low parenting skills (1 = baseline parenting skills lower than the median value, 0 = no).
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

We find weaker evidence that the parenting center program had a larger effect on children with low levels of baseline cognitive skills (Column 2). The program increased the cognitive skills of children with skills below the median of the sample at baseline by 0.21 *SD* (significant at 5%) but had no detectable effect on those with skills above the median. This difference, however, was not significant.

We find the opposite pattern for heterogeneous effects by baseline investment (Columns 3 and 4). The program had positive and significant effects on those with higher than median levels of material and time investments at baseline (0.22 *SD* and 0.15 *SD*, respectively), but it had no effect on those with low levels of investment at baseline. The difference in treatment effects is significant for material investments: We estimate that the skills of children with high baseline material investments improved by 0.20 *SD* more than did those with low initial investments (significant at the 5% level).

4. Comparison Between the Center-Based and Home-Based Programs

Despite the significant positive impacts of the center-based intervention, the magnitude of the average treatment effect on child cognitive skills (0.11 *SD*s) was estimated to be approximately half that of a previous home-visiting intervention (0.26–0.27 *SD*s) evaluated in Sylvia et al. (2018) (*p*-value for different between effects: 0.320). Similarly, the average treatment effects of the center-based parenting intervention on material investments (0.09 *SD*) and time investments (0.25 *SD*) were significantly smaller than were the effects of the home-visiting intervention on parental investment (0.85 *SD*).

We sought to determine what was driving the differences in impacts between the two interventions. The evaluations of the two interventions were similar in a number of important ways, which, while absent a head-to-head comparison as part of the same trial, allows us to narrow the scope of possible reasons for the difference in average treatment effects. First, both interventions worked with the local NHC officials to recruit trainers in the same way. Second, both interventions used the same age-appropriate, stage-based curriculum. Third, the total time of one-on-one instruction with parenting trainers was similar (~26 hours total). Fourth, both interventions were implemented in the same region and followed a similar sample selection procedure. Finally, the interventions were evaluated by the same research team, using largely the same instruments and approaches.

There are five main differences that could be large enough to drive the substantial difference in average impacts that we observe: (1) the population of children and caregivers involved in each study; (2) program duration; (3) potential for peer effects; (4) aspects of outcome measurement; and (5) method through which the interventions were delivered, specifically, through home-based or center-based parenting sessions. We examine each of these in turn.

4.1 Differences in the samples

As shown in Appendix D, the baseline characteristics of the two samples are similar in terms of the static characteristics of children and caregivers. In both programs, approximately half of the sample children were male (49% in home-based program vs. 52% in center-based program, *p*-value = 0.737); few had low birthweight (4%, *p*-value = 0.956); the majority were natural births (69% vs. 64%, *p*-value = 0.797); and the educational level of caregivers was similar (8.55 years vs. 8.15 years, *p*-value = 0.347). Although the share of households that received social security was higher in the home-based program (27% vs. 11%, *p*-value < 0.001), this difference is almost certainly due to the rise of China's poverty alleviation program, which was beginning to replace part of the social security program in the years that the center-based program was implemented (which was two years after the implementation of the home-based program).

The primary difference between the samples is the mean age of the children. By design, children in the home-based intervention were, on average, 10 months older when they were enrolled at baseline (24.45 months vs. 14.36 months, *p*-value < 0.001). This most likely also drives the five-percentage-point difference in the probability that the mother was the primary caregiver at the start of the intervention, as many mothers out-migrate to work in urban areas when their children grow into toddlerhood (Yue et al., forthcoming).

The difference in the age profiles of children in the center-based and home-based programs, however, appears unlikely to be the primary source of difference in the average impacts that we observe. The center-based program included children from age cohorts that were both younger (6–17 months) than and the same (18–30 months) as the home-based program. As shown in Table 7, Column 1, in the case of the center-based program, the intervention had larger (and significant) effects on younger children (6–17 months at baseline) than on older children (18–30 months at baseline). There were no detectable effects on the older cohort, which corresponds to the age of the children in the home-visit evaluation, and the difference in average effects among this age group in the center- and home-based interventions is larger than the difference observed in the full samples of the two studies. This difference in age profiles across the two studies, therefore, cannot be what is driving *smaller* average impacts for the center-based program.

4.2 Differences in outcome measurement

Although the constructs, or domains, of child development that were used as primary outcomes in both studies were these same, there were some differences in how these were measured. The center-based trial measured child skills at endline using the BSID-III. Version 3, however, was only recently adapted for use in China and had not been available during the home visiting intervention. The evaluation of the home-based intervention used Version 1 of the Bayley Scales of Infant Development (BSID-I) in the younger half of the sample (children below 30 months) and the Griffith Mental Development Scales (GMDS) for children above 30 months at endline.

