Report No: AUS0000253

Evaluating Integration in the ICDS:

Impact Evaluation of an AWC-cum-creche pilot in Madhya Pradesh

September 2018

Health, Nutrition, Population Global Practice and the Strategic Impact Evaluation Fund

Project Team:

Principal Investigators: Harold Alderman, Senior Research Fellow, International Food Policy Research Institute and Jed Friedman, Senior Economist, World Bank

Task Team Leader and Technical specialists: Mohini Kak, Health Specialist, World Bank and Ramesh Govindaraj, Lead Health Specialist, World Bank

Data Collection: Oxford Policy Management Limited

Table of Contents

List of Tables and Figures

Abbreviations and Acronyms

ASQ	Ages and Stages Questionnaire
AWC	Anganwadi Centre
AWCC	Anganwadi Centre cum Creche
AWH	Anganwadi Helper
AWW	Anganwadi Workers
BSID	Bayley's Scales of Infant Development
CREDI	Caregiver Reported Early Development Index
DWCD	Department of Women and Child Development
ECCE	Early Childhood Care and Education
ECE	Early Childhood Education
ECD	Early Childhood Development
GoI	Government of India
GoMP	Government of Madhya Pradesh
HAZ	Height for age z scores
ICDS	Integrated Child Development Services
ICC	Intraclass correlation coefficient
IFPRI	International Food Policy Research Institute
ITT	Intent to Treat
IQ	Intelligence Quotient
MC	Mobile Creches
MWCD	Ministry of Women and Child Development
OLS	Ordinary least squares
SC	Scheduled Caste
SD	Standard Deviation
SES	Socio economic status
SIEF	Strategic Impact Evaluation Fund
ST	Scheduled Tribe
WAZ	Weight for age z scores
WHZ	Weight for Height z scores

Executive Summary

India is home to approximately 164 million children 0-6 years of age, constituting about 20 percent of the world's child population in this age group. Their optimal development is critical to human capital development in the world. Yet, with 38.4 percent children under 5 years of age in India stunted, 21 percent wasted and only 9.6 percent children aged 6-23 months receiving an adequate diet, optimal child development in India is still a goal for the future. Indicators on cognitive development, while currently not available at the national level, are also likely to exhibit a similar status, as conditions for adequate cognitive development of the child are similar to those required for adequate health and nutrition.

Investments in early childhood development (ECD) are thus critical. They are known to positively impact school attendance, retention, learning and lifetime earnings. Recognizing the relevance of strengthening their investments in ECD, India in 2012 restructured their Integrated Child Development Services (ICDS) program framework to focus more on children under 3 years of age. One element of this strengthening was a proposal to introduce on an experimental basis, the provision of day care crèches for care and development of children 6-72 months of age whose mothers go to work, in 5% of existing Anganwadi Centres (AWCs).

It was in this context that the Department of Women and Child Development (DWCD), Government of Madhya Pradesh (GoMP), initiated an AWC-cum-creche pilot in 100 centres across 2 districts (Dhar and Singrauli) in the State. With the objective of supporting the government develop a strong evidence base for ECD in the country, the World Bank collaborated with the DWCD, MP to build a rigorous impact evaluation around the pilot. Supported by the Strategic Impact Evaluation Fund (SIEF), the evaluation was designed as a randomized control trial to study the impact of an integrated package of nutrition and early child stimulation delivered by an additional worker (a creche worker) at the AWC on nutrition and cognitive development outcomes in children under 3 years of age.

The model involved expanding the scope of the existing services provided at the AWC under the ICDS scheme to include crèche services for children under 3 years of age. The primary objective of adding creche services to the existing nutrition and pre-school education services provided at the AWC was to strengthen the focus on early childhood care and stimulation among children under 3 years of age. In other words, the model sought to enhance child care by including child stimulation along with a package of child feeding and health care being delivered through the Government's ICDS program.

The pilot or the AWC-cum-creche model was implemented entirely by and through Government systems, with some technical assistance from IFPRI and the World Bank in the form of capacity building of the newly recruited creche workers. Being completely government managed and owned makes the model not only potentially scalable but also provides a clear understanding of the barriers, limitations and opportunities if it is adopted and scaled up.

In addition to the primary outcomes of nutrition and cognitive development in children two to three years of age, the evaluation also explored an auxiliary outcome of change in labor force participation and time allocation facilitated by the availability of creche services provided through the AWC-cum-creche program. This was considered important to explore as the inability of a mother in poor households to work and thus contribute to household income could lead to a worsening of a child's well-being.

The evaluation further explored the emergent issue of maternal depression and its association with ECD outcomes. Maternal depression is known to be a risk factor for undernutrition as well as for delayed cognitive development in many low and middle-income countries and is an issue that has not received due attention in India. Additionally, given the challenge of implementing developmental assessments for

infants and small children through large household surveys, the team took the opportunity at the endline survey to contribute to the validation of an easy to implement instrument - the Caregiver Reported Early Development Instrument (CREDI) - for measuring the developmental status of children under 3 years of age at the population level.

While the Bayley's Scales of Infant and Toddler Development (BSID) are considered the most reliable measure of early child development in the ages up to 42 months, due to the rigor, cost and time required to implement them, they are often not practical to implement through large scale population based surveys. The team thus used CREDI to measure cognitive development of children in the endline survey, while BSID was administered to a sub-sample for validation. The inclusion of the validation exercise has the potential to contribute significantly to future policy and program discussions on assessments of cognitive development outcomes at scale.

The baseline survey was carried out from September-December 2014, while end-line survey was completed from January-February 2018. Both the baseline and endline surveys collected data on anthropometric and cognitive development outcomes on the index children, information on service provision from the Anganwadi worker (at baseline and endline) and creche worker (only at endline), household characteristics, knowledge and practice of child care as well as depression among primary caregivers in both treatment and control arms of the study.

The baseline survey reached 2934 primary caregivers and index children between the ages of 6 to 48 months while the endline survey was completed with 2856 primary caregivers and index children between the ages of 18 and 42 months. A sub-set of 1037 index children falling within the age range of 24-36 months were additionally administered the Bayley-III at the endline.

The results show, contrary to initial hopes, that the impact of the AWC-cum-creche has been modest. There is a small increase in the receipt of early childhood education services in treatment communities, on the order of 8.2 percentage points, or approximately an increase in service utilization for 1 out of every 12 children in the eligible age range. There is no observed increase in the physical development of the children in treatment communities (in terms of HAZ, WAZ, or WHZ) and no increase in assessed cognitive ability. There is, however, an increase in the socio-emotional development levels of children in these communities, and a slight but precisely estimated increase in the labor force participation of caregivers, on the order of a 4-percentage point increase.

In a sense the results reflect the quality of implementation of the intervention. There was minimal investment by the government, mainly in terms of creating a platform where smaller children could come; the training of workers was fairly basic. Yet, with this minimal investment a small positive change was seen in both labor force participation and in socio-emotional development. This points to potential possibilities if the government had addressed the myriad issues that persisted through the course of implementation and serves as lessons that could contribute to the National Creches Scheme for children of working mothers that is being formulated by the government.

Additionally, while the CREDI validation exercise did not provide clear answers or resolve which of a set of instruments designed to utilize caregiver responses is the most suitable, it showed that caregivers can provide assessments in keeping with the more interactive (and, thus, more time consuming and training intensive) BSID approach. This is important as it points to alternate feasible tools to measure cognitive development outcomes through large scale surveys.

I. Introduction

1. The first few years of a child's life largely determine her/his growth, learning and holistic development. With 38.4 percent children under 5 years of age in India stunted, 21 percent wasted and only 9.6 percent children aged 6-23 months receiving an adequate diet, optimal child development in India is a goal still for the future[1]. While data to assess performance along key developmental domains - cognitive, motor and language – are not available, health and nutrition indicators are suggestive of poor development along cognitive dimensions as conditions for adequate cognitive development of the child are similar to those required for adequate health and nutrition, namely appropriate child care practices.

2. With India being home to approximately 164 million children 0-6 years of age, constituting about 20 percent of the world's child population in this age group, investing in child development in India is critical to human capital development in the world. Early childhood development (ECD) is known to positively impact school attendance, retention, learning and lifetime earnings. Early nutrition programs can increase school completion by one year and can raise adult wages by 5-50 percent. Furthermore, children who escape stunting are 33% more likely to escape poverty as adults. Evidence from both developed and developing countries suggests a potential annual rate of return of 7–16 percent from high-quality ECD and nutrition programs targeting vulnerable groups, primarily due to positive impacts in education and health, reduced risky behaviors in adolescence, and increased productivity in adulthood. At a national level, reduction in stunting has the potential to increase GDP by 4-11 percent, especially in Asia and Africa[2].

3. India's investments in ECD have mainly been through the Integrated Child Development Services Scheme (ICDS). The ICDS is one of the world's largest and most unique program for early childhood development. It aims at improving the nutritional and health status of children in the age-group 0-6 years, laying the foundation for proper psychological, physical and social development of the child and breaking the vicious cycle of malnutrition, morbidity, reduced learning capacity and mortality.

4. The ICDS began as a pilot serving 33 development blocks in 1975 and today serves approximately 60 million children and 13 million mothers, roughly 40% of Indian children 0-6 years of age. It was designed as an integrated scheme providing health, nutrition and pre- school education services from a single point of service in the community, the Anganwadi Center (AWC), managed by an Anganwadi worker (AWW) and an Anganwadi helper (AWH).

5. While designed well, the ICDS suffered from several operational limitations[3]. One of its key limitations has been its inability to reach out to children under 3 years of age, which is the critical age for maximizing the impact of both nutrition and early child stimulation interventions. There also has been an over-emphasis on food supplementation relative to the other critical aspects required in the provision of an effective nutrition package. Capacity of the workers to deliver the full complement of services and of supervising officials to monitor and provide supportive supervision has been inadequate, program stewardship weak, and funding, though large and growing, has been insufficient to meet demands, particularly in the absence of measures to improve cost-effectiveness. Furthermore, behavior

[1] NFHS-IV, 2015-16.
[2] Shekar, Meera, Eberwein, Julia Dayton and Kakietek, Jakub. The costs of stunting in South Asia and the benefits of public investments in nutrition. Maternal & Child Nutrition (2016), 12 (Supp l. 1), pp. 186–195.
[3] Gragnolati, Michele; Shekar, Meera; Das Gupta, Monica; Bredenkamp, Caryn; Lee, Yi-Kyoung. 2005. India's Undernourished Children: A Call for Reform and Action. HNP Discussion paper series; World Bank, Washington, DC.

change communication and community response mechanisms have been under-developed and under-utilized.

6. Recognizing these lacunae, the Government of India (GoI) initiated a drive towards reform and action around 2010, when the Prime Ministers National Nutrition Council recommended the need to strengthen and restructure the ICDS program to enable a special focus on pregnant and lactating women and children under 3 years. Responding to this policy direction, the MWCD, in 2012, announced a framework for restructuring the program. Under the restructured ICDS program the AWC was repositioned, envisioning a vibrant ECD centre with focus on children under 3 years of age and care and nutrition counseling particularly for mothers of children under 3 years. One of interventions proposed under this restructured ICDS program was the initiation on an experimental basis, the provision of day care crèches for care and development of children 6-72 months of age whose mothers go to work, in 5% of existing AWCs.

7. Additionally, to facilitate action and provide greater direction on early childhood care and education in the country, the MWCD announced a National Early Childhood Care and Education (ECCE) Policy in 2013. The policy not only aimed at strengthening ECCE within the ICDS program but also provided a framework within which all institutions - governmental, non-governmental and the private sector – are to deliver ECCE, ensuring quality in service delivery across the country. The policy document defined ECCE as inclusive of both care and early stimulation and interaction needs for children below 3 years and developmentally appropriate preschool education for 3 to 6 years old. It also highlighted the need for developing innovative models and carrying out impact evaluations for ensuring an evidence-based approach to implementation.

8. It was in this context that the Department of Women and Child Development (DWCD), Government of Madhya Pradesh (GoMP), decided to initiate an AWC-cum-creche pilot in 100 centres in the State. The World Bank, with the objective of supporting the government develop a strong evidence base for ECD in the country, collaborated with the DWCD, MP to build a rigorous impact evaluation around the pilot. The ICDS is ubiquitous in India. Thus, any innovation that proves practical in one state such as Madhya Pradesh can potentially be scaled up across the state as well as in other parts of the country.

9. The evaluation is a randomized control trial of an AWC-cum-creche service model. The model involves expanding the scope of the existing Anganwadi Centres (AWC) under the GOI's Integrated Child Development Services Scheme (ICDS) to include crèche services for children under 3 years of age with the objective of strengthening the focus on early childhood care and stimulation in addition to nutrition services. The model thus seeks to enhance parental care by including child stimulation along with a package of child feeding and health care being delivered through the Government's ICDS program.

10. The evaluation studies the impact of this integrated nutrition and early child stimulation package on nutrition and cognitive development outcomes in children under 3 years of age, in particular, those provided with an opportunity to attend a day care center. As the model was implemented entirely by and through Government systems, it not only presents the possibility of a potentially scalable model but also provides a clear understanding of the barriers, limitations and opportunities if adopted and scaled up.

11. Insights from this evaluation are thus not only expected to contribute to strengthening the existing program by providing knowledge on what works, what doesn't, and identifying actions to improve implementation but also contribute to the global evidence on ECD. It is also expected to build technical

capacity on ECD within the country, which has received a limited focus and potentially contribute to the development of a national creche policy.

12. The lack of knowledge on the relevance of early childhood stimulation as an important component of child development has resulted in child development efforts in India being targeted at improving health and nutrition with minimal focus being placed on early childhood stimulation. This is an area of concern as the first two years are critical for brain development and function. Appropriate care and stimulation during this period provides enabling conditions for optimal growth and development, in the absence of which the child is likely to suffer wide-ranging consequences such as a lower IQ, greater susceptibility to illness, lower educational attainment, lower productivity and lower earnings.

13. Emergent developmental research points to a potential complementarity between child stimulation and nutrition interventions. A few studies have tested this with mixed results. The most detailed long-term study was from Jamaica, where both child development and food supplementation interventions were provided to stunted children of one year or older. The study found that it is the former rather than the latter that helped close the gap in cognitive skills. The supplementation in this study was moderate – 1 kg of baby formula per week – and the stimulation – weekly home visits by trained community workers - comparatively intense and costly.

14. A replication of the Jamaica study in Colombia found little or no synergy between ECD and nutrition[4]. The study assessed the effect of a pyscho-social home-visiting intervention and multiple micronutrient supplementation on children aged 12 to 24 months, both separately and combined. Both interventions lasted 18 months and while the psycho-social intervention led to gains in children's cognitive development and in receptive language skills. The micronutrient supplements had no significant effect, neither alone nor when paired with the stimulation. A follow-up survey to assess whether the gains were maintained two years after the program ended found that the gains faded out. This was either because the initial effects on child development were too small to be sustained, or the lack of continued home stimulation support could not consolidate the gains made.

