Common Prosperity in Rural China Begins at 0-3 Years Old

Paper for the 2024 Allied Social Science Associations
Session for the Association of Comparative Economics Society

Dorien Emmers
Assistant Professor
KU Leuven

Scott Rozelle
Senior Fellow/Director
Rural Education Action Program
Stanford University

Common Prosperity in Rural China Begins at 0-3 Years Old

Despite rapid economic growth in China since 1978, rural-urban inequality has widened. High levels of socioeconomic inequality can have profound implications for child development and lifelong educational equity. Using a dataset containing early childhood development (ECD) outcomes of 0- to 3-year-olds (N = 9,053) from study sites in Eastern, Central, and Western China, the study finds that the risks of cognitive, language, and motor delay are, respectively, 43.2, 18.3, and 20.7 percentage points higher in rural study sites than in urban Shanghai (ps < .01). Impact evaluation of cluster-randomized experiments shows that parental training (focusing on child psychosocial stimulation and caregiver-child interaction) can improve parenting beliefs and practices (or investments) and ECD outcomes of disadvantaged rural children (p < .01). Such programs can play an important role in advancing progress toward more social equality and economic equity, the stated goals of China's "Common Prosperity" policy.

Keywords: early childhood development, Common Prosperity, rural-urban inequality, intergenerational transmission of disadvantage, parental training, parenting beliefs and practices

1. Introduction

The introduction of the 1978 reforms in China led to four decades of sustained economic growth (Junsen Zhang 2021) and steady improvement in educational attainment (Li et al. 2017). During this same period, however, rural-urban inequality increased sharply (Rozelle and Hell 2020). As of 2016, the average urban disposable income in China was, on average, 170% higher than the average disposable income of the rural population (Junsen Zhang 2021). Furthermore, the Gini coefficient in China increased from about 0.30 in 1978 to 0.47 in 2016 (Junsen Zhang

2021). Luo and Xie (2020) and Yang and Gan (2020) show that, since the 2010s, the level of income inequality in China is among the highest in the world.

While trends in income inequality in China have long been a topic of academic scrutiny and debate, few studies have investigated the implications of socioeconomic inequalities for child development and lifelong educational equity in China. Research findings from other countries have shown that children in lower socioeconomic status (SES) families are at a higher risk of developmental and educational delays than are children from higher-SES families (Bornstein and Bradley 2002; Falk et al. 2021; Heckman and Mosso 2014; Schady et al. 2015), which may contribute to the intergenerational transmission of disadvantage (Qin, Wang, and Zhuang 2016). For example, Schady et al. (2015) finds that gaps in the cognitive ability between children from low-SES and high-SES families in Chile, Colombia, Ecuador, Nicaragua, and Peru emerge by age 3 and do not change substantively during later stages of childhood. Fernald et al. (2012) and Rubio-Codina et al. (2015) extend this research by showing that, when comparing high- and low-SES families, gaps in the cognitive skills of young children in low- and middle-income countries (LMICs) emerge as early as at 7 months of age and (in the absence of intervention) tend to persist as the children age throughout their childhood.

It is well established in the literature that the first years of life are foundational for human development (Currie and Almond 2011; Heckman 2006) and that SES differences during this sensitive and critical stage for human development may have lifelong consequences. Delays in early childhood development (ECD) have been associated with lasting impacts on a range of adult human capital outcomes, including lower cognitive functioning (Walker et al. 2011), lower incomes (Goodman, Joyce, and Smith 2011), increased risk of mental health problems (Walker et al. 2022), and decreased labor productivity (Gertler et al. 2014). Research from the United

States has shown that SES gaps in cognitive and non-cognitive abilities that emerge during early childhood tend to remain constant into adulthood (i.e., the gaps remain high or low across time—Heckman 2008).

The human development literature highlights the role of the family environment as one of the key mechanisms behind the impact of family SES on ECD outcomes (Young and Hannum 2018). This means that SES inequalities in children's outcomes originating in early childhood are to a large extent determined by inequalities in how stimulating, nurturing, and safe the children's home environments are (Attanasio, Cattan, and Meghir 2022). Caregiver investments in a cognitively stimulating home environment that is sensitive to the needs of the child have been identified as key inputs for the cognitive and noncognitive development of young children. Furthermore, Attanasio, Cattan, and Meghir (2022) argue that, whereas investments in the home environment are strongly correlated with household SES, financial resources do not play a large role for children's developmental opportunities per se. Empirical evidence has shown that the SES gap in caregiver investments in the family environment is, to a large extent, driven by differences in parenting beliefs, knowledge, and skills (Cunha 2015). Hence, providing low-SES households with advice, guidance, or training on effective parenting strategies has been shown to improve equity in child development and reduce intergenerational transmission of disadvantage.

To develop a scalable remediation strategy that can offset the intergenerational transmission of disadvantage in diverse cultural and socioeconomic settings,[1] evidence on the effectiveness and efficacy of promising policy interventions and their underlying mechanisms

[1] Caregiver-child interaction and communication appear to be universal aspects of effective parenting and healthy child development. However, it is also known that parents and children can interact and communicate in certain ways that are adapted to and consistent with their cultural context (Bornstein 2012).

are needed. Pioneering studies from the 1970s and 1980s, with long-term follow-ups, including the Jamaican Nutrition and Cognitive Stimulation Program, found long-lasting positive effects of small-scale parenting interventions (with a focus on child psychosocial stimulation and caregiver-child interaction) on a range of adult human capital outcomes, including lifetime educational attainment (Walker et al. 2022). Such promising results incentivized researchers and policymakers in the 2010s, in particular, after the introduction of Target 4.2 of the Sustainable Development Goals, to "ensure that all girls and boys have access to quality early childhood development" by 2030 and to replicate these types of intervention studies in resource-poor settings around the world (e.g., the studies of Attanasio et al. 2014 in rural Colombia; Singla, Kumbakumba, and Aboud 2015 in Uganda; Yousafzai et al. 2014 in Pakistan). Over the past several decades, evidence from randomized controlled trials in LMICs in Africa, Asia, and Latin-America have confirmed that parental training focusing on interactive caregiver-child activities can improve ECD outcomes in resource-poor settings (Emmers et al. 2022). Therefore, global policymakers no longer debate the effectiveness of parenting programs for improving the home environment and ECD outcomes. Instead, the main focus of the debates in these countries is how quality child and family services can be delivered in cost-effective, sustainable, and at scale ways.

Evidence on the underlying mechanisms of treatment impacts can inform policymakers about ways to design cost-effective policies that can support the development of disadvantaged children at scale and advance progress toward more social equality and economic equity, the stated goals of China's "Common Prosperity" policy. The international literature has shown that the impacts of parental training programs on ECD outcomes are partially mediated by improvements in parental investments in the quality of the early learning and the home environment (Amaro Da Costa Luz Carneiro et al. 2019; Attanasio et al. 2020). In addition, studies have shown that changes in parenting beliefs—the perceptions, attitudes, knowledge, ideas, and values of parents with regard to child rearing and child development—can positively mediate treatment impacts of parental training programs on ECD outcomes (Amaro Da Costa Luz Carneiro et al. 2019; Attanasio, Cunha, and Jervis 2019). Changes in parenting beliefs can affect ECD outcomes directly (e.g., via changes in the responsiveness of caregivers) and indirectly via changes in the nature of parental investment (e.g., improved parenting beliefs and knowledge may induce more investment in the time that parents spend on interactive reading, story-telling, and play activities with young children; Attanasio et al. 2019).

For the case of China, empirical evidence has demonstrated that skill gaps between rural and urban children can already be observed starting from an early age onwards when rural children start to lag behind urban peers in their development (e.g., Wang et al. 2019; Yue et al. 2017; Zhang et al. 2021). An important limitation of the evidence base, however, is that earlier population-level ECD assessments have focused mainly on children between 3 and 6 years of age, whereas population means for children younger than age 3 remain to be assessed (Clark et al. 2020). A systematic review and meta-analysis of studies that utilize empirical primary data that were collected in rural study sites across Central and Western China illustrated that the risk of early cognitive and language delay for rural children younger than 5 years is high: 45% and 46%, respectively (Emmers et al. 2021). Each of the studies included in the meta-analysis used standardized instruments, such as the Bayley Scales of Infant and Toddler Development (Bayley 2006) to assess ECD outcomes. In line with the guidelines of the Bayley Scales (the gold standard of ECD assessment) and definitions of the Global Research on Developmental Disabilities Collaborators (2018), developmental delay was defined as a development score of 1 or more standard deviations (SDs) below the mean of a reference population whose developmental trajectory is expected to be normal (i.e., a population of children in healthy/developed regions and did not include prematurely born, severely malnourished, or severely diseased children). Note that observational assessments, such as the Bayley Scales, are unlikely to be sensitive to SES differences. Furthermore, for the standardized, caregiver-reported ECD measures used in the studies, evidence has suggested that parents across different SES groups can provide accurate measurements of the development of children as validated by direct and observational assessments (Fernald et al. 2017). Emmers et al.'s meta-analysis, unfortunately, compares observational assessments of ECD outcomes of children under 3 years of age in Central and Western China (as well as migrant communities) only and does not examine the difference in ECD outcomes of samples either across rural and urban study sites or between inland rural China (i.e., rural samples from less-developed areas in Central and Western China) and Eastern China (i.e., historically, the more developed segment of China's rural population; Zhang 2021). Therefore, the current study sets out to construct a unique dataset containing Bayley test scores from birth cohorts in urban areas as well as rural areas of Eastern, Central, and Western China in order to assess whether ECD outcomes of rural children (from Eastern, Central, and Western regions) are significantly different from their urban peers before they reach the age of 3.