Focusing on cognitive skills, where we find evidence of positive impacts of the center-based intervention, our results suggest that this difference also is not a significant driver of the difference in average outcomes. First, the home-visiting program had a significant positive effect of 0.26 *SD* on the BSID-I Mental Development Index (MDI) among children who were administered the BSID-I at endline, which is more directly comparable to the BSID-III, used to evaluate the impact of the center intervention, than is the GMDS. The main difference is that, although the BSID-I MDI was designed to measure both cognitive and language development, these two constructs were measured separately (as a cognitive skill index and language skill index) in the BSID-III. That we find no effect on the language skill index in the center intervention, but see effects on the cognitive skill index, suggests that, if the BSID-I were to have been used to evaluate the center-based intervention, effects on the MDI would have been smaller than the effects that we find on the BSID-III cognitive skill index, assuming that the splitting of cognitive and language items is the main difference in the measures and depending on the degree to which language items are weighted in the MDI.

4.3 Differences in program duration

Having ruled out differences in the sample populations enrolled in each study and in measurement of outcomes, we turn to differences in features of the interventions as possible sources of the different impacts. A first major difference between the two programs was the duration of the intervention. The center-based program lasted for 12 months, while the home-based intervention lasted only 6. This means that children were exposed not only for a longer time in the center-based evaluation but also to more of the curriculum content in the one-on-one sessions. The older age group that was common to both programs was exposed to the final year of the curriculum in the center-based intervention but only to the final six months in the home-based intervention. All else equal, for this to be the reason for larger effects in the home-based intervention, it would need to be the case that children in the center-based intervention benefited less (in absolute terms) than did children in the home-based intervention from the last six months of the curriculum because they had already participated for six months. This scenario is unlikely, particularly given evidence from other contexts of dynamic complementarities or that there is a higher marginal return to human capital from a given investment for children with a larger stock of initial skills (Aizer and Cunha, 2012).

4.4 Negative peer effects

Another clear difference between the interventions is that the center-based intervention may have encouraged more interaction between children and caregivers in the same village. It is possible that the parenting centers led to negative peer effects that offset positive effects of increased investments from the intervention. This could occur, for example, if peer interactions cause children to exhibit behavior problems that affected their ability to benefit from investments. Alternatively, increased investment may cause increases in externalizing behavior in some children that somehow negatively affected the cognitive development of peers. Although some negative peer effects of this type are plausible, we also believe that it is unlikely that these were large enough to explain much of the difference in average effects compared to the home-based intervention. In separate analyses reported elsewhere (Qian, 2020), the effect of the parenting centers on children in a younger cohort in the same village who did not have access to the center was found to be positive. Assuming that these positive spillovers are due to peer interactions, this result suggests that negative peer effects are unlikely in our setting.

4.5 Differences in delivery method

A final major difference between the interventions is, of course, how they were delivered. In the center-based intervention, parenting trainings were delivered in a centralized location in the village to which caregivers had to bring their children, which can be considered a more *passive* delivery method, as it relies on caregivers' choosing to bring children to the centers. In contrast, the home-based intervention was delivered directly to caregivers in their own homes, a more *active* method of delivery, as caregivers did not have to choose to travel to the training.

The most plausible way through which this difference in delivery models could affect program effectiveness is through uptake or compliance (i.e., visits to the parenting center or in-home sessions with a parenting trainer). That the center-based intervention was delivered in a passive manner, as opposed to the more active delivery of home visits, raises participation costs on households and creates more scope for selection into and out of program participation. As a result, average effects could differ between the two programs due to how individuals selected into participation and how that affected who was treated by the intervention. We explore this possibility by first comparing the patterns of participation between the interventions and then by examining how this pattern maps onto intervention effects.

Differences in compliance could operate in terms of both the overall average level of compliance and the composition of those participating in each intervention. In terms of average participation, treatment households in the center intervention completed a mean of 6.3 center visits per month during the study, more than participants in the home-visiting intervention would have received under full compliance (1 visit per week). For one-to-one sessions, average participation in the center-based intervention was similar to that of the home-based intervention.

The composition of those taking up the intervention, however, differed substantially between the two programs. Tables 8 and 9 present the correlation between participation and baseline cognitive skills/parenting outcomes in the center-based and home-based programs, respectively. In the center-based program (Table 8), caregivers of children with low baseline cognitive skills were less likely to visit the center, while baseline investments and parenting skills were not significantly correlated with compliance. In contrast, in the home-based program (Table 9), children with lower baseline investments received a greater number of home visits, while baseline cognitive skills and parenting skills were not significantly correlated with program compliance. These correlations show that children with low initial cognitive skills tended to select out of the center-based program, whereas the home-visiting program tended to select out children who were already receiving a relatively high level of investment by caregivers.