15. A less resource intensive trial in Pakistan[5] also assessed the synergy of micro-nutrient supplementation and nutritional counseling and stimulation provide during monthly visits to the homes of children 0-24 months. Set in rural Pakistan, the study found benefits to cognition, executive function, pre-academic skills, and behaviour from monthly group sessions and home visits 2 years after the intervention ended, when children were 4 years of age. However, no apparent synergy with the nutritional counseling was found. Moreover, there was no impact on nutritional status from ECD by the time a child was two years of age.

[4] Attanasio OP, Grantham-McGregor SM, Fernandez C, Fitzsimons EO, Rubio-Codina M, Meghir C. Enriching the home environment of low-income families in Colombia: a strategy to promote child development at scale. Early Child Matters. 2013; 1:35±9.
Andrew AR, Attanasio OP, Fitzsimons EO, Rubio-Codina M. Why is multiple micronutrient powder ineffective at reducing anaemia among 12±24 month olds in Colombia? Evidence from a randomised controlled trial. SSM Popul Health. 2016; 2:95±104. https://doi.org/10.1016/j.ssmph.2016.02.004 PMID: 29349132.
Andrew A, Attanasio O, Fitzsimons E, Grantham-McGregor S, Meghir C, Rubio-Codina M (2018) Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in Colombia. PLoS Med 15(4): e1002556. https://doi.org/10.1371/journal.pmed.1002556.
[5] Yousafzai AK, Obradović J, Rasheed MA, Rizvi A, Portilla XA, Tirado-Strayer N, et al. Effects of responsive stimulation and nutrition interventions on children's development and growth at age 4 years in a disadvantaged population in Pakistan: a longitudinal follow-up of a cluster-randomised factorial effectiveness trial. Lancet Glob Health. 2016; 4:e548±58. https://doi.org/10.1016/S2214-109X(16)30100-0 PMID: 27342433.

16. In India, a longitudinal, cluster randomized trial[6] among beneficiaries of the ICDS program enrolled during pregnancy was carried out. The study compared three groups: 1) ICDS only; 2) ICDS + breastfeeding and complementary feeding counseling; and 3) ICDS + breastfeeding and complementary feeding counseling + responsive complementary feeding skills and psychosocial stimulation. Compared to the ICDS only, both intervention packages had positive impacts on dietary intake, hemoglobin and reduced morbidity in children 15 months of age. Stimulation did not increase the impact of the nutrition intervention on these outcomes, and only the nutrition intervention had a positive impact on length gain. An additive effect of nutrition and stimulation was found, however, on development outcomes. The study cautioned that the number of messages promoted was counter-productive and speculated that fewer messages might have had a larger impact, which is noteworthy in terms of lessons for future intervention design.

17. As a generalization, overall, there is limited evidence of synergy between stimulation and nutrition interventions in the small-scale programs and trials specifically designed to test this possibility; running both types of interventions jointly often leads to that same outcomes as running both independently[7]. Thus, integrated programs appear additive rather than synergistic. However, even if programs are additive, the continuity of implementation can provide a life cycle orientation that improves management as well as participation of target populations.

18. More specific to the program under study, not only is the literature on day care programs mixed (as it is for many ECD interventions), there is also evidence that poorly designed programs can be detrimental to child development. For example, while day care centers in Ecuador increased mother's labor force participation, they had a negative impact on cognitive development of those children who attended[8]. Similarly, labor supply increased, and socioemotional behavior worsened in a daycare project in Quebec.[9] Moreover, while homebased day care in Colombia had measurable benefits for children[10], the transition to center-based day care has a negative effect on child development.[11]

19. This evaluation, by focusing on a potentially scalable model of integrated nutrition and early childhood stimulation service delivery, through the GoI's national nutrition program under its own management, will not only contribute to the knowledge base regarding the additional value of ECD services attached to the current nutritional program of the ICDS, but also provide insights into how to improve ECD service delivery. Few studies have contributed information on the potential role of integrated ICDS services in enhancing both cognitive development and nutrition outcomes.

20. Another auxiliary outcome that was explored is the change in labor force participation and time allocation facilitated by the availability of child care or creche services provided through the AWC-cum-creche program. Labor force participation rates of mothers are known to be significantly lower than those of other women. A mother's decision to join the labor force could depend on the cost of day care compared to her wages from work, and insufficient or inadequate child care options could be

[6] Vazir S, Engle P, Balakrishna N, Griffiths PL, Johnson SL, Creed-Kanashiro H, Fernandez Rao S, Shroff MR, Bentley ME. Cluster-randomized trial on complementary and responsive feeding education to caregivers found improved dietary intake, growth and development among rural Indian toddlers. Matern Child Nutr. 2013 Jan;9(1):99-117. doi: 10.1111/j.1740-8709.2012.00413.x.Epub 2012 May 24.
[7] Alderman, Harold and Lia Fernald. The Nexus Between Nutrition and Early Child Development. Annual Review of Nutrition. 37: 447-476. 2017.
[8] Rosero, José, and Hessel Oosterbeek. "Trade-offs between different early childhood interventions: Evidence from Ecuador." (2011).
[9] Baker, Michael, Jonathan Gruber, and Kevin Milligan. "Universal child care, maternal labor supply, and family well-being." *Journal of political Economy* 116, no. 4 (2008): 709-745.
[10] Bernal, R., & Fernández, C. (2013). Subsidized childcare and child development in Colombia: Effects of Hogares Comunitarios de Bienestar as a function of timing and length of exposure. Social Science & Medicine, 97, 241-249.
[11] Bernal, Raquel, Orazio Pietro Attanasio, Ximena Peña, and Marcos Vera-Hernández. *The Effects of the Transition from Home-Based Community Nurseries to Child-Care Centers on Children in Colombia*. Working paper, Universidad de los Andes, Bogotá, Colombia, 2016.

barrier to labor force participation of women. This is especially relevant for poor households where the lack of child care options limiting labor force participation could lead to a worsening of child well-being. In general, increased access to day care and lower prices for such services facilitates labor force participation of caregivers (and sometimes their own mothers), as shown in studies in settings as diverse as Kenya, Chile, and Russia as well as within the European Union.[12]

21. The evaluation also delves into the emergent issue of the role of maternal depression on ECD outcomes. This issue has not received due attention in India. Maternal depression is known to be a risk factor for undernutrition as well as for delayed cognitive development in many low and middle-income countries[13,14]. It potentially affects child outcomes from very early during pregnancy (through altered placental function, epigenetic changes, and stress reactivity) to postnatal period, infancy and childhood (via altered mother – child interactions, less affection and responsiveness, poor psychosocial stimulation, inadequate feeding, poor hygiene and health-seeking practices[15]. Exploring the prevalence of this risk in the Indian context will contribute to an improved understanding of existing barriers to nutrition and child development outcomes in India.

22. Furthermore, given the challenge of implementing developmental assessments for infants and small children through large household surveys, the team took the opportunity at the endline survey to contribute to the ongoing validation of an easy implement instrument - the Caregiver Reported Early Development Instrument (CREDI) - for measuring the developmental status of children under 3 years of age at the population level[16]. Existing instruments, such as the Bayley's Scales of Infant and Toddler Development (BSID) are typically done by clinically trained personnel and provide detailed information on individuals' developmental status that can be used for informing clinical decisions, understanding developmental processes, or testing the efficacy of early interventions[17]. While these have previously been used to assess developmental status at scale[18], implementing them is quite costly as is the time taken to administer them. In addition, researchers need to pay excessive attention on the interviewers' level of education, training field interviewers to administer clinical assessments, finding suitable near-clinical settings in which to administer the scale in the field, while trying to establish reliability and maintain quality control[19]. All these elements often pose obstacles to using clinical scales like the Bayley's at scale.

23. The CREDI, designed by an international committee of specialists, allows for quickly and easily measuring the motor, cognitive, and socioemotional skills of children under 3 years old living in low-

[12] Garcia-Moran EM. Child Care Costs, Female Labor Force Participation and Public Policy. Job Market Paper, Nov 2010.
Lokshin MM, Glinskaya E. and Garcia M. The effect of early childhood development programmes on womens labor force participation and older childrens schooling in Kenya. Journal of African Economies, Volume 13, Number 2, PP 240-276 doi:10.1093/jae/ejh009
Bick A. The Quantitative Role of Child Care for Female Labor Force Participation and Fertility. Journal of the European Economic Association, Volume 14, Issue 3, 1 June 2016, Pages 639–668, https://doi.org/10.1111/jeea.12143. Attanasio, Orazio, Ricardo Paes de Barros, Pedro Carneiro, David Evans, Lycia Lima, Rosane Mendonca, Pedro Olinto, and Norbert Schady. 2017. *Free Access to Child Care, Labor Supply, and Child Development*. #ie Impact Evalauito Rep[ort #58. Lokshin, Michael. Household Childcare Choices and Women's Work Behavior in Russia *J. Human Resources October 2, 2004 XXXIX:1094-1115*
[13] Surkan, P. J., Kennedy, C. E., Hurley, K. M., & Black, M. M. (2011). Maternal depression and early childhood growth in developing countries: Systematic review and meta-analysis. Bulletin of the World Health Organization, 89, 608–615.
Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., … Bhutta, Z. A. (2017). Nurturing care: Promoting early childhood development. Lancet, 389, 91–102.
[14] Phuong Hong Nguyen, Jed Friedman, Mohini Kak, Purnima Menon and Harold Alderman. Maternal depressive symptoms are negatively associated with child growth and development: Evidence from rural India. Maternal Child Nutrition. Forthcoming.
[15] Herba, C. M., Glover, V., Ramchandani, P. G., & Rondon, M. B. (2016). Maternal depression and mental health in early childhood: An examination of underlying mechanisms in low-income and middle-income countries. Lancet Psychiatry, 3, 983–992.
[16] https://sites.sph.harvard.edu/credi/.
[17] Snow CE, Van Hemel SB. Early childhood assessment: Why, what, and how. Washington: The National Academies Press; 2008.
[18] Sprachman, Susan & van Kammen, Welmoet & Salem, Margo. (2018). Considerations for Implementing a Child Assessment in a Field Survey Research Project. Minneapolis, MN: Society for Research in Child Development Conference.
[19] McCoy DC, Sudfeld CR, Bellinger DC, Muhihi A, Ashery G, Weary TE, Fawzi W and Fink G (2017). Development and validation of an early childhood development scale for use in low-resourced settings Population Health Metrics 15:3 DOI 10.1186/s12963-017-0122-8.

resourced settings, and was used to measure cognitive development of children in the endline survey, while the BSID was administered to a sub-sample for validation.

24. The BSID is considered the most reliable measure of early child development in the ages up to 42 months. However, due to the cost and time requirements, these are not often used in impact evaluation of an intervention implemented at a relatively large scale nor in direct policy dialogue. The inclusion of the validation exercise within the evaluation can thus both increase the power of the policy dialogue on day care but also will provide an assessment of the reliability of caregiver response modules for future policy discussions.

II. The Intervention

25. The Department of Women and Child Development, Madhya Pradesh in 2014, initiated the process of setting up 100 pilot AWC-cum-creches (AWCCs) across 2 districts - Dhar and Singrauli – of Madhya Pradesh. 63 centres were set up in Dhar and 37 in Singrauli. The primary stated purpose for setting up the AWCCs was to provide nutrition and care facilities to children of working mothers. The two districts, on opposite sides of the state, were selected by the department on the basis of relative deprivation.

26. ***The AWC-cum-creche Model***[20]: The AWCC as planned and financed by the DWCD entailed an additional creche worker trained to provide nutrition and early stimulation and care to children under 3 years of age at the AWC and to work under the guidance of the existing AWW. This was expected to allow infants and young children (below 3 years of age) of working mothers to be provided the requisite care and nutritional support; as well as strengthen the programs ability to improve outreach activities by providing additional support to the AWW. Children of non-working mothers were also welcomed at the AWCC, although the idea for creche services emerged from the need to provide child care support to working mothers.

27. As initially planned, the role of the World Bank was primarily to support the impact evaluation and it advised the DWCD on the random selection of AWCs in the 2 districts identified by them. The selected AWCs were converted into AWCCs by the DWCD. The conversion involved (a) the provision of toys and other relevant materials such as cradles, bedding and cleaning materials for smaller children; (b) rent for additional space to accommodate creche children (if space in the existing AWC was too small, the additional rent would allow a bigger space to be rented); and (c) additional funds for provision of a meal in the late afternoon (the morning meals were already being provided through the ongoing ICDS program).

28. The creche worker was selected and appointed by the health committee under the local Panchayat and paid an honorarium of INR 3000/- per month. The AWCC was directed to remain open for 7.5 hours per day and each centre was provided the flexibility to decide the opening time based on local conditions. The centres were advised to synchronize the timings with those of the primary schools so that it would be convenient for older siblings to support their parents in dropping and picking up their younger siblings from the AWCC.

[20] Circular issued by the DWCD, Madhya Pradesh on January 30, 2014 to the 2 pilot districts.

29. ***Mechanism of change/impact:*** The model was expected to bring about a change in nutrition and cognitive development through the following mechanisms or pathway.

Table 1. Pathway of Change

Input	**Activity**	**Output**	**Outcome**	**Impact**
Additional resources to ensure implementation of second AWC worker	Recruitment and training of additional AWC worker	Increased focus on ECD at AWC, including home visits from worker to mother	Increased incidence of stimulative learning and play at home	Increased cognitive ability
Technical support and capacity building for development of appropriate tools for ECD for children under 3	Ensuring adequate supervision structure of AWW centres Development of supportive learning materials Conduct process monitoring for program implementation	Nutrition counselling and parenting lessons on early child stimulation offered to all mothers of young children	Increased time allocated (primarily or secondarily) to parent/child interaction Increased rates of exclusive breastfeeding (up to 6 months) Increased rates and improved quality of complementary feeding for malnourished	Increased nutritional status

30. ***Training of the creche workers*****:** As initially planned, the evaluation was conceived as an assessment of the effectiveness of the state's program. That is, the intervention was not viewed as a proof of concept or a state of the art program designed by researchers. However, given the limited experience of the state in service provision through creches to children under 3 years of age, technical assistance in the form of training to newly hired workers was facilitated by the World Bank and the International Food Policy Research Institute (IFPRI). IFPRI engaged Mobile Creches (MC), a well-known organization, that works for the right to Early Childhood Development for marginalized children to provide the DWCD, MP the required capacity building support. MC, which has done extensive work spanning from grassroot level interventions to policy advocacy at the national level, designed and implemented a 12-day pre-service training for the 91 creche workers that had been recruited for the pilot sites. They also facilitated the setup of 4 centres (2 in Dhar and 2 in Singrauli) as demonstration centres as part of the capacity building on creche operationalization. They further oriented the existing AWWs from the pilot sites on the new model to enable them to work in coordination with the new creche workers as well as oriented supervising officials from the 2 districts to effectively monitor and support these pilot centres. Through this process, they oriented 82 AWWs, 14 Child Development Project Officers, 33 supervisors, 1 ECCE coordinator, 2 District Programme Officers, 1 Additional Director and 1 Deputy Director. They also supported creche workers undertake community communication events and conducted follow-up visits to assess quality, efficacy of the services and

identify gaps. The community events were held in about 22 villages in January 2016, where about 15-25 parents and self-help group members were mobilized in each village and oriented on the importance and need for creches. The focus was to evoke a dialogue with working parents on their children and their care while they were at work. The meetings highlighted that about half the parents were aware of creche facilities at the centres, working parents felt that they lacked time for early child stimulation activities, and the distance between their home and the centre was a barrier in their child's enrolment in the creche facility as dropping them to the centre took time.