The high prevalence of developmental delay in rural areas of China may be explained, at least partly, by the comparatively high prevalence of low-SES families in rural areas. SES differences in China have empirically been linked to differences in educational outcomes (Young and Hannum 2018). Zhang et al. (2021), for example, used large-scale field data containing early Human Capability Index (eHCI) scores of 63,559 children between 3 to 6 years of age to show that children in low-SES families and those in families with lower maternal educational attainment are, on average, 12 and 8 percentage points less likely to be developmentally on track than are children in higher-SES families, respectively. This large sample of 3- to 6-year-olds was selected using a carefully designed stratified randomized sampling strategy across four provinces in China: Shanghai (i.e., a megacity), Zhejiang (i.e., representing the most-developed region of the country in Eastern China), and Gansu and Yunnan (i.e., representing China's least-developed regions). To the best of our knowledge, however, at-scale evidence remains to be provided of the SES differences in ECD outcomes for children under the age of 3 years in rural China. In order to address this research gap, the current study tests whether ECD outcomes differ significantly between children under age of 3 years in low-SES and high-SES households. In addition, the study tracks SES differences over age (i.e., how the SES gradient evolves for each child between the age of 6 months and the age of 3 years).

Over the past decade, parental training has been encouraged and tested as a scalable remediation strategy both to give more disadvantaged children in rural areas of China a fair start in life and to offset the intergenerational transmission of disadvantage. Evidence from 10 randomized controlled trials in rural China demonstrated that parental training programs focusing on caregiver-child interactions can improve ECD outcomes of disadvantaged children in rural areas (Emmers et al. 2021). The intervention programs were adapted to the local culture and settings in order to encourage program delivery, participation, and take-up. For example, considering that approximately 30% of the children under the age of 3 in rural China grow up in the care of a grandparent while their parents migrate to cities for work (Bai et al. 2022), grandparent caregivers in charge of the daily care of young children were encouraged to participate in the programs as well. Moreover, the underlying mechanisms of the treatment impacts of the parenting experiments on ECD outcomes remain to be investigated. In order to shed light on the underlying mechanisms, this study pools data from three cluster-randomized controlled trials that tested parental training programs in study sites in Central and Western China. The pooled data set allows us to test the aggregate treatment impacts of parental training on Bayley cognition, language, and motor scores of young children in rural China, as well as to explore the share of the treatment impacts on ECD outcomes that are mediated via changes in parenting beliefs and parental investment in caregiver-child interactions.

This study primarily makes three contributions to the literature. First, the study provides evidence of the rural-urban gap in ECD outcomes that is more representative for the whole of China than earlier studies. To be more specific, the current study is the first study that pools data on ECD outcomes of young children from study sites in rural and urban areas in all three major geographical regions: Eastern, Central, and Western China. We are able to accomplish this by using a unique dataset with ECD outcomes of 0- to 3-year-olds (N = 9,053) from rural study sites in Eastern, Central, and Western China as well as from urban Shanghai. Second, the study quantifies the SES gradient in ECD outcomes and tracks its age pattern (i.e., how the gradient evolves between the age of 6 months and 3 years). To the best of our knowledge, no earlier study has provided evidence of the SES gradient in skills for children under age 3 in China. In particular, no earlier study provided evidence of age patterns in the SES gradient via regular tracking of a birth cohort between the ages of 6 months and 3 years. Third, this is the first study that conducts a mediation analysis of the mechanisms behind treatment impacts on ECD outcomes of parental training programs in rural study sites in China.

The remainder of this paper is structured as follows. Section 2 focuses on the importance of equality and equity in the opportunity of ECD for Common Prosperity. The data and methodology are described in Sections 3 and 4, respectively. Section 5 presents the findings on the rural-urban gap in ECD outcomes, the SES gradient in ECD outcomes over time (between the ages of 6 and 30 months), the impacts of cluster-randomized parenting experiments on ECD outcomes, and the mediating role of parenting beliefs and parental investments. Section 6 provides a discussion and concludes.

2. Background: The Importance of ECD for Common Prosperity

China's government has announced its intention to steer China on a path of people-centered development toward more social equality and economic equity in the context of the "Common Prosperity" policy. This path is markedly different from China's developmental trajectory during the past couple of decades (X. Wang and Peach 2019). During the beginning of the reform and opening-up period, one of the slogans advocated by Deng Xiaoping was, "Let some people get rich first" (Naughton 1993). After decades of steady economic growth and institutional development, China achieved several notable milestones: China became the second largest economy in the world after the United States in 2010; China completed the eradication of extreme poverty in 2020; and China announced that the country had achieved its first Centenary Goal of becoming a *moderately prosperous society* in 2021 as the nation celebrated the 100th anniversary of the founding of the Communist Party (Kakwani et al. 2022). Despite achieving these milestones (specifically, China is now a moderately prosperous society without extreme poverty), the nation's GDP per capita remains below the average GDP per capita of the world (World Bank 2022), and regional and rural-urban inequalities have widened (Rozelle and Hell 2020; Junsen Zhang 2021).

Facing these remaining challenges, China's central leadership has announced that it is no longer acceptable to allow some people to fall too far behind (Kakwani et al. 2022). In its most high-profile recognition of the problem, since 2021, China's top leaders have given prominence to the phrase *Common Prosperity*, which is used to describe a policy shift that is intended to curtail income inequality and excessive wealth accumulation by individuals. For example, during the opening speech to the 20th National Congress of the Communist Party of China on October 16, 2022, General Secretary Xi Jinping stated:

> Achieving Common Prosperity is a defining feature of socialism with Chinese characteristics and involves a long historical process. The immutable goal of our modernization drive is to meet the people's aspirations for a better life. We will endeavor to maintain and promote social fairness and justice, bring prosperity to all, and prevent polarization. (p. 18)

Equality and equity of opportunity are the foundations of the Common Prosperity policy. A high degree of inequality of opportunity in China is due to the historic legacies of the country. According to their policy announcements, the current leadership appears to recognize that a large share of public investment has been channeled toward urban areas in the post-reform era at the dispense of rural areas. As a result of these systematic trends in investment and policy efforts, rural-urban inequality explains over 70% of income inequality in China today (Afridi, Li, and Ren 2015). Further, if the explanation in Xi Jinping's statement is valid and that rural residents are faced with a substantial inequality of educational opportunity that stems from developmental differences during early childhood is true, then it follows that providing young children with a fair start in life is crucial to address the "principal contradiction . . . between unbalanced and inadequate development and the people's ever-growing needs for a better life" (Kakwani et al. 2022, p. 29) that China's society faces today.

3. Data

The dataset contains information on 9,053 unique caregiver-child dyads from rural areas of Central and Western China (*n* = 6,205) and rural and urban areas in the provincial-level municipality of Shanghai (*n* = 2,848; Table 1). The Shanghai Birth Cohort data were collected between 2015 and 2018 by the Shanghai Key Laboratory of Children's Environmental Health and the Shanghai Children's Medical Center at Shanghai Jiaotong University. The sample contains 2,145 and 703 Shanghai residents and their young children living in urban and rural areas of Shanghai, respectively (see Table 1; for a more extensive description of the sampling strategy and data collection procedures, please refer to Zhang et al., 2019). The birth cohort data from Central and Western China were collected by an international research collaboration between the Rural Education Action Program and local partners in China between 2013 and 2020. For the samples from Central and Western China, repeated rounds of observations were collected for each child in the context of ongoing intervention studies (see Section 4.2 for further details). The subjects of the Shanghai Birth Cohort were not enrolled in a randomized controlled trial that included an intervention to the children assigned to the treatment arm. The pooled sample contains a total of 14,807 'untreated' observations (i.e., baseline or control group observations).

[Insert Table 1 about here]

Information on ECD outcomes and child and household characteristics were collected for each observation in the pooled dataset. In addition, for the observations from Central and Western China, we collected measures of parenting beliefs (e.g., belief that reading books is important for child development) and parental investments in child psychosocial stimulation (e.g., engagement of children in reading, storytelling, and interactive play). For the sample from

Shanghai, no secondary outcome measures of parenting beliefs and parental investment were collected.