Table 8. Correlation between program compliance and baseline child cognitive skills/parental outcomes in center-based intervention

Average visit times per month	(1)	(2)	(3)	(4)
Cognitive skills	0.486*			0.472*
	(0.279)			(0.278)
Material investment		0.157		0.090
		(0.435)		(0.436)
Time investment		0.077		0.050
		(0.288)		(0.281)
Parenting skills			0.199	0.128
			(0.348)	(0.337)
Baseline characteristics	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	792	792	792	792

Note. OLS estimates are reported, and robust standard errors are presented in parentheses, clustered at the village level.
*p < 0.10, **p < 0.05, ***p < 0.01

Table 9. Correlation between program compliance and baseline child cognitive skills/parental outcomes in home-based intervention

Average visit times per month	(1)	(2)	(3)	(4)
Cognitive skills	-0.031			-0.013
	(0.085)			(0.086)
Investment		-0.126*		-0.130*
		(0.066)		(0.066)
Parenting skills			0.010	0.032
			(0.082)	(0.080)
Baseline characteristics	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	210	210	210	210

Note. OLS estimates are reported, and robust standard errors, clustered at the village level, are presented in parentheses.
*p < 0.10, **p < 0.05, ***p < 0.01

Whether this difference in the pattern of compliance affected the average impacts of each program, however, depends on how different children were affected by participation in each intervention. More specifically, it depends on how the technology of skill formation that relates increased investment to skills varies by baseline characteristics. Table 10 presents the heterogeneous treatment effects by baseline characteristics in the home-based intervention, analogous to the results presented for the center-based intervention in Table 7. Similar to what was seen in the center-based program, children with low baseline cognitive skills improved significantly more than did those with high baseline skills. The estimated effect on low-ability children was 0.37 *SD* due to the home-visiting program, significant at the 1% level, whereas there was no effect on those with high baseline skills. The impact on children with low initial skill levels is slightly larger in magnitude to what we find for the center-based intervention.

Table 10. Heterogeneous effects of home-based intervention

Outcome: Cognitive skills at endline	(1)	(2)
Baseline characteristics	Low Cognitive Skills	Low Investments
Treatment	0.052 (0.097)	-0.006 (0.101)
Baseline characteristics	-0.745*** (0.112)	-0.145 (0.122)
Treatment * baseline characteristics	0.317** (0.159)	0.416** (0.169)
County FE	Yes	Yes
Baseline cognitive skills	--	Yes
Treatment effect on those with baseline characteristics (a + b)	0.369*** (0.129)	0.410*** (0.128)
Observations	503	473

Note. OLS estimates are reported, and robust standard errors, clustered at the village level, are presented in parentheses. Low cognitive skills (1 = baseline cognitive skills lower than the median value, 0 = no), low investments (1 = baseline investments lower than the median value, 0 = no), and low parenting skills (1 = baseline parenting skills lower than the median value, 0 = no).
**p* < 0.10, ***p* < 0.05, ****p* < 0.01

The results for the heterogeneous impacts by baseline cognitive skills, presented above in Tables 7 (Column 2) and 10 (Column 1), can be summarized graphically. Figure 3 shows that the treatment effects of the two programs across the entire baseline distribution of cognitive skills, from developmentally delayed on the left-hand side of the graph to developmentally normal on the right-hand side. The red and blue lines plot the non-parametric estimates of the treatment-control difference in outcomes for the home-based and center-based programs, respectively. Both are downward sloping, indicating that both provided the greatest benefit to those children with the greatest need (that is, those with the lowest levels of cognitive development at baseline). However, for children at the low end of the baseline skills distribution, the effects of home-based intervention are significantly higher than those of the center-based program.