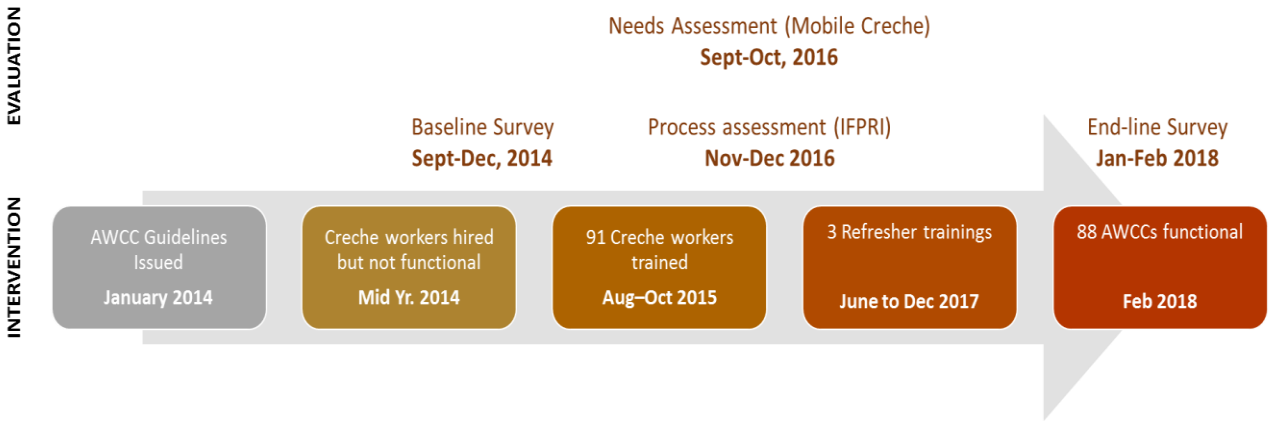
31. One year after the induction training, a series of refresher trainings for the creche workers, AWWs and AW helpers were planned and supported by the World Bank and delivered through MC. These were:

- A 4-day refresher training for creche workers and AWWs (with AWWs coming in for 2 days of the 4-day training) in June 2017. The sessions for the creche workers focused primarily on understanding dimensions of early childhood development (physical, cognitive, social and emotional) and skill development on early childhood stimulation and nutrition; while joint sessions with AWWs focused on understanding each other roles, coordination and teaching children using play-way methods.

- A 4-day refresher training for creche workers and AW helpers (with AW helpers coming in for 2 days of the 4-day training) in July 2017. The sessions for the creche workers focused on understanding issues faced by creche workers in delivering services and refreshing their understanding of ECD and skills for delivering early childhood stimulation. The joint sessions with AW helpers focused on coordination between the two workers in ensuring cleanliness, child care, nutrition and early child stimulation services at the centre.

- A 4-day refresher training for creche workers and AWWs (with AWWs coming in for 2 days of the 4-day training) in December 2017. The sessions for the creche workers focused on assessing their knowledge on ECD and then focusing on identified gap areas in knowledge; and building skills in new games and methods of early stimulation. The joint sessions with AWWs also focused on first assessing gaps in understanding, addressing these gaps and then orienting them in new methods of teaching-learning.

Figure 1. Intervention and Evaluation Timeline



Key Implementation Issues:

32. The AWC-cum-creche model as defined above was made effective through the circular issued in January 2014. However, given the lack of orientation or detailed guidance provided to the districts on how to implement the intervention, there was negligible movement on the ground. Recognizing this gap, the World Bank provided technical assistance to the DWCD, MP towards the development of detailed implementation guidelines for the AWCC model. While these guidelines were further translated and adapted into Hindi in October 2014 by the DWCD, they were never issued due to a change in leadership and the belief that these were too detailed and would never be absorbed by the field functionaries. Thus, the model remained loosely defined as per the January 2014 circular.

33. Creche workers were hired by mid-late June 2014 but were not provided with any training or orientation on their newly designed roles at the AWC for over a year. They therefore started functioning as additional hands to the AWWs. Additional materials, including learning and play materials were provided to the 100 centres, but no guidance was provided on how to use them for early play, stimulation and learning. Many of them had not received their honorarium as well and they remained largely un-motivated and directionless. Therefore, in actuality the intervention only started in October 2015 when the newly hired workers were trained on early childhood development and their specific roles and responsibilities. This was almost 1.5 years after the AWCC guidelines were issued and at the time of the endline survey, the intervention had been in implementation for approximately 2.5 years.

34. As is evident, constraints in implementation were a constant from the inception of the AWCC pilot. Although the DWCD sustained the pilot, the investment, focus and priority placed on the pilot diminished soon after its initiation. This was in line with the leadership and larger programmatic changes that were happening at the national level. Over the period of implementation, the limitations and lacunae in AWCC functioning were highlighted several times to the State leadership. In fact, a process evaluation of the AWCC was carried out in November 2016 and concrete feedback provided to the DWCD based on its findings. While these issues were acknowledged and the commitment to improving implementation made, budgetary constraints were continuously presented as binding factors preventing them from taking concrete actions.

III. Methodology

35. ***Research Questions and Policy Relevance.*** Specific questions addressed with this research include:

- Does the establishment of a creche in conjunction with the AWC lead to increase utilization of early childhood education services?
- Does adding an additional worker dedicated to mentoring on child care lead to measurable improvements in cognitive development of children by the time the child is three years old?
- Does the sharing of responsibilities free up time of AWW such that they achieve measurable improvements in weights and heights of children?
- Does the availability of day care lead to increased work on farms or for wages?

36. These questions all address the issue of the quality of service delivery. As indicated, the ICDS has floundered in places on the issue of capacity. There is no obvious flaw in its underlying holistic approach but the ability of the existing program framework to achieve desired results has been constrained by the limited programmatic focus on children under 3 years of age. By studying an innovation that will be delivered through the existing ICDS framework in a manner that can be replicated at the state level using minimal additional resources, the study can point to a potentially different service delivery pathway.

37. ***Evaluation Design.*** Identification of program impact is based on a randomized treatment assignment. Specifically, treatment status was determined through multi-stage stratified randomization according to the following steps:

- The randomization is based on the administration structure of the two-selected high-burden districts district. These are administratively subdivided into blocks. Within the ICDS structure, each block is further subdivided into sectors, with approximately 8 sectors per block. Each sector in turn, has on average 20-25 AWCs, which are monitored and supported by a sector Supervisor. A sector supervisor plays a key administrative role in the conduct of ICDS activities.

- Given this structure, we randomly allocated each sector within each block to either the treatment or the control (that is, the existing AWC structure). Given the high coverage of the ICDS, no attempt was made to select among communities without an AWC as these would be atypical of the larger universe. Hence the unit of treatment randomization was the sector and each sector faced an equal likelihood of assignment to the treatment or the control. Potential contamination from messaging or informational spillovers was a major reason why we randomized at the ICDS sector level rather than the village level. As the ICDS sector supervisor plays a key role in AWC and community support, it is quite likely that equal treatment towards intervention and control AWCs in her catchment area would be difficult to maintain. In addition, AWWs are often networked within their sector and may easily share messages during joint supervision or training meetings.

- 100 communities (an average of three per sector) were assigned to the treatment and an equal number assigned to the control, selected with the best covariate balance following Bruhn and McKenzie[21]. Covariates considered include the rate of child wasting as reported by ICDS program, as well as AWC catchment size. Power calculations utilized intraclass correlation coefficient

[21] Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." American Economic Journal: Applied Economics, 1 (4): 200-232.

(ICCs) of anthropometry from previous survey work in Madhya Pradesh by IFPRI. The choice between number of clusters and individuals within each cluster was also influenced by the expected number of age eligible children within each community.

- Prior to household selection into the interview sample, every household in each community (as proxied by the primary sampling unit) was listed to determine the presence of an age eligible child. 15 households per cluster were randomly selected for interviews from the list of all households with age eligible children if the community had more than 15 children otherwise all children in the age category were selected.

38. Analysis utilized repeated cross section surveys. The main analysis provides intent-to-treat (ITT) estimates although treatment on the treated results can also be explored at a later date. As is standard in RCTs, we included unadjusted results based on treatment assignment as well as results that include covariates such as household and individual characteristics. All outcomes are estimated with ordinary least squares (OLS). Standard errors are clustered at the sector level, employing the Huber-White sandwich estimator.

39. Analysis with the baseline data (described further below) suggests that this study design is relatively well powered to identify moderate or even fairly small impacts of the program on priority outcomes. The table below (Table 2) suggests that the study design is powered to identify a change in the rate of AWC use of 0.18 standard deviations (SDs), a change in the rate of 0.15 SDs in the rate of stunting, 0.22 SD change in the female labor force participation rate, and a 0.16 SD change in the estimated prevalence of developmental delay (as assessed with the ASQ).

Table 2. Estimated minimum detectable effects sizes (MDES) anticipated given the study design, for priority outcomes, as informed by baseline data values

Outcome	MDES	Mean	SD
AWC use	0.179	0.3719	0.4834
Stunted	0.1501	0.6444	0.4788
Maternal labor supply	0.2155	0.5396	0.4985
Development delay - problem solving (ASQ)	0.155	0.1977	0.3983

40. The trial, Making Integration the Operative Concept in the Indian Integrated Child Development Services, is AEARCTR-0000967 registered at the AEA RCT Registry. Both rounds of data collection were approved by IFPRI's IRB, approval # 00005121 and # 00007490 respectively.

IV. Data Collection

41. Data collection for this evaluation was completed over two rounds. Baseline data collection was undertaken by Oxford Policy Management Ltd. (OPML) and commenced on September 17, 2014 in Dhar and concluded on November 30, 2014 in Singrauli. There were breaks in the middle of the data collection process to accommodate the festivals of Dussehra and Diwali which occurred between late October and early November. Endline data collection was also led by OPML and commenced 38 months after the completion of the baseline study. Field teams were deployed first in Dhar starting on January 12, 2018 after which they moved to Singrauli where data collection concluded on February 26, 2018. One week of time was allotted after this for revisits in Dhar district to account for sample shortfalls in a few study areas of the district. These revisits were completed between March 1-8, 2018.

Table 3. Timeline of Data Collection

Districts	Baseline		Endline	
	Start Date	End Date	Start Date	End Date
Dhar	Sep 17th, 2014	Nov 4th, 2014	Jan 12th, 2018	Feb 5th, 2018
Singrauli	Nov 7th, 2014	Nov30th, 2014	Feb 8th, 2018	Feb 26th, 2018
Dhar Revisit	--	--	Mar 1st, 2018	Mar 8th, 2018

42. The respondent groups engaged for this study were:
    i. Mothers/primary caregivers of children
    ii. Sampled index children between the aged 6-48 months (at baseline) and 18-42 months (at endline).
    iii. Anganwadi workers at study AWCs
    iv. Creche workers stationed at the 100 pilot AWCC centers.

43. The tools used for data collection from these respondent groups are described in Table 4 below:

Table 4. Data Collection Tools

	Study Tool	Description	Person Administered To
1	**Household Roster**	The household roster captured information on all active members of the selected households. Information was gathered on age, gender, dates of birth for children under 6 years of age, education levels and employment information. The roster was also used to identify the biological mother and father of index children and the primary caregiver for the child (if different from the mother).	Primary caregiver or household member best equipped to provide information.
2	**Primary Caregiver Questionnaire**	Captured family characteristics (religion, social category, ownership of assets), time allocation across work and non-work activities, knowledge of and engagement in child development activities, use of Anganwadi Center / AWCC services, index child's	Primary caregiver of index child

		diet intake, and measures of primary caregiver depression.	
3	**Caregiver Reported Early Childhood Development Instrument (CREDI)**	The CREDI was administered to all caregivers who answered the primary caregiver questionnaire. The instrument captured caregiver reports of behaviours in index children that are indicative of cognitive, motor, language, socio-emotional, and mental health development during early childhood. The CREDI was only utilized at the endline survey. The baseline survey employed the Ages and Stages (III) development instrument. Baseline balance will be assessed with the ASQ, but impacts with the CREDI will only be estimated with the single-difference analysis.	Primary caregiver of index child
4	**Bayley's Scales of Infant and Toddler Development, Third Edition (Bayley-III)**	The Bayley-III is an in-depth test of a child's cognitive, motor and language skills. The test was administered to a sub-sample of 1000~ index children in the age range of 24 to 36 months by specially trained testers.	A sub-sample of index children aged 24-36 months
5	**Anthropometry**	Anthropometric measurements (height/length and weight) were taken for all children under 6 in sampled households by trained investigators.	Children under 6 years in sampled households.
6	**Anganwadi Worker Questionnaire**	Captured information on AWW demographics, workload and service provision, early childhood development and nutrition knowledge, implementation of pre-school education activities at the AWC, trainings and supervision.	Anganwadi Worker in study area
7	**Creche Worker Questionnaire**	Captured information in the study's treatment arm on Creche Worker demographics, workload and service provision, child feeding and early childhood development knowledge, availability of materials for provision of AWCC services, training and supervision.	Creche Worker in study area
8	**Primary Sampling Unit Questionnaire**	Captured demographic information and measures of public service delivery in the study areas.	Any local official with records of information about the area (Mukhiya, Ward Member, Anganwadi Worker etc.)