3.1 ECD Outcomes

The ECD outcomes for each of the observations in the pooled sample were previously assessed via behavioral assessments using the first or third edition of the Bayley Scales of Infant and Toddler Development (Bayley 2006). The Bayley scale is generally considered the gold standard for measuring ECD outcomes (Del Rosario et al. 2021). The first edition of the Bayley (Bayley-I) test contains two scales for the assessment of 0- to 30-month-olds in the domains of cognition and motor development (Bayley 1933). The third edition of the Bayley (Bayley-III) test contains scales for the observational assessment of 1- to 42-month-olds in the developmental domains of cognition; (expressive and receptive) language; and (fine and gross) motor skills. The Bayley Scales were translated in Mandarin Chinese and validated for a healthy Chinese sample (Yi 1995; Xu et al. 2011). The observational assessments require the child to engage in a series of increasingly difficult tasks. If the child passes a task, the item receives a score of 1. Otherwise, the item receives a score of 0. When the child fails to perform six consecutive tasks, the test ends, and the raw score is found by the aggregation of the raw item scores. Studies have shown that the versions of the Bayley tests that are used in China have good reliability and validity (Hua et al. 2019; Yi 1995; Yue et al. 2019). The Bayley tests were all administered one-to-one via behavioral assessment by trained enumerators. Prior to test administration, the enumerators completed one week of intensive training, including 2.5 days of experiential learning in the field.

For use in the analysis of this study, the Bayley raw scores were converted into three intuitive ECD indicators. First, we converted the Bayley raw scores into Bayley composite scores, following the scoring guidelines in the Bayley manual (Bayley 2006; Yi 1995). The composite scores have a mean of 100, an SD of 15, and a range of 40–160. Second, we used the composite scores to construct indicators of developmental delay. We define developmental delay as a composite score of 1 or more SDs below the mean of a reference population whose developmental trajectory is expected to be normal (i.e., a child population in a developed region who was not born prematurely, severely malnourished, or severely diseased; Emmers et al. 2021).[2] This approach is in line with the usage guidelines of the Bayley test, which conventionally defines children with a score of more than 1 SD below the normative sample mean as mildly delayed in their development (Bayley 2006).

Third, with a view to comparability with treatment impact estimates reported in published studies, we convert the Bayley raw scores to standard normally distributed *z*-scores. Considering that Bayley raw scores increase with age, we eliminate the age effect by internally standardizing raw scores within age (month) groups. This is done by first estimating age-conditional means and SDs using non-parametric regression and then using these estimated statistics to compute age-adjusted internal *z*-scores for each subscale. This non-parametric standardization procedure has the advantage that it is less sensitive to outliers and small sample sizes within age categories than are parametric procedures (Rubio-Codina, Caridad Araujo et al. 2016).

3.2 Parenting Beliefs

In addition to measures of child development outcomes, the research team collected measures of parenting beliefs and parental investments in child psychosocial stimulating

---

[2] For the norm-referenced Chinese version of the Bayley-I test with a mean of 100 and SD of 15, the conventional cut-off score of 85 is used. In the absence of a Chinese norming sample, Bayley-III cut-off scores are defined based on the means (SDs) of Bayley-III composite scores for healthily developed populations that have been reported in the international literature (i.e., the mean (SD) is expected to be 105 (9.6), 109 (12.3), and 107(14) for the cognitive, language, and motor scales, respectively; Wang et al. 2019). Hence, the Bayley-III cut-off points for moderate cognitive, language, and motor delay are 95.4, 96.7, and 93, respectively.

activities for the rural samples from Central and Western China (Table 1). These measures were collected using household interview questionnaires administered to the household's main caregivers, whereby the main caregiver was defined as the person who is responsible for the daily nutrition and care of the child. To measure parenting beliefs, we asked a series of questions related to the caregiver's beliefs and attitudes toward parenting, including whether the caregiver thinks it is important to play interactively or read with the child. This set of questions was also used to assess parenting knowledge and beliefs in earlier published studies (e.g., Luo, Emmers et al. 2019; Sylvia et al. 2022).

3.3 Parental Investment

To measure parental investment in a cognitively stimulating home environment, we administered a subset of items extracted from the Family Care Indicators scale (FCI; Hamadani et al. 2010). Specifically, we assessed the number of picture books at the home and caregiver engagement in four at-home, play activities with the child: reading books, telling stories, singing songs, and playing with toys. Previous studies demonstrated that the FCI provides a reliable measure of parenting and the home environment in developing countries (Hamadani et al. 2010). The FCI has been adapted to the Chinese language/context and has been used in earlier studies across rural China (Wang et al. 2022). We then estimated a dedicated measurement system to relate the observed measures of parenting beliefs and parental investment to their corresponding latent factors using factor analysis (Appendix A). Finally, we internally standardized the factor scores within the age (in months) group of the child to obtain standard normally distributed factor *z*-scores that are uncorrelated with the age of the children (Rubio-Codina, Caridad Araujo et al. 2016).

3.4 Child and Household Characteristics

Information on basic child characteristics (e.g., sex, age, gestational age at birth) and caregiver characteristics (e.g., age, educational attainment) also was collected for the full sample. Finally, information on household SES was collected. For individuals in the Shanghai sample, survey respondents were asked to rate their socioeconomic situation on a 5-point Likert scale, for which 1 indicates a very comfortable life and 5 indicates that they have a difficult time financially. Scores of 1, 2, 3, and 4/5 corresponds to very well off, fairly well off, fairly poor, and very poor, respectively. For the rural samples from Central and Western China, the survey team collected information on the presence of a number of household assets at home. A household asset index was constructed based on six indicators: access to air conditioning, flush toilet, boiler, car, TV, and internet. Similar SES indices have been used extensively in the medical, demographic, nutritional, and economics literatures (e.g., Rubio-Codina, Attanasio et al. 2016; Schady et al. 2015). A household is defined as very (fairly) poor if the household has access to one to three (four to six) of these household assets.

4. Methods

4.1 Descriptive Analysis of Rural-Urban and SES Differences

We conduct the complete statistical analysis with Stata/MP, version 17.0. First, we summarized demographic characteristics and ECD outcomes for the rural and urban subsamples. Then, we compared the means of the demographic characteristics and ECD outcomes of children from rural study sites in Eastern, Central, and Western China with the respective means of their urban counterparts. Two-tailed *t*-tests were used to test for the significance of the differences in means.

Second, in line with the methodology of earlier studies (e.g., Fernald et al. 2012; Rubio-Codina et al. 2015; Schady et al. 2015), we provided descriptive evidence of the SES gradient in ECD outcomes by comparing ECD scores across different SES groups. Specifically, we compared ECD outcomes across households with high versus low access to household assets and different levels of maternal educational attainment. We used two-tailed *t*-tests to determine the significance of the differences in means. In addition, using a panel dataset with four waves of ECD scores (n = 1,801; Table 1), we tracked the evolution of this SES gradients of children between the ages of 6 and 30 months.

4.2 Impact Evaluation and Mediation Analysis

For treatment impact estimation and the mediation analysis, we pooled data from three cluster-randomized parenting experiments in rural study sites in Central and Western China. The study sample from Yunnan Province (see bottom row of Table 1) was excluded for the impact evaluation and mediation analysis, because of the lack of secondary outcome measures of parenting beliefs and parental investment (as described in Sections 3.2 and 3.3). First, we describe the three randomized parenting experiments from which data are pooled for the subsequent part of the analysis. We then explain the estimation strategy for the assessment of treatment impacts on primary (i.e., ECD outcomes) and secondary outcomes (i.e., parenting beliefs and parental investment) and the mediation analysis.

4.2.1 Intervention Design and Procedures

The pooled dataset contains data from three cluster-randomized intervention studies in rural study sites in Central and Western China. Taken together, the three randomized controlled trials cover 274 villages in 28 nationally designated poverty counties in rural Shaanxi, Yunnan, and Hebei Provinces (Appendix Figure B.1). In each study, treatment assignment was randomized at the village level to reduce the risk of contamination (among families within a village) across experimental groups. The impact evaluation sample includes 2,800 children with complete records of ECD scores at baseline and follow-up. Appendix B contains a detailed description of the study design, participants, randomization, masking, and procedures.

During the intervention period, all children and caregivers who resided in clusters of the intervention group were enrolled in a parental training program that focused on child psychosocial stimulation that was delivered at home or in a child center at a central location in the community. In the first home-based programs in rural Shaanxi, caregiver-child dyads were invited to participate in 24 weekly training sessions over a 6-month period between November 2014 and April 2015. In the second home-based program in Yunnan and Hebei Province, caregiver-child dyads were enrolled in 24 bi-weekly training sessions over a one-year period between October 2015 and September 2016. In the center-based program in Shaanxi, caregiver-child dyads were invited to participate in 48 weekly training sessions over a period of one year. During each of these parental training sessions, community health workers (paid and managed by the research team) trained caregivers on interactive caregiver-child activities (e.g., reading, storytelling, interactive playing with toys). Training sessions were designed to last for one hour.