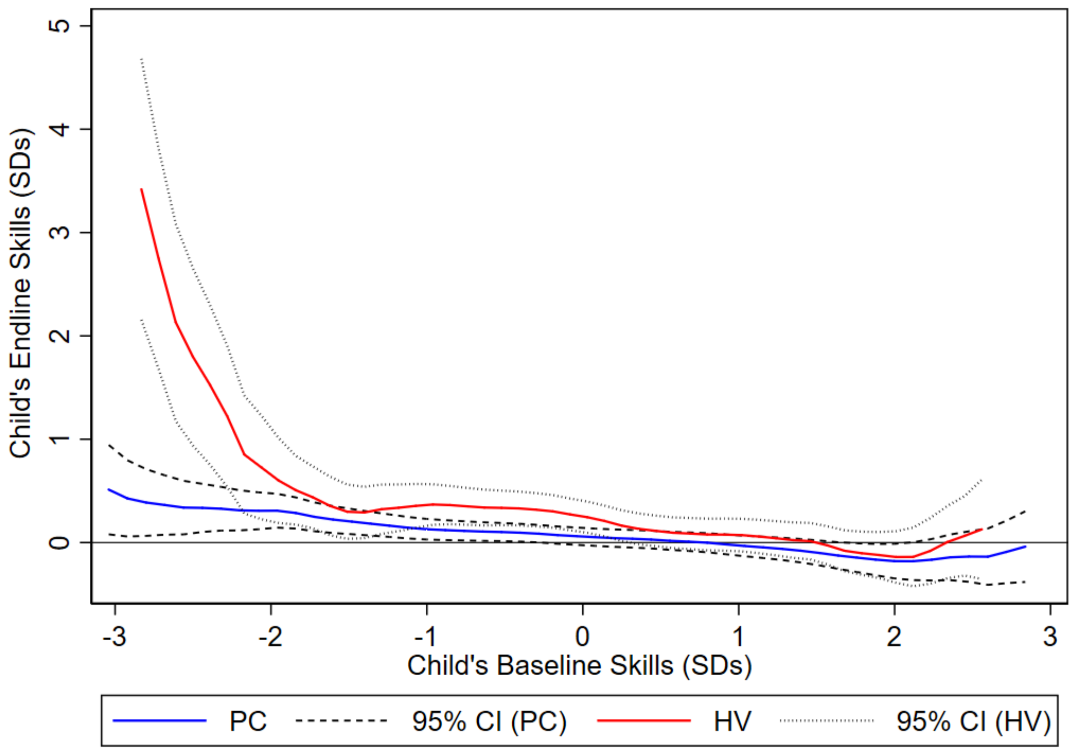

Figure 3. Nonparametric estimates of treatment effects of the parenting center program and the home-visiting program on child's endline skills with respect to baseline skills.

Note. Local polynomial regressions of the endline skill factor score on baseline skill factor score are performed across treatment arms. The Epanechnikov kernel function is used in the regression. The treatment effect of parenting center program (PC) on child's endline skills is plotted in the blue solid line, with the 95% confidence interval (black dashed line), and the treatment effect of home visit program (HV) on child's endline skills is plotted in the red solid line, with the 95% confidence interval (black dotted line).

The pattern of heterogeneous effects differs even more in terms of initial investments. The effects by baseline investments are, in fact, opposite in the center-based and home-based programs: The effects of the home-based program on children with low baseline investments are 0.42 *SD* higher than children with high baseline investments, significant at the 5% level (Table 10, Column 2). This is in contrast to the center-based intervention, which, as discussed, had significantly larger impacts on children with higher levels of initial investment (Table 7, Columns 3 and 4).

In summary, we find that, in the center-based intervention, caregivers of children with low levels of baseline skills tended to select out of participation, whereas they did not in the home-based intervention. In the home-based intervention, those who selected out tended to be caregivers with higher levels of initial investment. Because children with lower levels of initial skills and whose caregivers are investing less tend to benefit more from parenting interventions, this is a likely cause for lower average impacts of the center-based program.

5. Conclusion

Little is known about the relative effectiveness of different delivery models for ECD interventions in LMICs. We evaluated the effects of a free center-based parenting intervention on ECD outcomes and parenting practices, using data from a randomized trial across 100 villages in rural northwestern China. We also compared the effects of the center-based intervention with those of a home-based intervention conducted by the same research team in the same region and using the same parenting curriculum and public service system as the center-based intervention (Sylvia et al, 2018).

We find that the center-based parenting intervention produced positive and significant impacts on caregiver material investments, time investments, and parenting skills, leading to improvements in children's cognitive skills after 12 months. The center-based program also had greater impacts on children with low baseline cognitive skills, indicating that the intervention provided the greatest benefit to children with the greatest need for ECD support. The effects of the center-based intervention on both child skills and parenting outcomes, however, are significantly lower than those of the home-based intervention. According to our analysis, this is likely due to the selection of families into the programs in terms of compliance: In the center-based program, caregivers of children with low baseline cognitive skills tended to participate at a lower rate, whereas, in the home-based program, caregivers who were already investing in their children at higher levels were the ones who tended to not participate. Our findings suggest that greater compliance by families with more vulnerable children in the home-based program contributed to the larger impacts (on average) than the center-based program. Thus, at least as implemented, it seems clear that the home-based program was more effective in raising the cognitive skills of the children who presumably are the ones who need the most attention.

Beyond differences in effectiveness, there are also clear differences in the cost associated with providing parental training to children in the study areas. The costs of the home-based intervention include mainly the salaries of the parenting trainers (and monitors of the trainers) and the supply of toys and books (both originals and replacements). The costs of the center-based intervention included these same costs—the salaries of parenting trainers and the supply of toys and books—and because the number of trainers, toys, and books per village were the same in both the home-based and center-based programs, these elements had nearly identical costs. In the case of the parenting centers, however, there were additional costs, including the salary of the day-to-day center manager and the cost of the utilities of the center, including the cost to turn the center into a child-safe play space. Full cost accounting also would include the rental fee for the center building/space (although in this particular intervention, villages provided it for free). In this way, then, it is obvious that the center-based intervention was more expensive than was the home-based one. Given the results from the above experiment, it is also clear that, in the case of rural China, the home-based intervention was unambiguously more cost effective.