44. The baseline survey for the study reached 2,934 primary caregivers and index children between the ages of 6 to 48 months. The endline survey was completed with 2,856 primary caregivers and index children between the ages of 18 and 42 months. A sub-set of 1037 index children falling within the age range of 24-36 months were additionally administered the Bayley-III. Both baseline and endline surveys also collected information from Anganwadi workers and creche workers in both treatment and control arms of the study. A separate team was employed for the BSID. RehabInsights was responsible for the adaptation of the instrument to local conditions and to administer it to the randomly selected subset of children.

45. The detailed sample distribution across different study components is described in Table 5 below.

Table 5. Sample Coverage (by phase and district)

Component	2014			2018		
	Dhar	**Singrauli**	**Overall**	**Dhar**	**Singrauli**	**Overall**
Primary Caregivers	1830	1104	2934	1722	1134	2856
Bayley-III	--	--	--	631	406	1037
CREDI	--	--	--	1722	1134	2856
ASQ	1830	1104	2934	--	--	--
Anganwadi Workers	117	64	181	119	69	188
Creche Workers	--	--	--	51	31	82

46. **Survey Non-Response** Reasons for non-participation in the survey during endline data collection are described in Table 6 below. The primary reason for non-participation was migration away from the study area during time of data collection.

Table 6. Reasons for Non-Participation

Result of interview	**Count**	**Percent**
Refused	3	0.09
Household locked	117	3.5
Entire household absent for extended period of time	83	2.48
Household not found/located	22	0.66
Primary caregiver and child absent for extended period of time	263	7.86

V. Sample Characteristics

Household Social Profile

47. Table 7 presents the socio-economic profile of sampled households in 2018 along with national and state-level estimates taken from the 2011 Census. Hindu households represent a large majority of households (96 percent) in the sample, followed by three percent Muslim households. This closely resembles the population-level religious distribution in Madhya Pradesh which has 91 percent Hindu and six percent Muslim populations, respectively.

48. Table 7 also presents the social categories of sampled households along with state and national estimates. Scheduled tribe (ST) households are the most commonly found in the sample, making up 49 percent of all households. Eight percent of households belong to scheduled castes (SC). The rest of the households (43 percent) belong to other/general categories. The sample has an over-representation of ST households and under-representation of SC households, compared to the state and national averages.

Table 7. Socio-economic profile: Sampled Households and State/National Estimates

Households	**2018**	**Madhya Pradesh 2011 Census[22]**	**Census 2011[23]/ SECC 2011[24]**
Religion			
Hinduism (%)	96.15	90.89	79.80
Islam (%)	3.68	6.57	14.23
Christianity (%)	0.07	0.29	2.30
Sikhism (%)	0.04	0.21	1.72
Other religions (%)	0.07	1.61	1.73
Social Category			
Scheduled Tribes (%)	48.51	25.29	10.98
Scheduled Castes (%)	8.27	15.20	18.45
Others (%)	43.22	59.06	68.50

[22] Madhya Pradesh Religion Census 2011. Accessed on 30th May 2018 at: https://www.census2011.co.in/data/religion/state/23-madhya-pradesh.html.
[23] All India Religion Census 2011. Accessed on 30th May 2018 at: https://www.census2011.co.in/religion.php.
[24] Socio-economic and Caste Census 2011 Caste Profile Report. Accessed on 30th May 2018 at: http://secc.gov.in/statewiseCasteProfileReport?reportType=Caste%20Profile#.

Household Economic Profile

49. Table 8 presents the employment distribution of heads of households sampled for the survey. The most common occupation amongst heads of households is in agriculture, followed by casual wage labor in non-agricultural activities. Over ten percent of households across the two phases of data collection reported working in governmental or non-governmental salaried jobs. Table 8 also presents employment information disaggregated by wealth quintiles[25] where '5' represents the richest and '1' represents the poorest households present in the sample. In the study sample, poorer households tend to participate more in agricultural employment or non-agricultural wage labor, while richer households are associated more frequently with running a non-agricultural family business or holding regular salaried employment.

Table 8. Household Head's Employment by Wealth

Wealth Quintile	**1**		**2**		**3**		**4**		**5**		**Overall**	
Employment (%)	**2014**	**2018**	**2014**	**2018**	**2014**	**2018**	**2014**	**2018**	**2014**	**2018**	**2014**	**2018**
Not working	3.2	8.3	3.7	4.4	4.8	5.8	5.0	7.3	9.4	10.8	5.15	7.49
Wage labor	26.3	34.1	15.6	30.9	19.6	28.6	19.1	25.6	8.5	6.9	17.86	24.96
Agriculture	61.6	43.6	70.0	47.2	66.0	46.1	50.3	34.5	29.4	29.8	56.17	39.74
Non-agricultural family business	2.5	4.4	2.6	7.3	2.7	7.2	8.9	12.0	25.0	22.9	8.01	10.96
Salaried (government or private sector)	5.3	6.8	5.9	7.7	5.6	10.8	13.7	18.7	22.7	26.9	10.33	14.6
Housewife/Domestic Duties	0.8	1.8	0.9	1.5	0.8	0.7	1.4	0.9	2.1	2.2	1.19	1.44

Primary Caregiver Demographics

50. Table 9 notes the age and educational and occupational characteristics of primary caregivers found in sampled households. The distribution of educational attainment amongst primary caregivers remains similar between the two phases of data collection, with never having attended school being the most common. Primary caregiver employment is concentrated in agriculture with over 50 percent reporting participation in agricultural labor across the two phases.

[25] Wealth quintiles have been determined based on the construction of a wealth index using Principal Components Analysis.

Table 9. Primary Caregiver Demographics

	2014	2018
Mean Age	26.20	26.45
Educational attainment (%)		
Never attended school	47.04	40.12
Up to primary school (grades 1-5)	21.41	18.03
Up to middle school (grades 6-8)	16.24	17.89
Up to secondary school (grades 9-10)	7.67	13.62
Up to senior secondary school (grades 11-12)	4.11	6.25
Undergraduate level and above (BA+)	3.53	4.02
Employment[26] (%)		
Agricultural labor	52.28	52.03
Non-agricultural labor	3.21	5.08
Household business	1.81	3.12

Index Children Demographics

51. Table 10 provides a summary of the distribution of index children by gender and a mean estimate of age in months across the two phases of data collection.

Table 10. Index Child Characteristics

Indicator	**Control**		**Treatment**		**Overall**	
	2014	**2018**	**2014**	**2018**	**2014**	**2018**
Index Child Sex						
% Female	50.61	48.53	50	50.45	50.31	49.49
% Male	49.39	51.47	50	49.55	49.69	50.51
Mean Child Age (in months)	28.56	29.95	28.69	30.04	28.62	29.99

[26] The reference period for employment is 3 months in 2014, and 1 month in 2018.

52. Table 11 provides a summary of the distribution of index children across age categories ranging from 6 to 54 months. This reflects that the selection criteria were wider at baseline (6 to 48 months) than at endline (18 to 42 months).

Table 11. Index child age range (months)

Age Range (in months)	**6-11**	**12-17**	**18-23**	**24-29**	**30-35**	**36-42**	**43-48**	**49-54**
2014	221	332	381	550	515	644	285	6
2018	1	6	501	873	771	702	1	0

VI. Results/Outcomes

53. This section presents the main findings of the study. As the surveys were repeated cross sections, most of the results are obtained by using difference in difference analysis. The cognitive and socio-emotional development results, however, are based on data which were not collected in the baseline. Thus, the analysis employs standard randomization. In the absence of repeated observations in the same individuals, we cannot employ standard ANCOVA[27] estimations. However, we build upon the logic of such models by including baseline sector averages of the outcomes in question as regressors in lieu of fixed effects. This generally improves the precision of results.

54. Although there were only three years between the surveys, there is evidence that substantial development has occurred through the districts. Comparisons of the 2006 and 2016 National Family Health Surveys [comparable to Demographic Health Surveys surveys] shows appreciable increases in ICDS utilization throughout the country. Our data also indicates that initiatives such as the provision of cooking gas and construction of latrines were also rolled out in this period. Even the average education of care giver was higher in the resurvey, reflecting investments made in the decade prior to the study. Thus, all treatment effects – or their absences – are occurring in the context of measurable secular development.

Balance of baseline characteristics

55. Before exploring the intervention results on priority outcomes of interest, the analysis first explores characteristic balance in baseline values, summarized in Annex 2. Treatment and control mean or percent, mean differences between treatment and control, and p-values adjusted for clustering at the community level are presented along with normalized mean differences. Normalized mean differences suggest an imbalance between treatment and control for any value over 0.2[28].

56. The balance analysis reviews sample demographics, household socio-economic state (SES) measures, AWC utilization, nutrition and immunization indicators, cognitive assessments from the ASQ, and AWW characteristics (including visits). Out of the various characteristics explored, none are significantly different in mean value between treatment and control communities nor do any have a normalized mean difference above 0.15 (except for the gender of the caregiver, where 99% of caregivers in treatment communities are female and 100% in control communities – a difference unlikely to be salient for intervention outcomes). This review of balance in baseline characteristics lends confidence that any observed difference between treatment and control in mean outcome represents a causal impact of the intervention.

Service Utilization

57. The first outcome investigated is utilization of Creche and AWC services, and whether overall utilization in the community increases as a result of the establishment of creche services (Table 12). In terms of access specifically to Early Childhood Education (ECE) services, the forte of the AWC-cum-

[27] McKenzie, David. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of development Economics* 99, no. 2 (2012): 210-221.
[28] Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47, no. 1: 5-86

creche, there is an increase of 8 percentage points as well as an increase of 0.47 days in terms of days visited the creche or AWC in the previous week. There is also a general trend in AWC and ECE usage in the sample.

Table 12. AWC (ECE) and Creche usage – Index child estimates only

VARIABLES	(1) Attends creche/ECE (binary)	(2) Days went to creche/AWC (past week)
Treatment-Period Interaction	0.0818*	0.471**
	(0.0470)	(0.200)
Period	0.0966***	0.619***
	(0.0357)	(0.142)
Treatment	0.0314	0.0940
	(0.0423)	(0.136)
Primary Caregiver's Education = 2, Primary (Less than 5 years)	0.0167	0.0235
	(0.0210)	(0.0970)
Primary Caregiver's Education = 3, Middle (6-8 years)	-0.0307	-0.368***
	(0.0243)	(0.102)
Primary Caregiver's Education = 4, Higher (9-10 years)	-0.108***	-0.589***
	(0.0279)	(0.120)
Primary Caregiver's Education = 5, Secondary (11-12 years)	-0.109**	-0.653***
	(0.0468)	(0.183)
Primary Caregiver's Education = 6, Undergraduate studies (13-15 years)	-0.212***	-0.922***
	(0.0406)	(0.208)
Primary Caregiver's Education = 7, Postgraduate (More than 15 years)	-0.249***	-1.230***
	(0.0735)	(0.170)
Primary Caregiver's Education = 99, Other/Unknown	-0.544***	-2.086***
	(0.0376)	(0.168)
Primary Caregiver's Age (years)	-0.000672	-0.00473
	(0.00154)	(0.00708)
Sex	-0.0271	-0.265**
	(0.0247)	(0.106)
Child Age (months)	0.0222*	0.0962*
	(0.0115)	(0.0492)
Child Age (months) Squared	-8.44e-05	-0.00106
	(0.000188)	(0.000815)
Wealth Index	0.0295**	0.0718
	(0.0118)	(0.0478)
HH Number of Males, Age 0-4	0.0183	0.252***
	(0.0144)	(0.0691)
HH Number of Females, Age 0-4	0.0277**	0.167***
	(0.0129)	(0.0620)
HH Number of Males, Age 5-15	-0.00710	0.0278
	(0.0125)	(0.0548)
HH Number of Females, Age 5-15	0.00586	0.0847**
	(0.00767)	(0.0329)

HH Number of Males, Age 16-59	0.000510	-0.0226
	(0.00896)	(0.0482)
HH Number of Females, Age 16-59	-0.00651	-0.0125
	(0.0107)	(0.0518)
HH Number of Males, Age 60+	-0.0207	-0.0576
	(0.0173)	(0.0812)
HH Number of Females, Age 60+	-0.00133	-0.0606**
	(0.00779)	(0.0241)
Constant	-0.128	-0.717
	(0.171)	(0.756)
Observations	4,932	4,932
R-squared	0.096	0.081

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

58. If we include all observed children (not merely the index child) and also control for sector baseline means in utilization rates (akin to an ANCOVA analysis) the point estimates and precision increase. These results are presented in Table 13. When we do this, children in treatment areas are 9.9 percentage points more likely to receive ECE services and spend on average 0.57 days per week at the creche or AWC. It is apparent that service usage, especially of the ECE type, has increased in both the intervention areas and the control. However, the increase is larger in the treatment communities.

Table 13. AWC (ECE) and Creche usage – All children and additional baseline control

VARIABLES	(1) Attends creche/ECE (binary)	(2) Days went to creche/AWC (past week)
Treatment-Period Interaction	0.0992**	0.573***
	(0.0447)	(0.205)
Period	0.112***	0.687***
	(0.0307)	(0.136)
Treatment	0.0104	0.0130
	(0.0125)	(0.0319)
Primary Caregiver's Education = 2, Primary (Less than 5 years)	0.0490***	0.150*
	(0.0184)	(0.0801)
Primary Caregiver's Education = 3, Middle (6-8 years)	0.0116	-0.155*
	(0.0193)	(0.0882)
Primary Caregiver's Education = 4, Higher (9-10 years)	-0.0614**	-0.301**
	(0.0251)	(0.116)
Primary Caregiver's Education = 5, Secondary (11-12 years)	-0.0510	-0.394**
	(0.0340)	(0.152)
Primary Caregiver's Education = 6, Undergraduate studies (13-15 years)	-0.131***	-0.442**
	(0.0367)	(0.195)
Primary Caregiver's Education = 7, Postgraduate (More than 15 years)	-0.157**	-0.812***
	(0.0751)	(0.227)
Primary Caregiver's Education = 99, Other/Unknown	-0.710***	-3.000***
	(0.0360)	(0.161)

Primary Caregiver's Age (years)	0.00220	0.00371
	(0.00189)	(0.00850)
Sex	0.0229***	0.0174
	(0.00701)	(0.0349)
Child Age (months)	0.00517***	0.0138***
	(0.000775)	(0.00362)
Child Age (months) Squared	-5.02e-05***	-0.000142***
	(1.15e-05)	(5.33e-05)
Wealth Index	0.00629	0.0513
	(0.00944)	(0.0371)
HH Number of Males, Age 0-4	0.0351***	0.225***
	(0.0121)	(0.0572)
HH Number of Females, Age 0-4	0.0641***	0.278***
	(0.0118)	(0.0506)
HH Number of Males, Age 5-15	0.0105	0.0345
	(0.0131)	(0.0598)
HH Number of Females, Age 5-15	0.00594	0.0666*
	(0.00801)	(0.0334)
HH Number of Males, Age 16-59	-0.0120	-0.0557
	(0.0103)	(0.0450)
HH Number of Females, Age 16-59	-0.0146	-0.0247
	(0.0101)	(0.0417)
HH Number of Males, Age 60+	-0.0150	-0.0628
	(0.0162)	(0.0753)
HH Number of Females, Age 60+	-0.00374	-0.0631**
	(0.0101)	(0.0311)
Village Baseline Proportion Attends ECE or Creche (all)	0.687***	
	(0.0368)	
Village Baseline Average Days went to AWC or Creche (last week, all)		0.799***
		(0.0376)
Constant	-0.0902	-0.473*
	(0.0631)	(0.273)
Observations	10,611	10,611
R-squared	0.153	0.177

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Anthropometry

59. While the creches provide food, there are no additional services such as behavioral change communication and parental interaction that might foster improved nutrition associated with the creches. However, as the creche worker often functions as an auxiliary AWW, plausibly the presence of a creche can improve nutrition indirectly by freeing up the time of the AWW. However, as indicated in Table 14, there is no evidence in the data of any improvement of nutrition of the index children above the clear secular trend. Similarly, there is no improvement associated with the treatment for the larger sample of children who were not deemed index children (not shown). Anthropometry does show a strong relation with maternal education, as expected, but only a slight association with relative wealth (again, perhaps, as expected if the absolute levels of wealth do not dramatically vary in the study

sample). The sector average in the baseline is also strongly associated with current anthropometry of the index children.