All interactive caregiver-child activities were stage-based and fully scripted in a curriculum. This curriculum was based loosely on the curriculum developed for a pioneering parenting intervention in Kingston, Jamaica, and adapted to the rural China setting by local child development psychologists and ECD experts (Sylvia et al. 2021). Each activity focused on child development in one of four skill domains: cognition, language, motor, or social-emotional development. Two new activities were introduced during each training session (i.e., parents were instructed on practicing these activities with their children) such that caregivers would encounter at least one new activity to stimulate development in each of the four targeted skill domains each month. In addition to developmental activities, the curriculum also included one module with basic information on child health and nutrition that was delivered at the end of each session.

Training sessions were delivered by community health workers in the service of China's National Health Commission. Before the start of the intervention, each of the community health workers had to complete a one-week intensive training course on basic theories of ECD and on communication, coaching, and counselling techniques. The training program consisted of five days of classroom-based instruction by ECD experts in combination with two days of field practice. Fully scripting the curriculum obviated the need for more extensive training of community health workers. Upon completion of the training course, each of the community health workers received a copy of the fully scripted ECD curriculum and a toolkit (which included a growth chart, development check list, counseling materials, and user handbook). To monitor service quality, a carefully designed supervisory system was set up that involved tracking the completion of training sessions (via phone, via mobile app, or by managers at the child centers), unannounced supervisory visits with feedback, and phone interviews with caregivers for feedback. Throughout the program, trainers received periodic refresher training by phone (in the home visitation programs) or in person (in the center-based program).

4.2.2 Empirical Estimation Strategy

First, we estimated the impact of the treatment assignment on the primary (i.e., ECD outcomes) and the secondary outcomes of interest (i.e., parenting beliefs and parental investment). Provided that treatment assignment and attrition were random, a comparison of the means of outcomes between treatment arms resulted in unbiased estimates of the treatment effect. We used ordinary least squares regression to estimate intention-to-treat (ITT) effects with the following ANOVA regressions specification:

$$Y_{ij1} = \alpha_0 + \alpha_1 T_j + \alpha_2 Y_{ij0} + \alpha_3 X_{ij1} + \pi_s + \varepsilon_{ij1}, \tag{1}$$

where Y_{ij1} is a primary or secondary outcome measure for child *i* in village *j* at the end of intervention period; T_j is a dummy variable that indicates treatment assignment of village *j*; Y_{ij0} is the outcome measure for child *i* at baseline; and X_{ij1} is a vector of covariates that includes children's sex, age (in months), an indicator for premature birth status, and the relationship between the child and the primary caregiver. π_s is a set of province, study, and cohort fixed effects. We adjusted standard errors for clustering at the village level, using the cluster-correlated Huber-White estimator.

Second, we conducted a mediation analysis to explore the underlying mechanisms of the treatment impacts on ECD outcomes. The goal in a mediation analysis is to decompose the total treatment effect on the primary outcome into the indirect and direct effects (Keele, Tingley, and Yamamoto 2015). Because we were estimating ITT effects, we decomposed the average total ITT effect ($\bar{\tau}_{ITT}$) into a mediated ITT effect ($\bar{\mu}_{ITT}$) and an unmediated ITT effect ($\bar{\upsilon}_{ITT}$). $\bar{\mu}_{ITT}$ measures the extent to which the effect of the parenting intervention on skill development is transmitted through parenting beliefs and/or parental investment. $\bar{\upsilon}_{ITT}$ reflects the proportion of the average ITT effect that results from all other treatment-induced changes, both observed and unobserved. Following the methodology of Heckman, Pinto, and Savelyev (2013), we first estimated the following augmented outcome equation, for which we added the mediation channels of interest (i.e., the secondary outcomes) to equation (1):

$$Y_{ij1} = \alpha_0 + \alpha_1 T_j + \alpha_2 M_{ij1} + \alpha_3 Y_{ij0} + \alpha_4 X_{ij1} + \pi_s + \varepsilon_{ij1} \tag{2}$$

where $Y_{ij1}$is the ECD outcome of child *i* in village *j* at the end of the intervention period, and M_{ij1} is the mediating outcome for child *i* in village *j*. In all of our specifications, we included a set of covariates (X_{ij1}) to account for shocks that potentially affect both the ECD outcomes and our mediating outcomes. We first added one mediator to estimate the indirect effect of this particular mediator. Next, we included both mediators (i.e., parenting knowledge and parental investment) simultaneously to examine the total indirect effect of the two mediators.

To attribute a causal interpretation to the mediation analysis, we needed to assume that the mediating variables are exogenous with respect to ECD outcomes, conditional on other observed covariates and treatment assignment. It is possible, however, that there were unobserved variables that are affected by the intervention and, in turn, affected ECD outcomes and the secondary outcomes of interest, which would lead to biased estimates of $\bar{\mu}_{ITT}$. To address this issue, we follow the methodology proposed by Oster (2019) to assess the sensitivity of $\bar{\mu}_{ITT}$ to the inclusion of additional controls. We calculate a statistic known as Oster's δ for each $\bar{\mu}_{ITT}$.[3] This statistic captures the ratio between the selection bias introduced by unobservable variables and the selection bias introduced by the observable controls that would be required such that the true value of $\bar{\mu}_{ITT}$ would be statistically insignificant from zero. This means, for example, that a value of Oster's δ = 1 or δ = 2 suggests, respectively, that the explanatory power of unobserved confounders would need to be at least as high or twice as high as the explanatory power of the observed controls to produce an insignificant μ_{ITT}. As a rule, values of Oster's δ larger than 1 indicate that $\bar{\mu}_{ITT}$ is unlikely to be biased due to unobserved heterogeneity.

Finally, to allow for multi-stage mediation effects, we conducted a multi-stage mediation analysis using structural equation modelling (Hayes 2009). In particular, we aimed to test

[3] As suggested by Oster (2019), we set R_{max} = 1.3$\tilde{R}$. R_{max} refers to the R^2-statistic that would be observed if observables and unobservables were all included in the regression. $\tilde{R}$ is the R^2-statistic obtained after estimating regression specification (2).

whether the parental training intervention led to an improvement in parenting beliefs, which, in turn, led to an improvement in parental investment and, as a result, improved ECD outcomes. We used structural equation modelling to simultaneously estimate three equations: (a) an augmented equation (2) to estimate the impact of treatment assignment and the mediators on the ECD outcome of interest; (b) a mediation equation (1) to predict the impact of the treatment assignment and parental investment on parenting beliefs; and (c) a mediation equation (1) to predict the impact of treatment assignment on parental investment. In this model, the total indirect effect consists of three indirect effects: the indirect effect through parenting beliefs; indirect effect through parental investment; and indirect effect through parenting beliefs and parental investment. We used a bootstrapping approach with 1,000 replications to construct bias-corrected 95% confidence intervals for the estimated effects (Preacher and Hayes 2008).

5. Results

5.1 Rural-Urban Differences in Demographics and ECD Outcomes

Table 2 provides a summary of the demographic characteristics of the pooled sample. We observe that 48.0% of the 0- to 3-year-olds in the study sample are male; 6.1% of the children were born prematurely (i.e., born before 37 weeks of pregnancy); 62.7% of the children were first born; and 37.3% had an older sibling. Mothers were, on average, 27.7 years old at the time of survey administration, and 42.1%, 22.8%, and 5.0% of the mothers graduated from senior high school, college or university, and postgraduate education, respectively. Finally, 24.5% of the households were relatively well off (in terms of family assets), while the remaining 75.5% were relatively poor.

[Insert Table 2 about here]

Table 2 also provides a comparison of summary statistics between Shanghai and the rural sample from Central and Western China. Children from Shanghai were, respectively, 1.5 and 33.9 percentage points less likely to be born prematurely and to have an older sibling than were their rural counterparts from Central and Western China (ps < .01). Shanghai mothers were, on average, 1.4 years older ($p < .01$). Finally, mothers from Shanghai were 78.5, 60.7, and 11.1 percentage points more likely to have graduated from senior high school, college or university, and postgraduate education than were mothers from the rural study sites in Central and Western China, respectively (ps < .01).

Table 3 presents the differences in means (and SDs) between the rural and urban in our China samples. A comparison of the means in Columns 1 and 3 demonstrates that Bayley cognition, language, and motor composite scores are, on average, 31.3, 5.7, and 11.5 points higher, respectively, for the children from the Shanghai urban sample than for the rural sample from Central and Western China (ps < .01). The table also provides evidence of the differences in developmental outcomes between urban children from Shanghai and rural children from Chongming, a peri-urban county in the municipality of Shanghai (Columns 1 and 2). The results show that, respectively, the cognition, language, and motor scores of children from urban Shanghai are 32.3, 16.8, and 11.7 points higher than the scores of their same-age peers from Chongming (ps < .01). Interestingly, the findings show that cognition and motor scores do not differ significantly between children from Chongming, a rural area in Eastern China, and children from rural areas in Central and Western China ($p > 0.10$). Language scores are 11.1 points lower for children from Chongming as compared to children from rural study sites in Central and Western China (ps < .01).