This is the first study to compare two popular delivery models for ECD interventions, using the same research team, program, and methods. The results provide insight into the effects of ECD interventions delivered through existing public resources in developing countries. Our findings also shed light on the underlying mechanisms for the differences in intervention effects, that is, differences in the patterns of compliance between the two delivery models.

We also acknowledge two main limitations to this study. First, because the experiment was conducted in only one underdeveloped rural area in Western China, the results may not be generalizable to all LMICs. In addition, this study compared only the short-term treatment effects of the center-based and home-based programs. Future research should follow up on the children in both studies to investigate differences in treatment effects of the two programs over time.

Despite these limitations, the results have important implications. Our findings suggest that program participation is critical to overall program effectiveness. Due to selection of participation in both programs, the home-based delivery model may be more effective, and cost effective, than is the center-based delivery model. Alternatively, increasing program participation of vulnerable children and their caregivers would substantially increase the benefits of the center-based intervention, which could play an active role in shaping early cognitive skills in poor rural areas. These results also may imply that, for other services for which there is an option for more passive delivery (as with parenting centers) or a more active form of delivery (as with home visits), relative participation in the different models should be taken into account.

References

Abdul Latif Jameel Poverty Action Lab. (2019). *Encouraging early childhood stimulation from parents and caregivers to improve child development*. J-PAL Policy Insights. Unpublished manuscript.

Aboud, F. E., & Yousafzai, A. K. (2015). Global health and development in early childhood. *Annual Review of Psychology*, *66*, 433–457.

Aizer, A., & Cunha, F. (2012). *The production of human capital: Endowments, investments and fertility* (No. w18429). New York, NY: National Bureau of Economic Research.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., & Rubio-Codina, M. (2014). Using the infrastructures of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: Cluster randomized controlled trial. *BMJ*, *349*, g5785.

Bai, Y., Zhang, S., Wang, L., Dang, R., Abbey, C., & Rozelle, S. (2019). Past successes and future challenges in rural China's human capital. *Journal of Contemporary Chin*a, *28*(120), 1–16.

Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology*, *28*(1), 97–104.

Bayley, N. (2006). *Bayley scales of infant and toddler development: Administration manual*. San Antonio, TX: Harcourt Assessment.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., . . . Devercelli, A. E. (2017). Early childhood development coming of age: Science through the life course. *Lancet*, *389*(10064), 77–90.

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, *1*(4), 200–232.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investment substantially boost adult health. *Science*, *343*(6178), 1478–1485.

Cattell, R. G. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276.

Chan, M. (2013). Linking child survival and child development for health, equity, and sustainable development. *Lancet*, *381*(9877), 1514–1515.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., & Heckman, J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, *43*(4), 738–782.

Cunha, F., & Heckman, J. J. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, *7*(2–3), 320–364.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1, 697–812.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation, *Econometrica*, *78*(3), 883–931.

Frongillo E. A., Sywulka S. M., & Kariger P. (2003). *UNICEF psychosocial care indicators project. Final report to UNICEF*. Ithaca, NY: Division of Nutritional Sciences, Cornell University.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., . . . Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, *344*(6187), 998–1001.

Hamadani, J. D., Tofail, F., Hilaly, A., Huda, S. N., Engle, P., & Grantham-McGregor, S. M. (2010). Use of family care indicators and their relationship with child development in Bangladesh. *Journal of Health, Population, and Nutrition*, *28*(1), 23–33.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, *104*(33), 13250–13255.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. Q. (2010). The rate of return to the HighScope Perry Preschool program, *Journal of Public Economics*, *94*(1–2), 114–128.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, *6*(1), 689–733.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Huggett, M., Ventura, G., & Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, *101*(7), 2923–2954.

Li, H., Loyalka, P., Rozelle, S., & Wu, B. (2017). Human capital and China's future growth. *Journal of Economic Perspectives*, *31*(1), 25–48.

Luo, R., Jia, F., Yue, A., Zhang, L., Lyu, Q., Shi, Y., . . . Rozelle, S. (2017). Passive parenting and its association with early child development. *Early Child Development and Care*, 1–15. https://doi.org/10.1080/03004430.2017.1407318

National Bureau of Statistics of China (2016). *China Statistical Yearbook 2006*. Beijing: China Statistics Press.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Qian, Y. (2020). *Indirect effect of an early childhood intervention and peer effects in parenting practice* (REAP working paper). Stanford University, Stanford, CA.

Richter, L. M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F. L., Behrman, J. R., . . . Darmstadt, G. L. (2017). Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet*, *389*(10064),103–118.