Table 14. Anthropometry of Index Children - ITT Results

VARIABLES	(1) Height/Age Z-Score	(2) Weight/Age Z-Score	(3) Weight/Height Z-Score
Treatment-Period Interaction	-0.0744	-0.0191	0.0276
	(0.104)	(0.0609)	(0.0692)
Period	0.420***	0.371***	0.191***
	(0.0776)	(0.0447)	(0.0485)
Treatment	0.0458	0.00981	-0.0128
	(0.0401)	(0.0241)	(0.0303)
Primary Caregiver's Education = 2, Primary (Less than 5 years)	0.147***	0.142***	0.0992**
	(0.0516)	(0.0356)	(0.0401)
Primary Caregiver's Education = 3, Middle (6-8 years)	0.247***	0.213***	0.122**
	(0.0446)	(0.0437)	(0.0484)
Primary Caregiver's Education = 4, Higher (9-10 years)	0.372***	0.305***	0.160***
	(0.0753)	(0.0514)	(0.0476)
Primary Caregiver's Education = 5, Secondary (11-12 years)	0.671***	0.507***	0.201***
	(0.0606)	(0.0678)	(0.0706)
Primary Caregiver's Education = 6, Undergraduate studies (13-15 years)	0.690***	0.648***	0.355***
	(0.105)	(0.0951)	(0.0862)
Primary Caregiver's Education = 7, Postgraduate (More than 15 years)	0.862***	0.793***	0.432***
	(0.134)	(0.150)	(0.135)
Primary Caregiver's Education = 99, Other/Unknown	-0.190**	-0.808***	-1.043***
	(0.0775)	(0.0418)	(0.0475)
Primary Caregiver's Age (years)	0.00880**	0.00295	-0.00262
	(0.00384)	(0.00294)	(0.00257)
Sex	-0.0361	0.0267	-0.0592**
	(0.0376)	(0.0295)	(0.0281)
Child Age (months)	-0.0932***	-0.0221**	0.0349***
	(0.0138)	(0.00989)	(0.0102)
Child Age (months) Squared	0.00146***	0.000300*	-0.000491***
	(0.000236)	(0.000170)	(0.000173)
Wealth Index	0.0531**	0.0210	0.00507
	(0.0218)	(0.0140)	(0.0129)
Village Average Baseline Height/Age Z-Score (Index Children)	0.540***		
	(0.0355)		
Village Average Baseline Weight/Age Z-Score (Index Children)		0.625***	
		(0.0323)	
Village Average Baseline Weight/Height Z-Score (Index Children)			0.674***
			(0.0237)
Constant	-0.196	-0.692***	-0.903***
	(0.246)	(0.162)	(0.139)

Observations	5,624	5,624	5,617
R-squared	0.148	0.168	0.108
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Cognitive and socio-emotional development

60. As these results are based on the CREDI, only single difference at endline results are explored and summarized in Table 15. The coefficients for treatment for all three indicators of child development – cognitive, motor, and socio-emotional – are positive, but only for the latter is the result marginally statistically significant. The effect size in terms of standard deviation is 0.11. Given the limited time that the creche workers have to focus on child stimulation, it is not surprising that the most noticeable impact is on socio-emotional development, a process that is promoted by a child gaining confidence in the absence of their main caregiver and in interaction with other children.

61. The overall results follow expected patterns: development increases with child's age, maternal education and wealth. There is no difference across genders. However, despite the significance of these factors, the overall regressions capture a modest share of the variance. Inclusion of the cluster average score from the previous round (Table 16) has little effect on the total estimates. Additional details on the measures of child development are provided in the Annex 1.

62. As indicted, the results in Table 15 are ITT results. However, many families did not utilize the creche either because there was no staff (due to a few positions not being filled and a few caregivers being on maternity leave) or because they did not value the service. Leaving aside these endogenous issues, an approximation of the impact on the participants and considering that a quarter of the eligible children had some contact (though not necessarily in the week prior to the survey), gives a rough expectation that children who spend time in the creche have a 0.4 effect size increase in socio-emotional scores.

Table 15. Indicators of Child Development – ITT results for Index children

VARIABLES	(1) CREDI - Cognitive	(2) CREDI - Motor	(3) CREDI - Socio-Emotional
Treatment	0.521	0.157	0.438*
	(0.357)	(0.146)	(0.225)
Primary Caregiver's Age (Years)	0.0336	0.00633	0.00749
	(0.0284)	(0.0115)	(0.0164)
Primary Caregiver's Education = 2, Primary (Less than 5 years)	2.021***	0.626***	0.850***
	(0.348)	(0.157)	(0.209)
Primary Caregiver's Education = 3, Middle (6-8 years)	3.116***	1.036***	1.324***
	(0.390)	(0.170)	(0.229)
Primary Caregiver's Education = 4, Higher (9-10 years)	3.523***	1.032***	1.446***
	(0.365)	(0.171)	(0.209)
Primary Caregiver's Education = 5, Secondary (11-12 years)	5.095***	1.658***	2.372***
	(0.567)	(0.243)	(0.315)
Primary Caregiver's Education = 6, Undergraduate studies (13-15 years)	6.553***	1.870***	2.733***
	(0.691)	(0.280)	(0.275)

Primary Caregiver's Education = 7, Postgraduate (More than 15 years)	8.270***	2.067***	3.711***
	(0.885)	(0.390)	(0.379)
Primary Caregiver's Education = 99, Other/Unknown	-3.138***	-1.690***	1.633***
	(0.448)	(0.206)	(0.283)
Sex	-0.950***	-0.298***	-0.456***
	(0.241)	(0.107)	(0.128)
Child Age (Months)	1.753***	0.849***	0.969***
	(0.169)	(0.0741)	(0.0926)
Child Age (Months) Squared	-0.0187***	-0.00946***	-0.0114***
	(0.00273)	(0.00118)	(0.00146)
Wealth Index	0.784***	0.184***	0.134
	(0.138)	(0.0532)	(0.0818)
Constant	-17.59***	0.602	-5.935***
	(2.590)	(1.215)	(1.539)
Observations	2,821	2,821	2,821
R-squared	0.364	0.348	0.261

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 16. Indicators of Child Development – ITT results for Index children – also controlling for baseline sector ASQ means in outcome (ANCOVA type analysis)

VARIABLES	(1) CREDI - Cognitive	(2) CREDI - Motor	(3) CREDI - Socio-Emotional
Treatment	0.466	0.134	0.426*
	(0.340)	(0.146)	(0.219)
Primary Caregiver's Age (Years)	0.0329	0.00619	0.00693
	(0.0282)	(0.0113)	(0.0163)
Primary Caregiver's Education = 2, Primary (Less than 5 years)	1.869***	0.585***	0.777***
	(0.343)	(0.156)	(0.206)
Primary Caregiver's Education = 3, Middle (6-8 years)	2.874***	0.960***	1.202***
	(0.383)	(0.166)	(0.219)
Primary Caregiver's Education = 4, Higher (9-10 years)	3.270***	0.960***	1.305***
	(0.360)	(0.162)	(0.201)
Primary Caregiver's Education = 5, Secondary (11-12 years)	4.791***	1.555***	2.215***
	(0.575)	(0.240)	(0.300)
Primary Caregiver's Education = 6, Undergraduate studies (13-15 years)	6.206***	1.757***	2.577***
	(0.658)	(0.262)	(0.259)
Primary Caregiver's Education = 7, Postgraduate (More than 15 years)	7.919***	1.938***	3.541***
	(0.939)	(0.397)	(0.387)
Primary Caregiver's Education = 99, Other/Unknown	-3.708***	-1.730***	1.309***
	(0.554)	(0.244)	(0.368)
Sex	-0.918***	-0.285***	-0.448***

	(0.246)	(0.106)	(0.128)
Child Age (Months)	1.734***	0.855***	0.951***
	(0.165)	(0.0730)	(0.0909)
Child Age (Months) Squared	-0.0184***	-0.00957***	-0.0112***
	(0.00267)	(0.00116)	(0.00144)
Wealth Index	0.554***	0.107	0.0376
	(0.142)	(0.0689)	(0.0951)
Baseline Average ASQ Communications	0.0187	0.0109	0.0124
	(0.0302)	(0.0133)	(0.0184)
Baseline Average ASQ Gross Motor	-0.000294	-0.000846	-0.0137
	(0.0504)	(0.0215)	(0.0305)
Baseline Average ASQ Fine Motor	-0.00917	0.00762	-0.0149
	(0.0270)	(0.0111)	(0.0181)
Baseline Average ASQ Problem Solving	0.108***	0.0200	0.0690***
	(0.0364)	(0.0196)	(0.0232)
Baseline Average ASQ Personal/Social	-0.0354	-0.0143	-0.0196
	(0.0394)	(0.0161)	(0.0228)
Constant	-20.55***	-0.345	-6.943***
	(2.874)	(1.380)	(1.849)
Observations	2,822	2,822	2,821
R-squared	0.370	0.352	0.267

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Labor force participation of caregivers

63. One goal of the AWC-cum-creche intervention, as stated by the GoMP, was to free up the time of the primary caregivers so that they may devote more time to income-earning and productive activities. The principal measure of economic activity is a weekly recall of main activity: whether the caregiver worked at all, the number of days worked in the last week, and whether the caregiver supplied household agricultural labor (the main category of economic activity in the two study districts). All three of these measures have positive coefficients in Table 17, indicating an increase in economic activity. However, only the coefficients for any work or work on agriculture are significant at conventional levels (p<.10). This result suggests that the overall labor force participation of caregivers of young children has increased by four percentage points in AWC-cum-creche communities relative to controls.

64. Controlling for baseline sector-level means in the outcome of interest (similar to an Ancova model) increases the precision of the estimates. As conveyed in Table 18, a caregiver in treatment communities is more likely to work in the last week (a relative increase of 4.5 percentage points) and work an additional 0.3 days. Furthermore, as Table 13 indicates, the program induced a 9.9% increase in creche use in the previous week. Based on this, one can infer that those families which were encouraged to switch to the creche by the program increased their labor force participation by over 40 percentage points. As participation in the creche is likely correlated with the desire to participate in the labor force, it reflects what relaxing a constraint has on a subset of households and is not the expectation of the program on the overall community, should there be a wider scale up.

Table 17. Indicators of female labor force participation – ITT estimates

VARIABLES	(1) Work (Past Week Binary)	(2) Days Worked (Past Week)	(3) Work on HH Ag Labor (Binary)
Treatment-Period Interaction	0.0426*	0.277	0.0637*
	(0.0226)	(0.166)	(0.0376)
Treatment	-0.00858	-0.0613	-0.0379
	(0.0157)	(0.157)	(0.0423)
Period	-0.0265	-0.314**	0.0556*
	(0.0179)	(0.135)	(0.0290)
Age (years)	0.000967	0.00508	-0.00333**
	(0.000818)	(0.00664)	(0.00138)
Wealth Index	-0.00575	-0.0580	-0.0715***
	(0.00513)	(0.0456)	(0.0171)
Education Category = 2, Primary (Less than 5 years)	0.0275***	0.182**	-0.0260
	(0.0103)	(0.0749)	(0.0198)
Education Category = 3, Middle (6-8 years)	0.0352***	0.473***	-0.0514*
	(0.0106)	(0.0960)	(0.0301)
Education Category = 4, Higher (9-10 years)	0.0446***	0.632***	-0.155***
	(0.0116)	(0.105)	(0.0331)
Education Category = 5, Secondary (11-12 years)	0.0555***	0.854***	-0.331***
	(0.0171)	(0.142)	(0.0395)
Education Category = 6, Undergraduate studies (13-15 years)	0.0626***	0.892***	-0.375***
	(0.0192)	(0.156)	(0.0429)
Education Category = 7, Postgraduate (More than 15 years)	0.0603**	0.834***	-0.367***
	(0.0281)	(0.215)	(0.0484)
Education Category = 99, Other/Unknown	0.132***	1.683***	-0.518***
	(0.0154)	(0.133)	(0.0508)
HH Number of Males, Age 0-4	-0.0117**	-0.0707	0.00197
	(0.00530)	(0.0487)	(0.0128)
HH Number of Males, Age 5-15	0.00982*	0.0657	0.00288
	(0.00528)	(0.0414)	(0.0108)
HH Number of Males, Age 16-59	0.00480	0.0453	0.0141
	(0.00480)	(0.0350)	(0.0103)
HH Number of Males, Age 60+	0.00397	0.115*	0.0233
	(0.00872)	(0.0666)	(0.0194)
HH Number of Females, Age 0-4	-0.0148***	-0.0945**	-0.00438
	(0.00557)	(0.0426)	(0.0112)
HH Number of Females, Age 5-15	-0.00258	-0.0284	0.0262***
	(0.00333)	(0.0291)	(0.00767)
HH Number of Females, Age 16-59	-0.00262	-0.0447	0.0507***
	(0.00502)	(0.0373)	(0.0113)
HH Number of Females, Age 60+	-0.0137***	-0.107***	0.0298**
	(0.00471)	(0.0215)	(0.0125)
Treatment = o,			
Period = o,			