[Insert Table 3 about here]

Figure 1 illustrates the prevalence of cognitive, language, and motor delay for the rural and urban samples in the study. For the Shanghai urban sample, the data show that the prevalence of cognitive and motor delay is low (<10%), whereas the prevalence of language delay is 35.6%. For the rural study sample, we find that the prevalence of delay is significantly higher (*p*s < .01). The prevalence of cognition delay is 56.1% and 46.7% for the rural samples from Chongming and Central and Western China, respectively. The prevalence of language delay is 86.2% and 50.2%, and the prevalence of motor delay is 25.8% and 26.9% for the respective rural samples. With an overall prevalence of cognition, language, and motor delay of 47.7%, 53.9%, and 26.8% in the pooled rural sample,[4] we find that the prevalence of cognition, language, and motor delay is, respectively, 43.2, 18.3, and 20.7 percentage points higher in the rural sample than in the urban sample (*p*s < .01).

[Insert Figure 1 about here]

Based on a subsample of the overall dataset (n = 1,801), which contains repeated observations of Bayley-I cognition and motor scores from rural areas in Central/Western China, Figure 2 presents trends in mean ECD scores and the prevalence of delay as the children age between 6 months and 29 months. With regard to cognition scores, the left-hand side of Figure 2 demonstrates that the mean Bayley cognition scores of children deteriorate during the first years of life (i.e., between the ages of 6 and 11 months and 24 and 29 months); the prevalence of cognitive delay increases from 19.9% to 57.0%. In contrast, with regard to motor scores, we observe an improvement in the composite scores and a decrease in the prevalence of delay from 32.1% to 17.1% during this early stage of life.

[4] By calculation of the sample-size weighted average of the prevalence of delay among young rural children in Chongming and areas in Central and Western China, we obtain an aggregate measure of the prevalence of delay across all rural study sites.

[Insert Figure 1 about here]

5.2 SES Gradient in ECD Outcomes

Table 4 presents differences in ECD outcomes between subgroups based on maternal educational attainment and household SES. The results show that ECD outcomes in the rural and urban samples are consistently higher for children with mothers who completed more years of education and in households with a higher SES. Each of the differences in the means of the ECD outcomes between subgroups, based on maternal educational attainment and household SES, is significant at the 5% level.[5] These results indicate that maternal educational attainment and household SES are important predictors of ECD outcomes.

[Insert Table 4 about here]

The results presented in Figure 3 add to these findings by showing that inequalities in ECD outcomes tend to persist as children age and may even persist over generations. Figure 3a shows that children who belong to the bottom (top) of the skill distribution at age 6–11 months, in general, still belong to the top (bottom) of the skill distribution at age 24–29 months (*p*s < .01).[6] Hence, in the absence of any intervention, on average, disadvantaged children who suffer from an early delay in development are unlikely to catch up with their more advantaged peers. Figures 3b and 3c also provide evidence of the intergenerational transmission of developmental

---

[5] In addition, Table 4 shows that ECD outcomes are consistently higher for young girls than for young boys ($p_s < .01$). Cognition, language, and motor scores are higher for term-born children than for preterm-born children in the rural samples from Central and Western China ($p_s < .05$). These differences in subgroup means for term versus preterm-born children are insignificant for the Shanghai sample.

[6] Note that the gap in cognitive skill development between the bottom quartile of children and the remainder of the sample narrow between the age of 6–11 months and 12–17 months. This early, small, but noticeable catch-up effect may be caused by a micronutrient fortification program that provided multiple micronutrient powders for young children to half of the caregivers in the sample. As reported in Luo et al. (2017), the micronutrient fortification program that was initiated between the age of 6–11 months had a small positive impact on child cognitive development, but this positive impact faded out afterwards (i.e., when the sample children were 18 months and 24 months)."

delay. If mothers completed fewer years of education or households have lower SES, then their children are more likely to have lower ECD scores at age 6–11 months (ps $< .01$). Moreover, these inequalities persist at least until the children reach 24–29 months of age.

[Insert Figure 3 about here]

5.3 Treatment Impacts of Parental Training on Primary and Secondary Outcomes

We conducted two-tailed *t*-tests to check whether outcomes were balanced across the treatment and control groups at baseline. First, we conducted a balance test for the full baseline sample (Appendix Table B.1 in Appendix B). According to the data, there were no significant differences between the baseline treatment and control subsamples (ps $> .10$). Second, we tested the balance between the subsample of stayers (i.e., the subsample that did not attrite by the time of follow-up data collection). The results in Appendix Table B.2 show that the characteristics of stayers are balanced across the treatment/controls arms as well. In addition, an omnibus test for the joint significance of all baseline characteristics confirms that treatment assignment cannot be predicted by the observed covariates for either the full sample or the subsample of stayers (ps $> .10$). Finally, attrition is insignificantly correlated with treatment status ($p > .10$). Hence, in this study, the randomization (and/or attrition bias) are unlikely to affect estimated ITT effect.

Figure 4 depicts the estimated ITT effects on primary and secondary outcomes. Parental training improves child cognition *z*-scores, on average, by 0.20 SD ($p < .01$). In addition, parental training has a small but significant and positive impact on child language and motor *z*-scores of 0.09 SD and 0.08 SD, respectively (ps $< .10$). With regard to the secondary outcomes of interest, the ITT effect estimates illustrate that parental training, on average, increases both parenting beliefs and parental investment by 0.17 SD and 0.26 SD, respectively (ps $< .01$).

[Insert Figure 4 about here]

5.4 Mediated and Unmediated Impacts

Table 5 presents the findings of the mediation analysis. In Column 1, we report the ITT effect on child cognition, language, and motor *z*-scores, controlling for a set of baseline child and household characteristics. In Columns 2 and 3, we report our estimations of direct ITT effects and the impact of each of the mediators. Finally, we estimate the direct ITT effects and the impacts of both of the mediators simultaneously, which we report in Column 4.

The results of this mediation analysis show that a significant share of the treatment impact is mediated via parenting beliefs and parental investment. After controlling for treatment assignment to the parental training program, we find that the belief and investment factor *z*-scores significantly predict the cognition, language, and motor *z*-scores of young children (*p*s < .10). When our analysis included the mediator outcomes in the regression specification, the estimated direct impact of the parenting program on Bayley cognition, language, and motor scores were reduced, respectively, from 0.20 SD to 0.15 SD (*p*s < .01), from 0.09 SD (*p* < .10) to 0.04 SD (*p* > .10), and from 0.08 SD (*p* < .10) to 0.06 SD (*p* > .10). Considering that the estimates for the Oster's δ of each regression specification range between 3.77 and 9.18, it can be concluded that there is evidence that the estimated coefficients of the mediated impacts are robust to unobserved heterogeneity.

[Insert Table 5 about here]

Table 6 shows the decomposition of τ_{ITT} into μ_{ITT} and υ_{ITT}. The mediated effect is further disaggregated in the share of the τ_{ITT} that is mediated via parenting beliefs, parental investment, and the multi-stage mediation effect via parenting beliefs and parental investment. As shown in Column 1 of Table 6, in total, 22.3% of the ITT effect on child cognition *z*-scores is mediated via the secondary outcomes of interest; 8.2%, 12.1%, and 2.0% of μ_{ITT} are attributable to the mediating role of parenting beliefs, parental investment, and the multi-stage mediation effect, respectively. For language development, we find that, in total, 53.0% of τ_{ITT} is mediated via the secondary outcomes of interest (Column 2); 16.5%, 30.8%, and 5.7% of μ_{ITT} are attributable to the mediating role of parenting beliefs, parental investment, and the multi-stage mediation effect, respectively. Finally, for motor scores, we find that, in total, 23.2% of τ_{ITT} is mediated via the secondary outcomes of interest (Column 2); 12.7%, 9.0%, and 1.5% of μ_{ITT} are attributable to the mediating role of parenting beliefs, parental investment, and the multi-stage mediation effect, respectively.

[Insert Table 6 about here]

6. Discussion and Conclusion

In this study, we show that rural-urban differences in educational opportunities in China emerge early in life. One major finding is that, even before the age of 3, children in rural study sites across rural areas in Eastern, Central, and Western China are already lagging behind in their development as compared to their urban counterparts.[7,8] On the one hand, we find that, overall, urban children in Shanghai are developmentally on track with a prevalence of cognition delay of only 4.5%. The prevalence of motor skill delays (6.1%) also demonstrates that urban children do not suffer from developmental delays. Only levels of language skills (35.6%) are high, as

---

[7] The reported risks of cognitive, language, and motor delay of 46.7%, 50.2%, and 25.8%, respectively, that are reported for rural children from Central and Western China are in line with the risks of delay reported in earlier studies from rural study sites in Central and Western China. For example, Wang et al. (2019) report a 49%, 52%, and 30% risk of cognition, language, and motor delay, respectively; and Emmers et al. (2021) report a 45% and 46% risk of cognition and language delay, respectively.