The State Council of the People’s Republic of China. (2006, November 19). *China Rural Poverty Alleviation and Development Outline (中国农村扶贫开发概要)*. Retrieved from http://www.gov.cn/zwhd/ft2/20061117/content_447141.htm

Sylvia, S., Warrinnier, N., Renfu, L., Yue, A., Attanasio, O. P., Medina, A., & Rozelle, S. (2018). *From quantity to quality: Delivering a home-based parenting intervention through China's family planning cadres* (LICOS Discussion Paper No. 402).

Walker, S. P., Chang, S. M., Vera-Hernández, M., & Grantham-McGregor, S. (2011). Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics*, *127*(5), 849–857.

Wang, B., & Yue, A. (2019). *The relationship between family environment and early child development in rural China* (REAP working paper). Stanford University, Stanford, CA.

Wang, L., & Zheng L. (2019). *Family care and early childhood development: evidence from china’s four rural subpopulations* (REAP working paper). Stanford University, Stanford, CA.

Wang, L., Liang, W., Zhang, S., Jonsson, L., Li, M., Yu, C., . . . Luo, R. (2019). Are infant/toddler developmental delays a problem across rural China? *Journal of Comparative Economics*, *47*(2), 458–469.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420–445.

Wu, K. B., Young, M. E., & Cai, J. (2012). *Early child development in China: Breaking the cycle of poverty and improving future competitiveness*. Washington, DC: The World Bank.

Yue, A., Bai, Y., Shi, Y., Luo, R., Medina, A., & Sylvia, S. (forthcoming). Parental migration and early childhood development in rural China. *Demography*.

Yue, A., Shi, Y., Luo, R., Chen, J., Garth, J., Zhang, J., . . . Rozelle, S. (2017). China’s invisible crisis: Cognitive delays among rural toddlers and the absence of modern parenting. *The China Journal*, *78*(1), 50–80.

Appendix A. Measurement System

In this study, we construct latent factor measures for child skills (a composite of the BSID-III cognitive, language, motor and social-emotional subscales) as well as caregiver material investments, time investments, and parenting skills. To preliminarily identify relevant measures of latent factors for the measurement system, we used exploratory factor analysis (EFA). We used Cattell's (1966) scree plot and Horn's (1965) parallel analysis, both of which are reported in Table B1. Both analyses show that one latent factor should be extracted from the measures for child skills, material investments, time investments, and parenting skills and at baseline.

Table A1. EFA to Determine the number of latent factors

	Cattell's scree plot	Horn's parallel analysis
Material investments at baseline	1	1
Time investments at baseline	1	1
Parenting skills at baseline	1	1
Child skills at baseline	1	1

Table A2 reports EFA-estimated rotated factor loadings on measures of child skills, material investments, time investments, and parenting skills at baseline. We find that all measures at baseline load strongly on the first factor.

Table A2. Estimated rotated factor loadings by EFA on child skills, material investments, time investments, and parenting skills at baseline

Model	First Factor
One-factor model of child skills	
Cognitive skills	0.612
Language skills	0.614
Motor skills	0.626
Social-emotional skills	0.210
One-factor model of material investments	
Number of play material sources	0.539
Number of play material varieties	0.775
Number of picture books	0.601
Number of play materials	0.714
Number of books (except picture books)	0.599
Number of magazines and newspapers	0.517
One-factor model of time investments	
Read books or looked at picture books with child in last 3 days	0.560
Told stories to child in last 3 days	0.614
Sang songs with child in last 3 days	0.586
Played with the child with toys in last 3 days	0.457
Spent time with child in naming things, counting, or drawing in last 3 days	0.489
One-factor model of parenting skills	
Caregiver feels duty to help baby understand the world	0.654
Caregiver finds it important to play with baby	0.722
Caregiver knows how to play with baby	0.630
Caregiver finds it important to read stories to baby	0.643
Caregiver knows how to read stories to baby	0.564

Table A3 shows the measurement system for the child skill factor, and Table B4 reports the measurement system for the latent factors of caregiver material investments, time investments, and parenting skills. The first column of each table contains the factor loadings. We normalized the factor loading of the first measure to equal 1 in both rounds, and, thus, this measure defines the scale of the latent factor. Following Attanasio et al. (2014) and Sylvia et al. (2018), we also calculated the signal-to-noise ratio to measure the percentage of each measure's variance that is driven by signal, which is reported in the second column. This assesses the amount of information contained in each measure. For example, we calculated this ratio for the k-th measure of the parenting skill factor as the following specification:

$$S_k^P = \frac{\alpha_k^2 Var(P)}{\alpha_k^2 Var(P) + Var(\delta_k)}$$

where the k-th measure of parenting skill factor is written as the following simplified notation:

$$m_k^P = \mu_k + \alpha_k P + \delta_k$$

We find that the information contained in different measures can vary significantly, even for the same latent factor, and signal cannot account for 100% of the variance for most measures. This indicates that the latent factor approach is useful to reduce measurement errors in modeling child skills and parenting outcomes.