Constant	0.902***	5.721***	0.438***
	(0.0258)	(0.257)	(0.0576)
Observations	5,781	5,781	5,781
R-squared	0.012	0.024	0.092

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18. Indicators of female labor force participation – also controlling for baseline sector means in outcome (ANCOVA type analysis)

VARIABLES	(1) Work (Past Week Binary)	(2) Days Worked (Past Week)	(3) Work on HH Ag Labor (Binary)
Treatment-Period Interaction	0.0448**	0.320*	0.0514
	(0.0222)	(0.162)	(0.0327)
Treatment	-0.00391	-0.0222	-0.0107
	(0.00790)	(0.0605)	(0.00799)
Period	-0.0274	-0.335**	0.0463*
	(0.0175)	(0.129)	(0.0266)
Age (years)	0.000831	0.00182	-0.000196
	(0.000815)	(0.00654)	(0.00126)
Wealth Index	-0.000358	-0.00678	-0.000326
	(0.00479)	(0.0369)	(0.00860)
Education Category = 2, Primary (Less than 5 years)	0.0220**	0.103	0.00407
	(0.0102)	(0.0678)	(0.0145)
Education Category = 3, Middle (6-8 years)	0.0169	0.206**	0.0279
	(0.0103)	(0.0846)	(0.0204)
Education Category = 4, Higher (9-10 years)	0.0250**	0.307***	-0.0433*
	(0.0114)	(0.0930)	(0.0220)
Education Category = 5, Secondary (11-12 years)	0.0345**	0.467***	-0.165***
	(0.0165)	(0.130)	(0.0343)
Education Category = 6, Undergraduate studies (13-15 years)	0.0353*	0.445***	-0.198***
	(0.0183)	(0.140)	(0.0395)
Education Category = 7, Postgraduate (More than 15 years)	0.0305	0.287	-0.138***
	(0.0293)	(0.218)	(0.0440)
Education Category = 99, Other/Unknown	0.0836***	0.955***	-0.159***
	(0.0157)	(0.137)	(0.0372)
HH Number of Males, Age 0-4	-0.0107**	-0.0526	-0.00365
	(0.00522)	(0.0421)	(0.0106)
HH Number of Males, Age 5-15	0.00849*	0.0498	0.00861
	(0.00454)	(0.0336)	(0.00924)
HH Number of Males, Age 16-59	0.00309	0.0295	0.00223
	(0.00481)	(0.0355)	(0.00913)
HH Number of Males, Age 60+	0.00216	0.0920	0.00534

	(0.00846)	(0.0665)	(0.0176)
HH Number of Females, Age 0-4	-0.0126**	-0.0550	-0.0116
	(0.00530)	(0.0364)	(0.00927)
HH Number of Females, Age 5-15	-0.00174	-0.0172	0.0182**
	(0.00328)	(0.0269)	(0.00709)
HH Number of Females, Age 16-59	-0.000134	-0.00846	0.0425***
	(0.00474)	(0.0373)	(0.0104)
HH Number of Females, Age 60+	-0.0133**	-0.0799***	0.0234**
	(0.00524)	(0.0262)	(0.00955)
Village Average Baseline Work (Past Week, Binary, PCG)	0.507***		
	(0.0354)		
Village Average Baseline Days Worked (Past Week, PCG)		0.614***	
		(0.0380)	
Village Average Baseline Work on HH Ag Labor (Binary, PCG)			0.818***
			(0.0342)
Constant	0.437***	2.178***	0.00885
	(0.0414)	(0.348)	(0.0453)
Observations	5,781	5,781	5,781
R-squared	0.041	0.091	0.265

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VII. Qualitative Evaluation

65. During the course of implementation of the AWCC pilot, a comprehensive needs assessment to guide refresher trainings of AWCC workers and a process evaluation to better understand implementation processes and progress was carried out by MC and IFPRI respectively. A summary of the key findings that emerged from these assessments is outlined below. It puts in perspective and provides important insights into the impact evaluation results.

66. **Mobile Creche Assessment**: Prior to initiating refresher trainings for the creche workers, a team from Mobile Creche undertook a detailed needs assessment by visiting 86 of the 91 functional AWCCs from September to October 2016 and observed their functioning. The team found that almost 75% of the centres were working for 6 to 7 hours per day and that the number of creche children per centre ranged from 5 to 12. Given the space constraints in the existing AWCs, they found that about 10 crèches (6 in Dhar and 4 in Singrauli) were running separately from AWC, although they were 0.5 kms or less from the AWC. In the remaining co-located AWCCs, they found the space per child was largely adequate (in 74 % centres in Dhar and 78% centres in Singrauli). In terms of appropriate utilization of the space, however, they found that approximately only half the creche workers were utilizing the space properly for learning and play. In fact, in 88% AWCCs the children were found sitting together and in only 12% centres (7 in Dhar and 4 in Singrauli), were workers conducting age appropriate activities separately with creche and AWC children. Toys for creche children were also only found in approximately 50% of the centres.

67. An equal, if not greater, concern was the less than adequate availability of clean drinking water and child friendly toilets at the AWCCs. Only about 69% centres in Dhar and 72% in Singrauli had water filters for storing drinking water while the rest stored water in drums or earthen pots; and only 5% centres in both the districts had clean and child friendly toilets within the premises. The creche workers endeavored to follow the practice of making the children wash their hands after using toilets and before/after meals, although even this was not universal (72% in Dhar and 59% in Singrauli).

68. Appropriate nutrition was an additional concern. No separate food was prepared for the younger creche children. The same food that was being provided for the older 3-6 years old at the AWC was made available for the younger children. While on some days this food was appropriate for younger children, (for example on days that *daliya* (porridge) or *khitchdi* and *kadhi chawal* was served), there were days when food like *poori sabzi* was hard and difficult for young children to consume. Also, no separate food was made available for the afternoons. Leftover food from the morning was served if children stayed beyond lunch.

69. **IFPRI's Process Assessment**: The process assessment focused on 5 domains of inquiry – facilities, human resources, information and demand creation, creche operations and utilization of creches by parents. The key research questions included (a) how are creche facilities prepared for children (b) how were the human resource (creche workers and AWWs) prepared for running the AWCC in terms of training, their role and their understanding and perceptions of their role (c) how were children selected to avail the creche facilities and parents informed/engaged (d) how was appropriate nutrition/feeding and child stimulation ensured at the AWCC; and (e) factors that influence the use or non-use of creche facilities by parents.

70. 20 intervention and 20 control centres were randomly selected for the process assessment. Data was collected through interviews with creche workers, AWWs, AW helpers, mothers of 6-24 months old children, Child Development Project Officer, lady supervisors, self-help group members, focused

group discussions with grandmothers and fathers; and through the observation of AWCC facilities and review of creche records.

71. The findings of the process assessment corroborated what emerged from the MC need assessment. There was a shortage of play material, water and sanitation facilities at the AWCCs. Nearly all the creche workers had received the 12-day induction training, a majority of the AWWs had received a shorter version of the training, while AW helpers had not been trained. The three workers (AWW, creche worker and AW helper) worked collaboratively to deliver services. The AWW raised awareness about the creche among mothers, played with the children, told them stories, helped in their care and monitored the weight (growth monitoring); the AW helper maintained the cleanliness of the creche, fed and cleaned the children and helped in picking and dropping them off to their homes.

72. Parents were motivated by the workers to send their child to the creche by highlighting that they would be fed and taken care of when the parents went to work (to work on their fields or elsewhere). Parents too saw benefits in sending their child to the creche so that their child is cared for, kept clean, is provided learning and play opportunities and food when they go to work (work could be household chores or external work); they were also motivated by other parents who were sending their children (trust, social norm and perceived benefit).

73. The centres remained open from 9 am to 3 pm (6 hours). Ten children attended the creche on average; this increased over time as per their records. Even children younger than age 1 were found at the creche. Attendance was affected by season, migration (from October to March) and festivities.

74. Food guidelines for creches were not being followed. Self-help groups supplied food to the centre once or twice a day, and children were given food either twice or thrice a day. No separate food was provided for creche children. As there were no reheating facilities at the centre, food once sent by the SHG was stored and given to the children at room temperature during meal times. On the positive side, the food at the centre was sufficient in terms of quantity and if there was shortage, the AWW prepared the take home ration provided for small children at their homes and fed the children.

75. Utilization of AWCs, specifically immunization of younger children, uptake of take home rations and attendance seemed to have increased. Parents seemed more confident about sending their children as there was a third worker, and they could send their younger children with the older ones. For non-using parents, the primary reasons for not using creche services included inadequate information about the creche, difficulty in dropping and picking up the child from the creche due to distance, irregular functioning and poor quality of services at the AWCC as well as the young age of the child.

76. On the whole, AWWs were positive about the addition of creche services. While some felt their workload had increased due to the addition, others felt it had decreased, as creche workers also care for the older children when the AWW is not present and help the AWW with maintaining registers. The AW helpers largely felt their workload had increased as there were more children to clean and take care of, and pick up and drop.

77. One of the biggest demotivating factors highlighted by the workers was irregular salary payments. The workers also highlighted inadequate materials and non-replacement of old materials, inadequate space, lack of toilets and water and inappropriate food for the younger children as challenges in implementation.

VIII. Conclusion and Recommendations

78. The pilot program offering creches in relatively poor districts achieved modest impacts with modest investments in institutional capacity. In particular, there were small increases in the receipt of ECE services, and equally modest increase in female labor force participation. Notably, there was a significant increase in socio-emotional development in the treatment communities. There was, however, no measurable change in the physical development of target children, and no change in cognitive scores. The negative impacts observed in some other daycare studies mentioned earlier, however, were not observed here.

79. These results point to the potential of early childhood care or creche services, but also highlight the challenges of doing so through ongoing government programs when there are design and resource constraints. There was minimal investment by the government, mainly in terms of creating a platform where smaller children could come; the training of workers was fairly basic. If with this minimal investment even a little positive change is evident in labor and in socio-emotional development and no inhibition of cognitive growth was observed, the results lead to potential lessons as a beginning of a dialogue on the National Creches Scheme for children of working mothers that has been formulated by the government. The limited impacts reported in the intention to treat analysis here do not account for quality of services.

80. In particular, the small scale of the program, relative to the number of ACWs in the state as well as changes in leadership contributed to benign neglect of many aspects of supervision that might have increased program quality. For example, although detailed implementation guidelines for the AWCC model were drafted and translated into Hindi in October 2014 by the DWCD, these were never issued, as the changed leadership believed that these were too detailed and would never be absorbed by the field functionaries. Thus, the model remained only loosely defined.

81. Creche workers were hired by mid-late June 2014 but were not provided any training or orientation on their newly designed role at the AWC for over a year. They therefore started functioning as additional hands to the AWWs. Additional material, including learning and play material was provided to the 100 centres, but no guidance was provided on how to use them for early play, stimulation and learning. Many of them had not received their honorarium as well and were largely unmotivated and directionless.

82. The program was not fully implemented in the target treatment communities and there were no plans for staff replacement and retraining. Thus, the intention to treat reflects a nascent program that may not fully reflect the potential. Findings from the ICDS program offers a perspective that may apply; the limited coverage of ICDS services as documented in the National Family and Health Survey 2005-2006 (NFHS-3) contributed to a divergence between average treatment effects on nutrition and intention to treat[29]. In other words, while the ICDS program was fairly effective for those beneficiaries that it reached, but as it didn't reach many households, the population wide impact (the intention to treat estimate) was diluted.

[29] Kandpal E. Beyond average treatment effects: Distribution of child nutrition outcomes and program placement in India's ICDS. *World Dev*. 2011;39(8):1410-1421. Jain M. India's struggle against malnutrition-is the ICDS program the answer? *World Dev*. 2015;67:72-89.

83. "Modest impacts with modest investments" is, however, not a firm basis on which to extrapolate to the impacts of more focused investments. Nor does this observation directly inform the dialogue on the National Creche Scheme. But the current pilot can be augmented by global studies that indicate the impact of day care on cognitive development is, not surprisingly, a function of quality[30]. Going further, the aspect of quality that appears to be most important and most cost effective is not infrastructure but the quality of activities and program structure. Routines and interactions between children and caregivers are particularly important for children in day care; young children need less in the way of structured content than older children but benefit especially from individualized care[31].

84. Given that, two observations from the Madhya Pradesh experience hint at prerequisites for more effective day care. First, few if any ICDS administrator have experience in managing early child education. To add a new task, that is creche services, to the ICDS program without extensive training, not only for the frontline functionaries but for the supervisors and officials from the sector supervisors to the block, district and state officials, is a failure to acknowledge that nurturing care[32] requires a different and more holistic framework focused on quality of care and not simply service provision which currently underpins the DWCD.

85. Secondly, along with creche workers, parents, too, tend to believe that children grow, rather than that they are nurtured; they know that children need to be fed and protected but generally do not recognize the need for stimulation. To make the most of any day care program, centers should be linked with parental enrichment communication. This can be provided on an outreach footing and not confined to the minority of parents whose children attend centers on a regular basis. Such programs go beyond behavioral change communication for health and nutrition and often focus on the needs of a different age cohort[33]. This was clearly demonstrated through a recent study in Odisha which found that individual home visits as well as group sessions with mothers and their children on early child stimulation, led to significant improvement on cognitive, language and fine motor development of the target children[34]. Best practices have been reviewed[35] and while they vary in impact and context, at least one generalization is clear: one does not get from current ICDS priorities to nurturing care as an afterthought. Again, this goes to training and program design as much as budgetary outlays.

86. Accounting for follow-up training, appropriate meal provision, play material for early stimulation and learning, improved infrastructure and other implementation issues documented in the qualitative evaluation may point to ways to improve the services provided, should day care stay on the agenda of state governments. Scaling up will not be easy; at the least it requires a commitment to investing in training – but a model exists that can be implemented if that commitment is forthcoming. The reason for a cautious optimism also stems in part from the overall trends observed in the three-year period between rounds, particularly in regard to nutrition as well as caregiver education. For decades these indicators of human capital development had been fairly stagnant. Thus, progress in these indicators – although by no means yet at a satisfactory level – can reinforce the goals of the 2012 restructuring towards a broader set of objectives for the ICDS and a greater focus on children under 3 years of age.

---

[30] Barros, R, Carvalho, M, Franco, S, Mendonca, R and Rosalem, A, 2011. *A Short-Term Cost Effectiveness Evaluation of Better Quality Daycare Centres*. Available at: IDB-WP-239

[31] Bóo, Florencia López, María Caridad Araujo, and Romina Tomé. 2016. .*How is Child Care Quality Measured?: A Toolkit*. Inter-American Development Bank.

[32] Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, et al. 2017. Nurturing care: promoting early childhood development. *Lancet 389*(10064), pp.91-102.