[8] Earlier evidence on the risk of developmental delay in urban areas of China is scarce. One study used the Bayley-III cognition scale to assess child cognitive development of 1,444 infants and toddlers in middle-sized Chinese cities (Hua et al. 2019). In another study, Wang et al. (2023) administered the Bayley-III cognition, language, and motor scales to 70 young children in Xi'an. The findings of the Hua et al. (2019) and Wang et al. (2023) studies confirm that children in urban areas of China resemble a healthily developed population with a prevalence of delay of approximately 15% or lower.

compared to international norms. The healthy levels of development of urban children in the sample are in line with the research of Zhang et al. (2021), who found that only 14% of all children in Shanghai were at risk of any developmental delay. On the other hand, our findings show that rural children are, respectively, 43.2, 18.3, and 20.7 percentage points more likely to be delayed in their cognition, language, and motor development than are their urban peers in Shanghai.

With regard to language development, we observe that, although language delay in urban China was high compared to international norms, the prevalence of language delay in rural China is even higher. Interestingly, the prevalence of language delay of rural sample children is more prevalent in the peri-urban sample from Chongming County in Eastern China than is the prevalence of language delay among sample children from rural areas in Central and Western China. Such a finding can perhaps be explained by earlier research that shows that the home language environment of young children in peri-urban areas tends to be less stimulating, with lower levels of child-directed speech than in the home language environment in remote rural areas; this explanation suggests that such patterns of delay may be due to more stringent time constraints and a higher risk of mental health issues of peri-urban caregivers as compared to caregivers in more remote rural communities (Feng et al. 2023; Ma et al. 2023).

The significantly higher rates of delay among rural children, compared to the rates of delay among urban children, are not surprising, given the income differentials in China between the relatively well-off families in China's cities and the relatively poorer families in rural China. In 2021, the average disposable income per capita in urban Chinese households was approximately 2.5 times higher than in rural households (i.e., the disposable income per capita amounted to CNY 47,412 in urban households as compared to CNY 18,931 in rural households—National

Bureau of Statistics of China 2022).[9] The finding that children growing up in lower-SES families are at an elevated risk of developmental delay as compared to children in higher-SES families is in line with the global literature showing that the elevated risk of exposure to inadequate nutrition and poor environments in low-SES households can cause developmental delay (Bornstein and Bradley 2002; Falk et al. 2021; Heckman and Mosso 2014; Schady et al. 2015).

The study also provides evidence of the persistence of inequalities in ECD outcomes over age and across generations. In this regard, when comparing high and low-SES families, gaps in developmental outcomes of young children emerge as early as 6 months of age and do not change substantively by age 2.5. These findings are in line with evidence from other LMICs, including Chile, Colombia, Ecuador, India, Indonesia, Nicaragua, Peru, and Senegal (Fernald et al. 2012; Rubio-Codina et al. 2015; Schady et al. 2015). Considering that children who grow up in rural, lower-SES families are at a higher risk of developmental delay than are children from higher-SES families, we conclude that the results provide evidence of the intergenerational transmission of disadvantage in China (Bornstein and Bradley 2002; Falk et al. 2021; Heckman and Mosso 2014; Schady et al. 2015).

Despite the fact that the current environment in rural China is associated with high levels of developmental delays, research shows that parental training programs can function as a promising strategy to support the developmental potential of at-risk populations of young children in rural communities in China. Based on evaluations of randomized experiments, the literature demonstrates that parental training programs implemented in rural study sites in

---

[9] For the study regions of the current study, 2021 statistical yearbook data show that the average disposable income per capita for urban Shanghainese households was CNY 51,295, while the average disposable income per capita in Chongming and rural areas of Yunnan, Hebei, and Shaanxi was CNY 33,140, CNY 14,197, CNY 18,179, and CNY 14,745, respectively (National Bureau of Statistics of China 2022).

Central and Western China, on average, can improve child cognition, language, and motor development by 0.20 SD, 0.09 SD, and 0.08 SD, respectively. The finding of a significant positive treatment impact on child cognition of 0.20 SD is in line with the findings of earlier studies from rural China, including those of Luo, Emmers et al. (2019), Sylvia et al. (2021), and Sylvia et al. (2022), who report treatment impacts of 0.11 SD to 0.28 SD on child cognition. Although the Luo, Emmers et al. and Sylvia et al. studies do not detect evidence of significant treatment impacts on language or motor development, the large size of our pooled evaluation sample (n = 2,800) does allow for the detection of the relatively small, but positive and significant, impacts on language and motor development. Further, the detected impacts of 0.17 SD and 0.26 SD on parenting knowledge and parental investment, respectively, also are in line with impact sizes reported in the literature (Emmers et al. 2021).

In addition, the current study, by investigating underlying mechanisms of the positive treatment impacts on cognition, language, and motor development, can inform policymakers about how the design of cost-effective parental training programs can support the development of disadvantaged children at scale and advance progress toward Common Prosperity. For each of these three developmental delays, the findings show that 22.3% to 53.0% of the impact is mediated via parenting beliefs in the importance of ECD and the early home environment as well as parental investments in a stimulating home environment. The overall mediated impact consists of three parts: the directly mediated impact via an increase in parenting beliefs, directly mediated impact via an increase in parental investment, and indirectly mediated impact via an increase in parenting beliefs and a subsequent increase in parental investment. Finally, it is noteworthy that the finding that the mediators play a more important role in language development (i.e., 53.0% of the impact on language z-scores) than for cognition or motor development (i.e., 22.3% and

23.2% of the impacts on development in the respective domains) is likely due to the measurement of the secondary outcomes. Respectively, two and four out of five input measures that were used to construct the parenting beliefs and parental investment factor *z*-scores were related to the home language environment (Appendix Table A.2). Therefore, we believe that it is more likely that our estimates underestimate rather than overestimate the mediated impacts. We recommend that future research gather more detailed information on parenting beliefs and parental investment related to specific developmental domains (e.g., cognitive or motor development) to produce a more highly refined mediation analysis of domain-specific beliefs and investments.

We acknowledge a number of limitations to this study. First, the study took place in sites in four Chinese provinces. Although it is possible that the results may differ in other regions, we are confident that the evidence provided in this study on the rural-urban gap in ECD outcomes is the most representative evidence available for the whole of China. In addition, the current study explores the mediating role of parenting beliefs and parental investments for impacts of parental training programs on the ECD outcomes of the sample children. We acknowledge that alternative mediating pathways may exist (e.g., through disciplining practices and/or through the mental health of caregivers). Due to a lack of secondary outcome measures for the sample from Shanghai, the current study is unable to compare mediating pathways across urban and rural contexts. Future studies are needed to investigate a more diverse set of mediating pathways.

In conclusion, the evidence shows that early developmental delays are prevalent across rural children in study samples from Eastern, Central, and Western China, whereas urban children in urban Shanghai are developmentally on track. Moreover, we find that these developmental inequalities tend to persist over time (as children age from 6 months to 30 months) and most likely will persist across generations. Providing young children with a fair start in life is foundational to making progress toward Common Prosperity. Parental training programs that focus on psychosocial stimulation are a promising tool to improve children's early cognitive, language, and motor development. The impacts are mediated via gains in parenting beliefs and parental investment in a stimulating home environment. To bring parental training programs to scale and deliver them as regular child and family services, further research is needed to determine which program features are key to improving cost-effectiveness, inclusiveness, and scalability of programs. For example, future studies are needed to investigate which program components (e.g., one-to-one interactive reading activities versus group-based play activities) or mediating parental investments (e.g., playing games of peekaboo versus making pegboard puzzles) can be more effective at stimulating development in certain human capital domains.

Declaration of Interest Statement

The authors have nothing to declare.

References

Afridi, Farzana, Sherry Xin Li, and Yufei Ren. 2015. "Social Identity and Inequality: The Impact of China's Hukou System." *Journal of Public Economics* 123 (March): 17–29. https://doi.org/10.1016/j.jpubeco.2014.12.011.

Amaro Da Costa Luz Carneiro, Pedro Manuel, Emanuela Galasso, Italo Xavier Lopez Garcia, Paula Bedregal, and Miguel Cordero. 2019. "Parental Beliefs, Investments, and Child Development : Evidence from a Large-Scale Experiment." WPS8743. The World Bank. http://documents.worldbank.org/curated/en/191061550167761091/Parental-Beliefs-Investments-and-Child-Development-Evidence-from-a-Large-Scale-Experiment.

Attanasio, Orazio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. 2020. "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia." *American Economic Review* 110 (1): 48–85. https://doi.org/10.1257/aer.20150183.

Attanasio, Orazio, Sarah Cattan, and Costas Meghir. 2022. “Early Childhood Development, Human Capital, and Poverty.” *Annual Review of Economics* 14 (1): 853–92. https://doi.org/10.1146/annurev-economics-092821-053234.

Attanasio, Orazio, Flávio Cunha, and Pamela Jervis. 2019. “Subjective Parental Beliefs. Their Measurement and Role.” NBER Working Paper 26516. Working Paper Series. https://doi.org/10.3386/w26516.

Attanasio, Orazio P., Camila Fernández, Emla O. A. Fitzsimons, Sally M. Grantham-McGregor, Costas Meghir, and Marta Rubio-Codina. 2014. “Using the Infrastructure of a Conditional Cash Transfer Program to Deliver a Scalable Integrated Early Child Development Program in Colombia: Cluster Randomized Controlled Trial.” *BMJ* 349 (September): g5785. https://doi.org/10.1136/bmj.g5785.