Table A3. Measurement system for child skills

Latent Factor	Measurement	Factor loading	% Signal
Child skills at baseline			
	Cognitive skills	1	43%
	Language skills	1.09	43%
	Motor skills	1.38	45%
	Social-emotional skills	0.37	4%
Child skills at endline			
	Cognitive skills	1	51%
	Language skills	1.24	60%
	Motor skills	1.13	40%
	Social-emotional skills	0.23	2%

Table A4. Measurement system for material investments, time investments, and parenting skills

Latent Factor	Measurement	Factor loading	% Signal
Material investments: B			
	Number of play material sources	1	29%
	Number of play material varieties	2.72	66%
	Number of picture books	1.15	39%
	Number of play materials	1.24	52%
	Number of books (except picture books)	1.32	33%
	Number of magazines and newspapers	0.89	25%
Time investments: B			
	Read books or looked at picture books with child in last 3 days	1	35%
	Told stories to child in last 3 days	1.07	44%
	Sang songs with child in last 3 days	1.29	37%
	Played with the child with toys in last 3 days	0.87	19%
	Spent time with child in naming things, counting, or drawing in last 3 days	1.01	23%
Parenting skills: B			
	Caregiver feels duty to help baby understand the world	1	47%
	Caregiver finds it important to play with baby	1.05	58%
	Caregiver knows how to play with baby	1.07	40%
	Caregiver finds it important to read stories to baby	0.97	42%
	Caregiver knows how to read stories to baby	1.04	29%
Material investments: E			
	Number of play material sources	1	15%
	Number of play material varieties	4.07	79%
	Number of picture books	2.06	43%
	Number of play materials	0.87	28%
	Number of books (except picture books)	1.41	19%
	Number of magazines and newspapers	0.99	14%
Time investments: E			
	Read books or looked at picture books with child in last 3 days	1	43%
	Told stories to child in last 3 days	1.07	50%
	Sang songs with child in last 3 days	1.07	46%
	Played with the child with toys in last 3 days	0.81	28%
	Spent time with child in naming things, counting, or drawing in last 3 days	0.69	20%
Parenting skills: E			
	Caregiver feels duty to help baby understand the world	1	16%
	Caregiver finds it important to play with baby	1.69	33%
	Caregiver knows how to play with baby	2.45	39%
	Caregiver finds it important to read stories to baby	2.12	37%
	Caregiver knows how to read stories to baby	2.86	39%

Note. B = Baseline; E = Endline

Appendix B: Data and Sample in the Center-Based Intervention

Table B1. Family Care Indicators (FCI) Scale

	Cronbach's Alpha	
Subscale	Baseline	Endline
Source of play materials		
Homemade toys	0.81	0.81
Household objects	0.81	0.82
Things from outside	0.81	0.82
Toys bought from store	0.81	0.82
Varieties of play materials		
Things that make/play music	0.80	0.81
Things for drawing/writing	0.80	0.81
Picture books for children (excluding schoolbooks)	0.80	0.81
Things meant for stacking, constructing, building (blocks)	0.80	0.81
Things for moving around (e.g., balls, bats)	0.80	0.82
Toys for learning shapes and colors.	0.80	0.81
Things for pretending (e.g., dolls, tea set)	0.81	0.82
Play activities		
Read books or looked at picture books with child in last 3 days	0.80	0.81
Told stories to child in last 3 days	0.80	0.81
Sang songs with child in last 3 days	0.80	0.81
Played with the child with toys in last 3 days	0.80	0.81
Spent time with child in naming things, counting, or drawing in last 3 days	0.80	0.81
Household books		
Number of books for adults in the home	0.80	0.81
Picture books		
Number of picture books for children in the home	0.80	0.80
Play materials		
Number of play materials in the home	0.80	0.81
Magazines		
Number of magazines and newspapers in the home	0.80	0.81
Total	0.81	0.82

Note. The items in the former three subscales (Source of play materials, Varieties of play materials, and Play activities) are scored as yes = 1 and no = 0 (Presence or absence of play material or activity). The four other items (Household books, Picture books, Play materials, Magazines) are scored in terms of the quantity.