[33] They may, however, also have favorable impact on nutrition. See, for example, Aulo Gelli, Amy Margolies, Marco Santacroce, Natalie Roschnik, Aisha Twalibu, Mangani Katundu , Helen Moestue, Harold Alderman and Marie Ruel. Using a community-based early childhood development center as a platform to promote production and consumption diversity increases children's dietary intake and reduces stunting in Malawi: A cluster randomized trial. Journal of Nutrition. Forthcoming.

[34] Meghir et. al (Draft report, 2017) Early Childhood Development for the Poor: Impacting at Scale, Odisha, India

[35] Britto et al. op cit.

87. Finally, it is important to note that focusing on ECD particularly in the ages between the most vulnerable period for nutritional shocks and the start of preschool will require monitorable indicators that can be tracked in national or even state level surveys, ideally within surveys that are regularly implemented. This is currently not the case especially in India. Despite the reputation of the BSID for research and for clinical studies, it is not suitable for this task. Although the experience with the CREDI cannot resolve which set of instruments designed to utilize caregiver responses is the most suitable, it shows that caregivers can provide assessments in keeping with the more interactive (and thus more time consuming and training intensive) BSID approach.

Annex 1. Validation of the CREDI relative to the BSID

The CREDI assesses three domains of development – motors skills, cognitive, and socioemotional. These domains can be subdivided into subdomains; for example, motor skills are either fine or gross motor development and the domain of cognitive skills includes subdomains of receptive, and expressive language as well as preacademic knowledge, and problem solving. The validation exercise concentrated only on the cognitive domain and the fine motor subdomain; the BSID was not employed for an assessment of socioemotional development.[36]

Following a similar validation exercise in Tanzania which assessed the cognitive domain inclusive of language,[37] we first verified that the instrument was suited to the age of the index children. This is especially relevant since the instrument is designed for children up to 36 months, although a portion of the index children in our sample were older than that. As many clusters sampled had relatively few children in the 24-36 month age bracket, the requirement that the index child was old enough to have spent time in day care required that 24.7 % of the index children were between 36 and 42 months. However, only 1 child in this age bracket was in the subset chosen for the comparison of the CREDI and the BSID.

As indicated in Annex Figure 1, there were ceiling effects for only a modest proportion of questions answered regarding those children older than 36 months and very few for other children. Moreover, as expected, as the questions refer to more challenging behaviors (that is, descending on the figure), the proportion of caregivers of children in all ages that respond positively declines. Annex Figure 2 graphs the correlation of the unadjusted CREDI and a local polynomial of BSID cognitive scores and Annex Figure 3 does the same for the motor scores.

The similar pattern in scores observed in figures 2 and 3, however, could be driven by the fact that both indicators are strongly associated by common factors such as maternal education and wealth as well as the expected improvement of unadjusted scores as a child ages. Indeed, annex table 1 shows strong statistical relationships of a number of covariates for both the full CREDI sample and the subsample used for comparison with the BSID as well as for the BSID results (the coefficients themselves are not directly comparable given different units). The only coefficient that might be considered unexpected is the negative association of gender, defined as male=1, in the CREDI something not observed in the cited Tanzania study.[38]

Annex table 2 uses the residuals of these regressions to assess the unique information on child development in these measures. The residuals from the regressions of the overall cognitive domain are significantly correlated with the BSID scores for language and for cognitive as well as these subdomains combined. These are all significant at $p< 0.1$ as is the correlation of residuals for motor development even though the CREDI score combines gross and fine motor and the BSID employed here does not. Moreover, the table shows that there are not substantial distinctions in the correlations when the sample is divided into education subgroups. This addresses a concern that less educated caregivers might report the activities of

[36] Previous work indicates that the BSID does not perform well relative to other measures (and vice versa). See, Rubio-Codina, Marta, M. Caridad Araujo, Orazio Attanasio, Pablo Muñoz, and Sally Grantham-McGregor. "Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies." *PloS one*11, no. 8 (2016): e0160962.

[37] McCoy DC, et al. op cit.

[38] The wealth coefficients are smaller and not statistically significant for children older than 36 months in the regressions for motor or socioemotional, possibly indicating ceiling effects.

their child in a manner that was less in conformity with independent observations than are observations by those with more education.

As neither of these domains was affected by the treatment using the full sample, the comparison provides little insight about whether these two measures are equally capable of capturing small effects of an intervention. The t-statistics for the coefficient of treatment for BSID cognitive was only 0.3. Thus, even multiplying the statistic by the square root of 3 as a crude indicator of what might have been the t statistic with a larger sample of 3,000 children rather than 1,000 does not provide a reason to expect that the absence of a significant impact on the BSID cognitive results reflects low power. This is also the case for motor skills. As the study did not assess the BSID as a tool to assess socio-emotional development no comparison can be made on that dimension.

Annex Table 1. Covariates of CREDI and BSID Scores

	CREDI cognitive Full sample	CREDI cognitive Subsample	BSID cognitive	BSID language and cognitive	CREDI motor Full sample	CREDI motor Subsample	BSID motor	CREDI socioemotional Full sample
Index Child Age (in months)	0.625***	0.619***	0.768***	1.829***	0.280***	0.305***	0.507***	0.281***
	(0.019)	(0.059)	(0.047)	(0.120)	(0.008)	(0.026)	(0.033)	(0.011)
Wealth Index	1.043***	1.034***	0.741***	3.007***	0.241***	0.272***	0.518***	0.231***
	(0.123)	(0.206)	(0.166)	(0.423)	(0.053)	(0.090)	(0.118)	(0.069)
Sex	-0.964***	-1.347***	0.923***	0.512	-0.305***	-0.525***	0.147	-0.465***
	(0.243)	(0.412)	(0.328)	(0.833)	(0.105)	(0.181)	(0.232)	(0.136)
Primary (Less than 5 years)	1.253***	1.181**	1.707***	5.439***	0.321**	0.380	0.859***	0.533***
	(0.332)	(0.559)	(0.445)	(1.131)	(0.143)	(0.245)	(0.315)	(0.186)
Middle (6-8 years)	1.258***	1.815***	1.365***	4.549***	0.599***	0.874***	1.088***	0.827***
	(0.342)	(0.585)	(0.467)	(1.186)	(0.147)	(0.256)	(0.330)	(0.192)
Higher (9-10 years)	1.639***	0.672	1.342**	3.396**	0.715***	0.765**	0.824**	1.094***
	(0.419)	(0.700)	(0.557)	(1.416)	(0.180)	(0.307)	(0.394)	(0.234)
Secondary (11-12 years)	2.910***	3.535***	1.659**	6.442***	1.128***	1.317***	1.247**	1.416***
	(0.532)	(0.884)	(0.706)	(1.792)	(0.229)	(0.388)	(0.499)	(0.297)
Undergraduate studies (13-15 years)	4.528***	2.102*	3.501***	10.741***	1.443***	1.524***	1.483**	2.101***
	(0.718)	(1.188)	(0.947)	(2.402)	(0.309)	(0.521)	(0.669)	(0.402)
Postgraduate (More than 15 years)	5.936***	5.885***	4.836***	12.942***	1.718***	2.062***	2.712***	2.919***

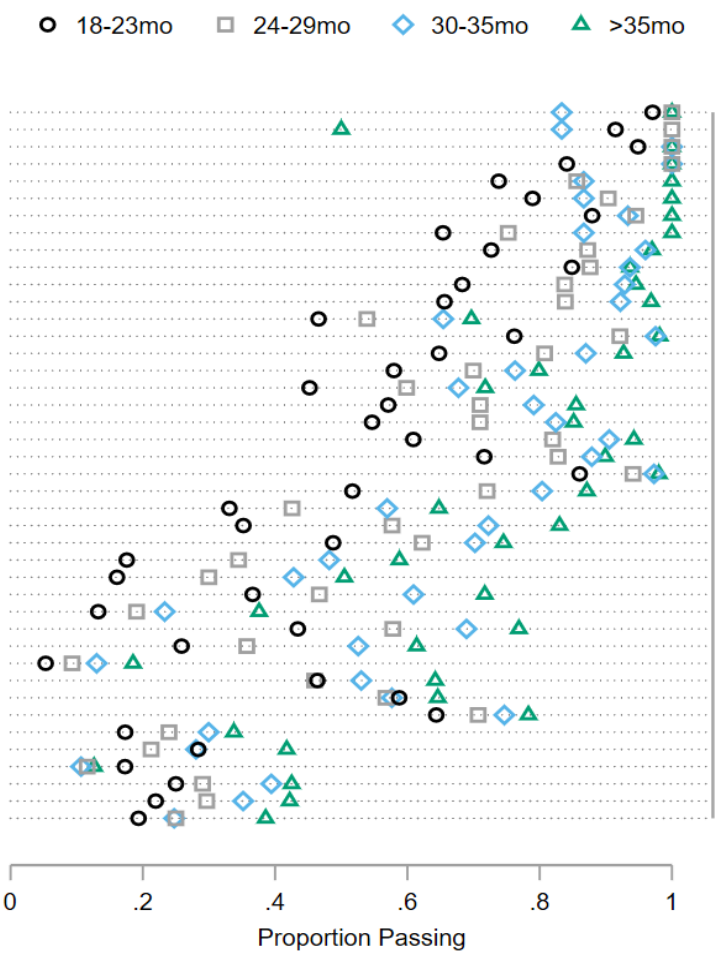
	(0.938)	(1.504)	(1.204)	(3.054)	(0.404)	(0.660)	(0.850)	(0.525)
Was the test observed?			-0.072	-0.193			-0.097	
			(0.087)	(0.220)			(0.061)	
Breaks			-2.588*	-7.790**			-1.314	
			(1.527)	(3.875)			(1.078)	
Observer dummy			Yes	Yes			Yes	
Constant	0.673	1.783	37.227***	58.062***	9.206***	8.814***	24.833***	4.645***
	(0.607)	(1.744)	(1.578)	(4.010)	(0.261)	(0.765)	(1.116)	(0.340)
Observations	2823	1023	1023	1018	2823	1023	1018	2822
R^2	0.322	0.157	0.310	0.320	0.318	0.160	0.283	0.229

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Annex Table 2. Correlations of BSID Residuals With CREDI Residuals

	Mother's education						Full sample	
	No education		Primary education		Higher than primary education			
Residuals	CREDI cog	CREDI motor	CREDI cog	CREDI motor	CREDI cog	CREDI motor	CREDI cog	CREDI motor
BSID Cog+LE+LR	0.3573		0.3554		0.3519		0.3512	
BSID LE+LR	0.3269		0.3317		0.3338		0.3190	
BSID Cog	0.2830		0.2727		0.2788		0.2678	
BSID LR	0.2914		0.2967		0.2982		0.2800	
BSID LE	0.3057		0.3101		0.3157		0.3004	
BSID Motor		0.1865		0.1800		0.1758		0.1581

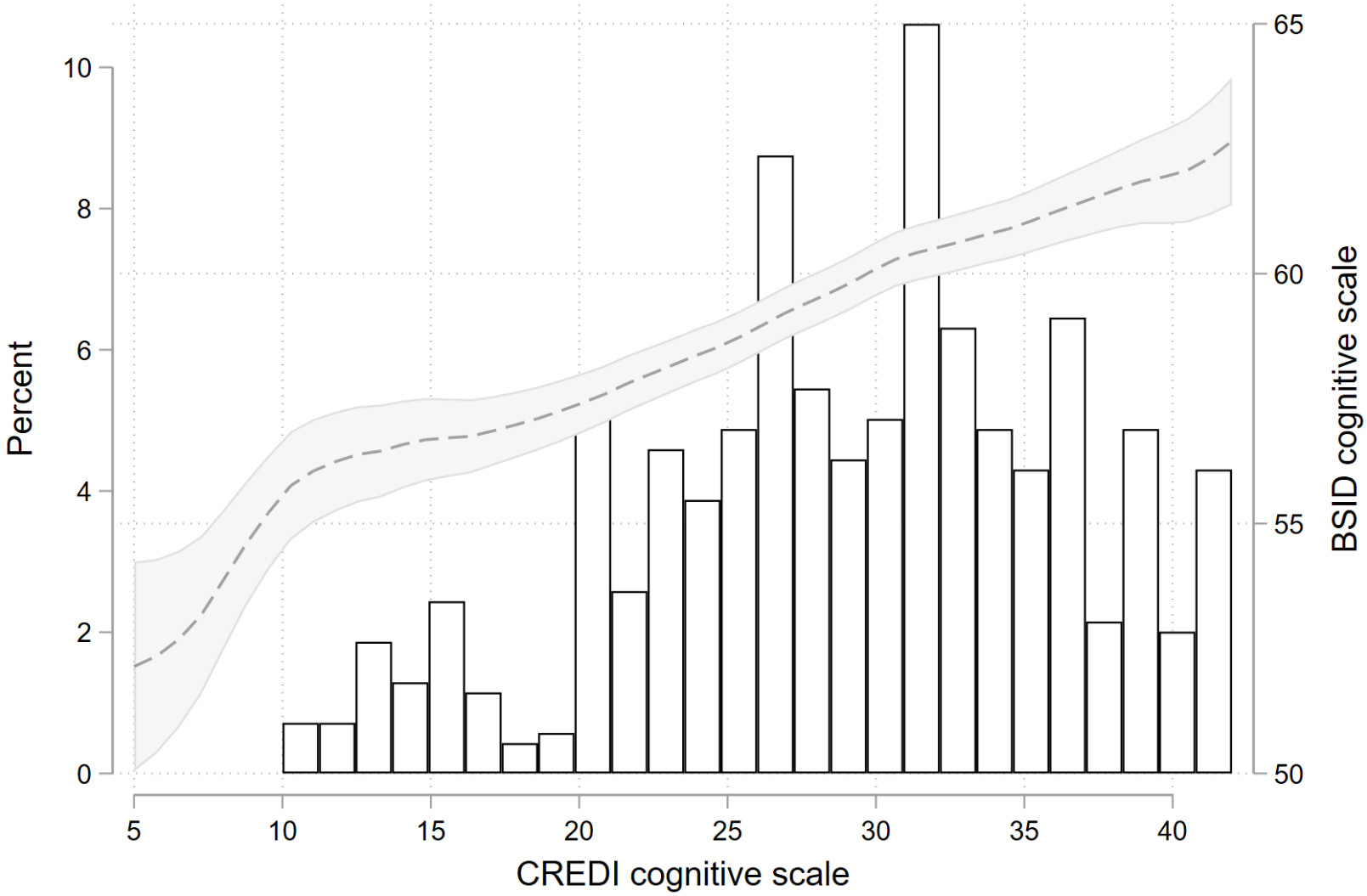
Annex Figure 1. CREDI: Ceiling Effects on Cognitive Questions