Bai, Yu, Dorien Emmers, Ying Li, and Lei Tang. 2022. “Parental Locus of Control and Early Childhood Development: Evidence on Parent and Grandparent Caregivers in Rural China.” *China Economic Review* 74 (August): 101814. https://doi.org/10.1016/j.chieco.2022.101814.

Bayley, Nancy. 1933. *The California First-Year Mental Scale*. University of California Syllabus Series 243. Berkeley, CA: University of California.

———. 2006. *Bayley Scales of Infant and Toddler Development– Third Edition*. Third edition. San Antonio, TX: Harcourt Assessment, Inc.

Bornstein, Marc H. 2012. “Cultural Approaches to Parenting.” *Parenting* 12 (2–3): 212–21. https://doi.org/10.1080/15295192.2012.683359.

Bornstein, Marc H., and Robert H. Bradley, eds. 2002. “Socioeconomic Status, Parenting, and Child Development: An Introduction.” In *Socioeconomic Status, Parenting, and Child Development*. Routledge.

Clark, Helen, Awa Marie Coll-Seck, Anshu Banerjee, Stefan Peterson, Sarah L. Dalglish, Shanthi Ameratunga, Dina Balabanova, et al. 2020. “A Future for the World’s Children? A WHO–UNICEF–Lancet Commission.” *The Lancet* 395 (10224): 605–58. https://doi.org/10.1016/S0140-6736(19)32540-1.

Cunha, Flávio. 2015. “Subjective Rationality, Parenting Styles, and Investments in Children.” In *Families in an Era of Increasing Inequality: Diverging Destinies*, edited by Paul R. Amato, Alan Booth, Susan M. McHale, and Jennifer Van Hook, 83–94. National Symposium on Family Issues. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-08308-7_6.

Currie, Janet, and Douglas Almond. 2011. “Chapter 15 - Human Capital Development before Age Five.” In *Handbook of Labor Economics*, edited by David Card and Orley Ashenfelter, 4:1315–1486. Elsevier. https://doi.org/10.1016/S0169-7218(11)02413-0.

Del Rosario, Chelo, Marie Slevin, Eleanor J Molloy, Jean Quigley, and Elizabeth Nixon. 2021. “How to Use the Bayley Scales of Infant and Toddler Development.” *Archives of Disease in Childhood - Education & Practice Edition* 106 (2): 108–12. https://doi.org/10.1136/archdischild-2020-319063.

Emmers, Dorien, Juan Carlos Caro, Scott Rozelle, and Sean Sylvia. 2022. “Early Parenting Interventions to Foster Human Capital in Developing Countries.” *Annual Review of Resource Economics* 14.

Emmers, Dorien, Qi Jiang, Hao Xue, Yue Zhang, Yunting Zhang, Yingxue Zhao, Bin Liu, et al. 2021. “Early Childhood Development and Parental Training Interventions in Rural

China: A Systematic Review and Meta-Analysis." *BMJ Global Health* 6 (8): e005578. https://doi.org/10.1136/bmjgh-2021-005578.
Falk, Armin, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. 2021. "Socioeconomic Status and Inequalities in Children's IQ and Economic Preferences." *Journal of Political Economy* 129 (9): 2504–45. https://doi.org/10.1086/714992.
Feng, Tianli, Xinwu Zhang, Lulu Zhou, Yue Ma, Yue Zhang, Lucy Pappas, Sarah-Eve Dill, and Scott Rozelle. 2023. "Variations in the Home Language Environment and Early Language Development in a Peri-Urban Community in China." *SCCEI Working Paper*.
Fernald, Lia C. H., Patricia Kariger, Melissa Hidrobo, and Paul J. Gertler. 2012. "Socioeconomic Gradients in Child Development in Very Young Children: Evidence from India, Indonesia, Peru, and Senegal." *Proceedings of the National Academy of Sciences* 109 (supplement_2): 17273–80. https://doi.org/10.1073/pnas.1121241109.
Fernald, Lia C. H., Elizabeth Prado, Patricia Kariger, and Abbie Raikes. 2017. *A Toolkit for Measuring Early Childhood Development in Low and Middle-Income Countries*. World Bank, Washington, DC. https://doi.org/10.1596/29000.
Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344 (6187): 998–1001. https://doi.org/10.1126/science.1251178.
Global Research on Developmental Disabilities Collaborators. 2018. "Developmental Disabilities among Children Younger than 5 Years in 195 Countries and Territories, 1990-2016: A Systematic Analysis for the Global Burden of Disease Study 2016." *The Lancet. Global Health* 6 (10): e1100–1121. https://doi.org/10.1016/S2214-109X(18)30309-7.
Goodman, Alissa, Robert Joyce, and James P. Smith. 2011. "The Long Shadow Cast by Childhood Physical and Mental Problems on Adult Life." *Proceedings of the National Academy of Sciences* 108 (15): 6032–37. https://doi.org/10.1073/pnas.1016970108.
Hamadani, J.D., F. Tofail, A. Hilaly, S.N. Huda, P. Engle, and S.M. Grantham-McGregor. 2010. "Use of Family Care Indicators and Their Relationship with Child Development in Bangladesh." *Journal of Health, Population and Nutrition* 28 (1): 23–33. https://doi.org/10.3329/jhpn.v28i1.4520.
Hayes, Andrew F. 2009. "Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium." *Communication Monographs* 76 (4): 408–20. https://doi.org/10.1080/03637750903310360.
Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312 (5782): 1900–1902. https://doi.org/10.1126/science.1128898.
———. 2008. "School, Skills, and Synapses." *Economic Inquiry* 46 (3): 289–324. https://doi.org/10.1111/j.1465-7295.2008.00163.x.
Heckman, James J., and Stefano Mosso. 2014. "The Economics of Human Development and Social Mobility." *Annual Review of Economics* 6 (August): 689–733. https://doi.org/10.1146/annurev-economics-080213-040753.
Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–86. https://doi.org/10.1257/aer.103.6.2052.
Hua, Jing, Yu Li, Kan Ye, Yujie Ma, Senran Lin, Guixiong Gu, and Wenchong Du. 2019. "The Reliability and Validity of Bayley-III Cognitive Scale in China's Male and Female

Children.” *Early Human Development* 129 (February): 71–78. https://doi.org/10.1016/j.earlhumdev.2019.01.017.

Kakwani, Nanak, Xiaobing Wang, Ning Xue, and Peng Zhan. 2022. “Growth and Common Prosperity in China.” *China & World Economy* 30 (1): 28–57. https://doi.org/10.1111/cwe.12401.

Keele, Luke, Dustin Tingley, and Teppei Yamamoto. 2015. “Identifying Mechanisms Behind Policy Interventions Via Causal Mediation Analysis.” *Journal of Policy Analysis and Management* 34 (4): 937–63. https://doi.org/10.1002/pam.21853.

Li, Hongbin, Prashant Loyalka, Scott Rozelle, and Binzhen Wu. 2017. “Human Capital and China’s Future Growth.” *Journal of Economic Perspectives* 31 (1): 25–48. https://doi.org/10.1257/jep.31.1.25.

Luo, Renfu, Dorien Emmers, Nele Warrinnier, Scott Rozelle, and Sean Sylvia. 2019. “Using Community Health Workers to Deliver a Scalable Integrated Parenting Program in Rural China: A Cluster-Randomized Controlled Trial.” *Social Science & Medicine* 239 (October): 112545. https://doi.org/10.1016/j.socscimed.2019.112545.

Luo, Renfu, Fang Jia, Ai Yue, Linxiu Zhang, Qijia Lyu, Yaojiang Shi, Meredith Yang, Alexis Medina, Sarah Kotb, and Scott Rozelle. 2019. “Passive Parenting and Its Association with Early Child Development.” *Early Child Development and Care* 189 (10): 1709–23. https://doi.org/10.1080/03004430.2017.1407318.

Luo, Renfu, Ai Yue, Huan Zhou, Yaojiang Shi, Linxiu Zhang, Reynaldo Martorell, Alexis Medina, Scott Rozelle, and Sean Sylvia. 2017. “The Effect of a Micronutrient Powder Home Fortification Program on Anemia and Cognitive Outcomes among Young Children in Rural China: A Cluster Randomized Trial.” *BMC Public Health* 17 (September): 738. https://doi.org/10.1186/s12889-017-4755-0.

Luo, Weixiang, and Yu Xie. 2020. “Economic Growth, Income Inequality and Life Expectancy in China.” *Social Science & Medicine* 256 (July): 113046. https://doi.org/10.1016/j.socscimed.2020.113046.

Ma, Yue, Lucy Pappas, Andrew Rule, Xinwu Zhang, Yujuan Gao, Sarah-Eve Dill, Tianli Feng, et al. 2023. “Does Urbanization Close the 30-Million Word Gap? A LENA Study from Rural and Peri-Urban China.” *Working Paper*, Yue Ma, Lucy Pappas, Andrew Rule, Xinwu Zhang, Yujuan Gao, Sarah-Eve Dill, Tianli Feng, Yue Zhang, Hong Wang*, Flavio Cunha, and Scott Rozelle, .