Table B2. Analysis of sample attrition

	(1)	(2)	(3)	(4)
Attrition	Full sample	Full sample	Control	Treatment
Treatment	-0.062	-0.067		
	(0.054)	(0.052)		
Male		0.038**	0.027	0.047
		(0.019)	(0.026)	(0.028)
Age in months		-0.001	0.001	-0.003
		(0.002)	(0.003)	(0.003)
Low birth weight		0.049	0.093	-0.011
		(0.072)	(0.112)	(0.080)
Natural birth		-0.033	-0.053	-0.017
		(0.024)	(0.037)	(0.030)
Premature		-0.042	-0.096	0.017
		(0.059)	(0.076)	(0.086)
Caregiver’s age		0.004**	0.007***	-0.0002
		(0.002)	(0.002)	(0.003)
Caregiver’s year of schooling		0.008**	0.011*	0.006
		(0.004)	(0.007)	(0.005)
Mother is the primary caregiver		0.078*	0.129**	0.002
		(0.043)	(0.051)	(0.075)
Child who has older siblings		-0.081***	-0.115***	-0.043
		(0.025)	(0.037)	(0.033)
Welfare household		0.036	-0.010	0.075
		(0.051)	(0.059)	(0.078)

Note. OLS estimates are reported in the table; robust standard errors are presented in parentheses, clustered at the village level. The p-value of the Chow test in Columns 3 and 4 is 0.36, which cannot reject that the correlates of attrition are similar in the control (Column 3) and treatment groups (Column 4).

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Appendix C. Dose-Response

We estimate the dose-response relationship between the average number of center visits per month and our outcomes of interest (child skills, material investments, time investments, and parenting skills), using the control function method (Wooldridge, 2015). We estimate the first stage equation for each of the outcome of interest using the following OLS specification:

$$V_{ijt} = \alpha_3 + \beta_3 D_{ij} + \beta_4 D_{ij} \times d_{ij} + \beta_5 d_{ij} + \gamma_3 Y_{ij(t-1)} + \tau_s + \varepsilon_{ij} \quad \text{(C.1)}$$

where V_{ijt} is the average number of center visits per month during the program, D_{ij} is a dummy that indicates treatment assignment, d_{ij} is the distance between household and village office, $Y_{ij(t-1)}$ is the outcome of interest at baseline, and τ_s is the county fixed effects. Table C1 presents the results of the first-stage estimation.

Table C1. First stage of dose-response estimation

	(1)	(2)	(3)	(4)
Excluded Instruments				
Treatment	8.064***	8.088***	8.080***	8.101***
	(0.669)	(0.668)	(0.668)	(0.675)
Distance to the village office	0.001	-0.004	-0.001	-0.003
	(0.039)	(0.041)	(0.039)	(0.038)
Treatment * Distance to the village office	-1.645***	-1.655***	-1.652***	-1.654***
	(0.349)	(0.354)	(0.352)	(0.352)
Lagged Outcome Variables				
Cognitive skills	0.210			
	(0.161)			
Material investments		-0.030		
		(0.232)		
Time investments			0.081	
			(0.147)	
Parenting skills				0.226
				(0.196)
N	1200	1200	1200	1200
F-statistic	95.78	105.42	93.03	85.60

Note. Robust standard errors are clustered at the village level and presented in parentheses.
$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Using the estimated residual in the first-stage equation, $\hat{\varepsilon}_{ij}$, we then estimate the second-stage equation as follows:

$$Y_{ijt} = \alpha_4 + \beta_6 V_{ijt} + \beta_7 \hat{\varepsilon}_{ij} + \gamma_4 Y_{ij(t-1)} + \tau_s + \xi_{ij} \quad \text{(C.2)}$$

Following the approach of Sylvia et al. (2018), we also test whether a concave dose-relationship relationship exists, using the squared number of center visits per month, V_{ijt}^2. In this case, we estimated the second-stage equation as the following specification:

$$Y_{ijt} = \alpha_5 + \beta_8 V_{ijt} + \beta_9 V_{ijt}^2 + \beta_{10} \hat{\varepsilon}_{ij} + \beta_{11} \hat{\varepsilon}_{ij}^2 + \gamma_5 Y_{ij(t-1)} + \tau_s + \xi_{ij} \quad \text{(C.3)}$$

where $\hat{\varepsilon}_{ij}$ is the estimated residual in the first-stage equation, and $\hat{\varepsilon}_{ij}^2$ is the squared residual. We adjust standard errors for clustering at the village level.

Appendix D. Comparison of Samples of the Center-Based and Home-Based Interventions

Table D1. Descriptive statistics of baseline characteristics in center-based and home-based program

Variable	Center-based program	Home-based program	*p*-value
Male (yes = 1)	0.52 (0.01)	0.49 (0.02)	0.737
Age in months	14.36 (0.15)	24.45 (0.14)	0.000
Low birth weight (yes = 1)	0.04 (0.01)	0.04 (0.01)	0.956
Natural birth (yes = 1)	0.64 (0.02)	0.69 (0.02)	0.797
Caregiver's completed years of schooling	8.15 (0.11)	8.55 (0.12)	0.347
Mother is the primary caregiver (yes = 1)	0.70 (0.01)	0.65 (0.03)	0.001
Household receives social security support (yes = 1)	0.11 (0.01)	0.27 (0.02)	0.000

Note. The statistics are the sample mean, with the standard error presented in parentheses. The *p*-values account for clustering at the village level.