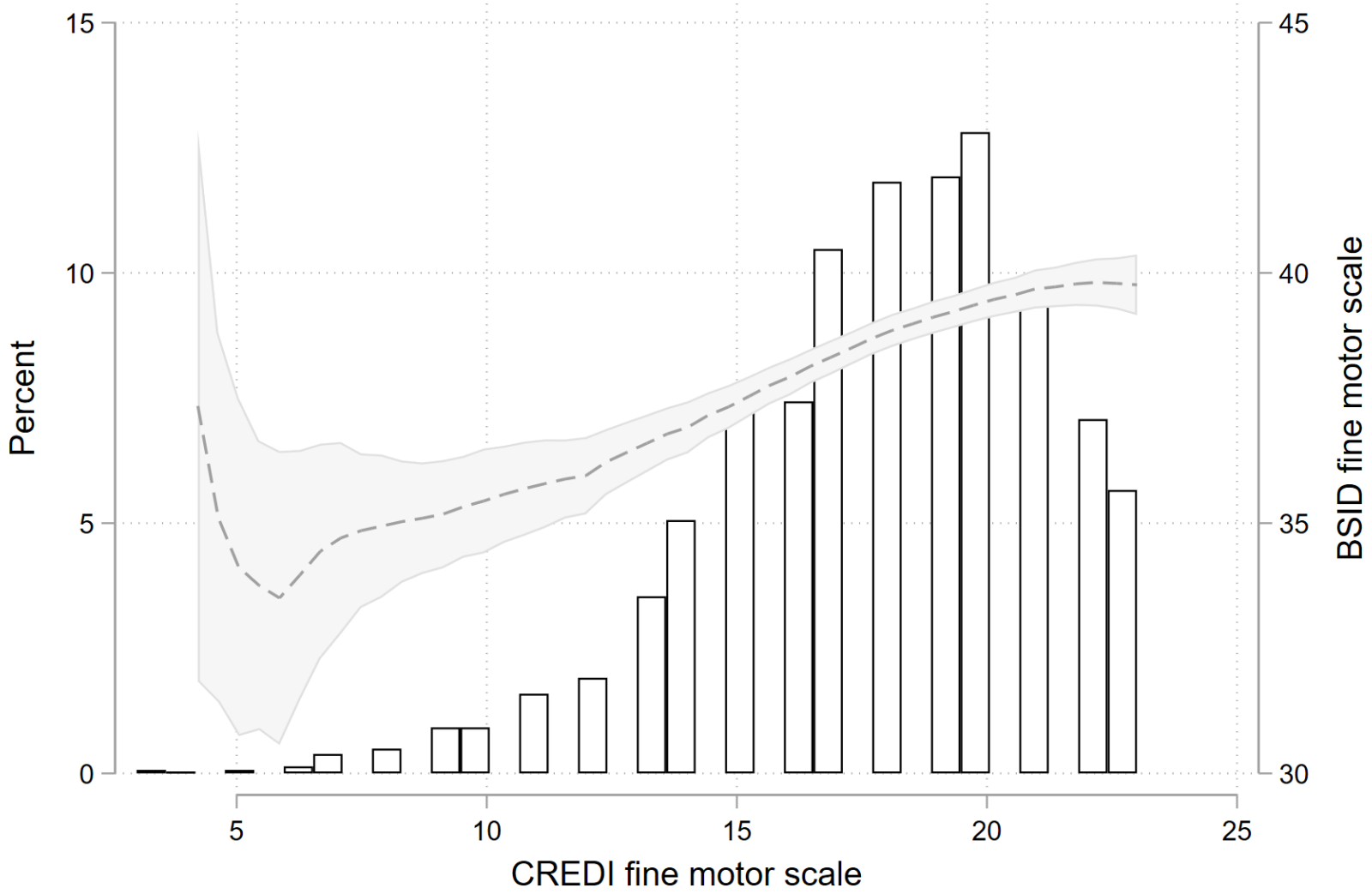
Can the child use gestures to indicate what he/she wants (e.g., put arms up to indicate that he/she wants to be held, or point to water)?
Can the child say one or more words (e.g., names like "Mama" or "ba" for "ball")?
Can the child follow simple directions (e.g., "Stand up" or "Come here")?
Does the child watch what other children do and try to copy them?
Does the child try to repeat sounds or words said by other people?
Can the child figure out how to turn a spoon or object if you give it to him/her the wrong way around?
Can the child point to a person or object when asked (e.g., "Where is mama?" or "Where is the ball?")?
Does the child imitate animal or other sounds (e.g., "vroom" for a car, "moo" for a cow)?
Can the child answer simple questions (e.g., "Do you want water?") by saying "yes" or "no", rather than nodding?
Does the child imitate others' behaviors (e.g., washing hands or dishes)?
Can the child follow orders or instructions that have more than one part (e.g., "Go get water and go to bed")?
Can the child say five or more separate words (e.g., names like "Mama" or objects like "ball")?
Does the child listen to someone telling a story with interest?
Can the child ask for something (e.g., food, water) by name when he/she wants it?
Can the child correctly name at least one family member other than mom and dad (e.g., name of brother, sister, aunt, uncle)?
Does the child play by pretending objects are something else (e.g., imagining a bottle is a doll, a stone is a car, or a spoon is an airplane)?
Does the child show curiosity to learn new things (e.g., by asking questions or exploring a new area)?
Does the child know the names of at least two body parts (e.g., arm, eye, or nose)?
If you show the child an object he/she knows well (e.g., a cup or animal), can he/she consistently name it?
Can the child speak using short sentences of two words that go together (e.g., "Mama go" or "Dada eat")?
Can the child use a tool (e.g., a stick or spoon) to reach objects that are far away?
Can the child indicate when he/she needs to go to the toilet?
Can the child say ten or more separate words (e.g., names like "Mama" or objects like "ball")?
Can the child sing a short song or repeat parts of a rhyme from memory by him/herself?
Can the child speak using sentences of three or more words that go together (e.g., "I want water" or "The house is big")?
Can the child whisper?
Can the child correctly ask questions using any of the words "what," "which," "where," or "who"?
Can the child correctly use any of the words "I," "you," "she," or "he" (e.g., "I go to store," or "He eats rice")?
Does the child pronounce most of his/her words correctly?
Can the child count up to five objects (e.g., fingers, people)?
Does the child ask about familiar people other than parents when they are not there (e.g., "Where is uncle/aunty?")?
If you show the child two objects or people of different size, can he/she tell you which one is the big one and which is the small one?
Can the child identify at least one color (e.g., red, blue, yellow)?
Does the child regularly use describing words such as "fast," "short," "hot," "fat," or "beautiful" correctly?
If you point to an object, can the child correctly use the words "on," "in," or "under" to describe where it is?
Can the child explain in words what common objects like a cup or chair are used for?
Does the child ask "why" questions (e.g., "Why are you tall?")?
If you ask the child to give you three objects (e.g., stones, beans), does the child give you the correct amount?
Can the child say what others like or dislike (e.g., "Mama doesn't like fruit," "Papa likes cricket")?
Can the child talk about things that will happen in the future using correct language (e.g., "Tomorrow he will attend school")?
Can the child talk about things that have happened in the past using correct language (e.g., "Yesterday I played with my friend")?
Does the child know the names of any letters (e.g., A, B, C)?

* indicates item that was reverse coded

Annex Figure 2. CREDI responses plotted against BSID: Cognitive.



Annex Figure 3. CREDI responses plotted against BSID: Motor.



Annex 2. Baseline Balance Tables

The following tables present baseline balance results between the treatment and control groups across two districts (Dhar and Singrauli) in Madhya Pradesh. Treatment and control mean or percent, mean differences between treatment and control, and p-values adjusted for clustering at the community level are presented along with normalized mean differences. Normalized mean differences suggest an imbalance between treatment and control for any value over 0.2.[39]

Annex Table 3 Baseline Sample Demographics

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Caregiver Female	0.99	1.00	-0.0093	0.320	-0.358
Ever attended school	0.49	0.49	-0.0022	1.000	-0.003
General education level	4.11	4.01	0.1000	0.534	0.022
Religion (Hindu)	0.97	0.95	0.0264	0.207	0.095
Child's Age (in months)	30.57	30.43	0.1400	0.759	0.009
Child Female	0.50	0.51	-0.0061	0.753	-0.009

Baseline demographics were balanced between the treatment and control groups –

p-values adjusted for community level clustering were not significant, indicating no significant differences between the treatment and control groups. The normalized mean difference on the caregiver female variable suggests an imbalance, but nearly all caregivers were female in both arms.

Annex Table 4 Child Development Indicators

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Child can identify 10 letters	0.34	0.28	0.0583	0.086	0.090

[39] Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47, no. 1: 5-86

Too sick to play	0.79	0.82	-0.0265	0.389	-0.047
Get along with others	0.94	0.95	-0.0072	0.444	-0.022
Kick, bite, hit others	0.75	0.76	-0.0154	0.509	-0.025
Gets distracted easily	0.97	0.98	-0.0039	0.586	-0.017
Attends AWC	0.42	0.39	0.0316	0.405	0.045
Received counseling at AWC	0.46	0.49	-0.0332	0.450	-0.047
Pay fee for AWC	0.82	0.75	0.0648	0.486	0.111
Meals at AWC	0.79	0.81	-0.0227	0.535	-0.040
Take home rations from AWC	0.75	0.76	-0.0173	0.619	-0.028
Member of mothers committee (AWC)	0.01	0.01	0.0002	0.961	0.002

AWC utilization indicators were balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering were not significant. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.

Annex Table 5 Nutrition Indicators

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Stunted	0.61	0.64	-0.0357	0.318	-0.052
Underweight	0.57	0.56	0.0038	0.922	0.005
Wasted	0.21	0.23	-0.0156	0.559	-0.027
Child ever breastfed	0.99	0.99	-0.0001	0.969	-0.001
Immediate breastfeeding	0.25	0.20	0.0424	0.087	0.072
Months exclusively breastfed	1.96	2.08	-0.1200	0.735	-0.012
Number of days in last 7 child attended AWC	2.88	2.86	0.0200	0.877	0.007
How old when fed water (months)	1.56	1.58	-0.0200	0.943	-0.007
How old when semi-solid food	6.21	6.21	0	0.996	0.000
How old when solid food	9.12	9.14	-0.0200	0.850	-0.005
How many times child ate food in 24 hours	3.26	3.19	0.0720	0.463	0.057

Child nutrition indicators were balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering were not significant. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.

Annex Table 6 Immunization Indicators

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Ever immunized	0.96	0.94	0.0214	0.306	0.068
BCG	0.93	0.91	0.0214	0.363	0.055
Polio 0	0.90	0.88	0.0215	0.327	0.048
Polio 1	0.93	0.91	0.0225	0.276	0.060
Polio 2	0.90	0.88	0.0259	0.350	0.058
Polio 3	0.85	0.81	0.0359	0.310	0.068
DTP 1	0.89	0.86	0.0353	0.270	0.076
DTP 2	0.85	0.80	0.0443	0.258	0.083
DTP 3	0.79	0.76	0.0307	0.484	0.052
Measles	0.79	0.73	0.0610	0.147	0.101
Vitamin A	0.74	0.68	0.0636	0.159	0.099
Hep B 0	0.68	0.64	0.0418	0.327	0.062
Hep B 1	0.74	0.67	0.0692	0.194	0.107
Hep B 2	0.67	0.61	0.0636	0.266	0.094
Hep B 3	0.62	0.58	0.0414	0.466	0.060
Deworming	0.34	0.31	0.0338	0.403	0.051

Child immunization indicators were balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering were not significant. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.

Annex Table 7 Cognitive scores for ages and stages

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Communication	40.23	39.01	1.2200	0.432	0.052
Gross motor skills	47.39	46.76	0.6300	0.572	0.031
Fine motor skills	35.64	33.91	1.7300	0.271	0.067
Problem solving skills	41.02	40.01	1.0100	0.434	0.049
Personal social skills	39.49	39.19	0.3000	0.803	0.014

Cognitive scores for the ages and stages module were also balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering were not significant. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.

Annex Table 8 AWW visits and characteristics

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Has AWW visited home	0.49	0.43	0.0590	0.189	0.084
How many times AWW visited home in last 3 months	2.22	2.44	-0.2230	0.089	-0.052
Heard about Bal Chaupal	0.01	0.01	-0.0005	0.884	-0.004
Heard about NRC	0.26	0.24	0.0222	0.469	0.036
Child referred to NRC	0.12	0.13	-0.0087	0.751	-0.018
Heard about Sneha Shivir	0.03	0.03	0.0013	0.906	0.005
AWW treats with respect	0.83	0.82	0.0080	0.809	0.015
AWW knowledgeable during pregnancy	0.65	0.64	0.0101	0.805	0.015
AWW knowledgeable about babies	0.59	0.57	0.0201	0.684	0.029
AWW directs to health service providers	0.50	0.49	0.0046	0.931	0.007
AWW is caring and friendly	0.76	0.71	0.0470	0.308	0.075
AWW engages children in learning and play	0.66	0.60	0.0655	0.206	0.096
AWW available when needed	0.69	0.62	0.0697	0.196	0.104

Variables identifying AWW visits and characteristics of AWWs were balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering were not significant. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.

Annex Table 9 SES Indicators

Characteristics at BASELINE	Treatment Percentage or Unadjusted Mean	Control Percentage or Unadjusted Mean	Mean Difference	p-value adjusted for clustering at community level	normalized mean difference
Own land	0.80	0.81	-0.0032	0.945	-0.006
Primary for lighting (kerosene)	0.26	0.32	-0.0686	0.255	-0.107
Electric fan	0.56	0.51	0.0538	0.465	0.076
Radio/transistor	0.05	0.06	-0.0098	0.287	-0.031
B&W TV	0.06	0.03	0.0233	0.105	0.079
Color TV	0.34	0.32	0.0230	0.729	0.035
Sewing Machine	0.15	0.15	0.0086	0.822	0.017
Mobile Telephone	0.81	0.79	0.0212	0.549	0.038
Refrigerator	0.10	0.11	-0.0067	0.844	-0.016
Bicycle	0.41	0.38	0.0300	0.503	0.043
Motorcycle/scooter	0.35	0.34	0.0105	0.822	0.016
Motor car	0.06	0.04	0.0120	0.401	0.039
Any animals	0.71	0.72	-0.0148	0.805	-0.023

Household characteristics of SES indicators were balanced at baseline between the treatment and control groups. P-values adjusted for community level clustering are not significant, indicating no difference in household characteristics and SES between the treatment and control groups. Additionally, normalized mean differences were below the .2 threshold that would suggest an imbalance.