National Bureau of Statistics of China. 2022. “China Statistical Yearbook 2022.” China Statistical Yearbook. 2022. http://www.stats.gov.cn/sj/ndsj/2022/indexeh.htm.

Naughton, Barry. 1993. “Deng Xiaoping: The Economist.” *The China Quarterly*, no. 135: 491–514.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 187–204.

Preacher, Kristopher J., and Andrew F. Hayes. 2008. “Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models.” *Behavior Research Methods* 40 (3): 879–91. https://doi.org/10.3758/BRM.40.3.879.

Qin, Xuezheng, Tianyu Wang, and Castiel Chen Zhuang. 2016. “Intergenerational Transfer of Human Capital and Its Impact on Income Mobility: Evidence from China.” *China Economic Review* 38 (April): 306–21. https://doi.org/10.1016/j.chieco.2014.10.005.

Rozelle, Scott, and Natalie Hell. 2020. *Invisible China: How the Urban-Rural Divide Threatens China's Rise*. Chicago, IL: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/I/bo61544815.html.

Rubio-Codina, Marta, M. Caridad Araujo, Orazio Attanasio, Pablo Muñoz, and Sally Grantham-McGregor. 2016. "Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies." *PloS One* 11 (8): e0160962. https://doi.org/10.1371/journal.pone.0160962.

Rubio-Codina, Marta, Orazio Attanasio, and Sally Grantham-McGregor. 2016. "Mediating Pathways in the Socio-Economic Gradient of Child Development: Evidence from Children 6-42 Months in Bogota." *International Journal of Behavioral Development* 40 (6): 483–91. https://doi.org/10.1177/0165025415626515.

Rubio-Codina, Marta, Orazio Attanasio, Costas Meghir, Natalia Varela, and Sally Grantham-McGregor. 2015. "The Socioeconomic Gradient of Child Development: Cross-Sectional Evidence from Children 6–42 Months in Bogota." *Journal of Human Resources* 50 (2): 464–83. https://doi.org/10.3368/jhr.50.2.464.

Schady, Norbert, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, et al. 2015. "Wealth Gradients in Early Childhood Cognitive Development in Five Latin American Countries." *Journal of Human Resources* 50 (2): 446–63. https://doi.org/10.3368/jhr.50.2.446.

Singla, Daisy R, Elias Kumbakumba, and Frances E Aboud. 2015. "Effects of a Parenting Intervention to Address Maternal Psychological Wellbeing and Child Development and Growth in Rural Uganda: A Community-Based, Cluster-Randomised Trial." *The Lancet Global Health* 3 (8): e458–69. https://doi.org/10.1016/S2214-109X(15)00099-6.

Sylvia, Sean, Renfu Luo, Jingdong Zhong, Sarah-Eve Dill, Alexis Medina, and Scott Rozelle. 2022. "Passive versus Active Service Delivery: Comparing the Effects of Two Parenting Interventions on Early Cognitive Development in Rural China." *World Development* 149 (January): 105686. https://doi.org/10.1016/j.worlddev.2021.105686.

Sylvia, Sean, Nele Warrinnier, Renfu Luo, Ai Yue, Orazio Attanasio, Alexis Medina, and Scott Rozelle. 2021. "From Quantity to Quality: Delivering a Home-Based Parenting Intervention Through China's Family Planning Cadres." *The Economic Journal* 131 (635): 1365–1400. https://doi.org/10.1093/ej/ueaa114.

Walker, Susan P., Susan M. Chang, Marcos Vera-Hernández, and Sally Grantham-McGregor. 2011. "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior." *Pediatrics* 127 (5): 849–57. https://doi.org/10.1542/peds.2010-2231.

Walker, Susan P., Susan M. Chang, Amika S. Wright, Rodrigo Pinto, James J. Heckman, and Sally M. Grantham-McGregor. 2022. "Cognitive, Psychosocial, and Behaviour Gains at Age 31 Years from the Jamaica Early Childhood Stimulation Trial." *Journal of Child Psychology and Psychiatry* 63 (6): 626–35. https://doi.org/10.1111/jcpp.13499.

Wang, Boya, Xia Luo, Ai Yue, Lei Tang, and Yaojiang Shi. 2022. "Family Environment In Rural China And The Link With Early Childhood Development." *Early Child Development and Care* 192 (4): 617–30. https://doi.org/10.1080/03004430.2020.1784890.

Wang, Lei, Dorien Emmers, Sean Sylvia, Yue Xian, Yu Bai, and Scott Rozelle. 2023. "Rural-Urban Differences in the Intergenerational Transmission of Cognitive Capabilities in China." SCCEI Working Paper.

Wang, Lei, Wilson Liang, Siqi Zhang, Laura Jonsson, Mengjie Li, Cordelia Yu, Yonglei Sun, et al. 2019. “Are Infant/Toddler Developmental Delays a Problem across Rural China?” *Journal of Comparative Economics* 47 (2): 458–69. https://doi.org/10.1016/j.jce.2019.02.003.

Wang, Xiaobing, and Terry Peach. 2019. “Understanding China’s Economic Development: Retrospect and Prospect.” *The Manchester School* 87 (2): 135–37. https://doi.org/10.1111/manc.12267.

World Bank. 2022. “GDP per Capita (Current US$) | Data.” 2022. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

Xi, Jinping. 2022. “Hold High the Great Banner of Socialism with Chinese Characteristics and Strive in Univty to Build a Modern Socialist Country in All Respects.” Report to the 20th National Congress of the Communist Party of China (Official English translation). Beijing: Xinhua News Agency. https://english.news.cn/20221025/8eb6f5239f984f01a2bc45b5b5db0c51/c.html.

Xu, Shanshan, Hong Huang, Jinsong Zhang, Xiaoyan Bian, Na Lü, Yingbo Lü, and Yao Chen. 2011. “Research on the Applicability of Bayley Scales of Infant and Toddler Development-Third Edition to Assess the Development of Infants and Toddlers in Shanghai.” *Chinese Journal of Child Health Care* 19 (1): 30–32.

Yang, Xintong, and Li Gan. 2020. “Bequest Motive, Household Portfolio Choice, and Wealth Inequality in Urban China.” *China Economic Review* 60 (April): 101399. https://doi.org/10.1016/j.chieco.2019.101399.

Yi, Shourong. 1995. *User’s Manual of the Bayley Scales of Infant Development (Revised Edition for Urban China)*. Hunan, China: Hunan Medical University.

Young, Natalie A. E., and Emily C. Hannum. 2018. “Childhood Inequality in China: Evidence from Recent Survey Data (2012–2014).” *The China Quarterly* 236 (December): 1063–87. https://doi.org/10.1017/S0305741018001303.

Yousafzai, Aisha K, Muneera A Rasheed, Arjumand Rizvi, Robert Armstrong, and Zulfiqar A Bhutta. 2014. “Effect of Integrated Responsive Stimulation and Nutrition Interventions in the Lady Health Worker Programme in Pakistan on Child Development, Growth, and Health Outcomes: A Cluster-Randomised Factorial Effectiveness Trial.” *The Lancet* 384 (9950): 1282–93. https://doi.org/10.1016/S0140-6736(14)60455-4.

Yue, Ai, Qi Jiang, Biaoyue Wang, Cody Abbey, Alexis Medina, Yaojiang Shi, and Scott Rozelle. 2019. “Concurrent Validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development III in China.” *PLoS ONE* 14 (9): e0221675. https://doi.org/10.1371/journal.pone.0221675.

Yue, Ai, Yaojiang Shi, Renfu Luo, Jamie Chen, James Garth, Jimmy Zhang, Alexis Medina, Sarah Kotb, and Scott Rozelle. 2017. “China’s Invisible Crisis: Cognitive Delays among Rural Toddlers and the Absence of Modern Parenting.” *The China Journal* 78 (July): 50–80. https://doi.org/10.1086/692290.

Zhang, Jun, Ying Tian, Weiye Wang, Fengxiu Ouyang, Jian Xu, Xiaodan Yu, Zhongcheng Luo, Fan Jiang, Hong Huang, and Xiaoming Shen. 2019. “Cohort Profile: The Shanghai Birth Cohort.” *International Journal of Epidemiology* 48 (1): 21–21g. https://doi.org/10.1093/ije/dyy277.

Zhang, Junsen. 2021. “A Survey on Income Inequality in China.” *Journal of Economic Literature* 59 (4): 1191–1239. https://doi.org/DOI: 10.1257/jel.20201495.

Zhang, Yunting, Le Kang, Jin Zhao, Prof. Yingquan Song, Prof. Fan Jiang, and Chunling Lu. 2021. "Assessing the Inequality of Early Child Development in China - A Population-Based Study." *The Lancet Regional Health - Western Pacific* 14 (September): 100221. https://doi.org/10.1016/j.lanwpc.2021.100221.