Preschool and Child Health: Evidence from China's Universal Child Care Reform*

Meiqing Ren†

February 18, 2024

Abstract

Early childhood education programs have been found to effectively promote children's social and cognitive development. However, the health impact of these programs is less understood. Using a quasi-experiment of the first universal child care reform in China from 2010, this paper aims to identify whether the preschool reform produces any short-term effects on health-related outcomes of preschoolers (3-6 years old). Specifically, this reform expands access to affordable preschools that provide full-day center-based education, with school meals and essential health services on campus. I exploit the variation in the number of newly established preschools across provinces and implement difference-in-differences and triple-difference strategies. Results confirm the effectiveness of this reform by showing a strong and positive impact on preschool attendance. This paper then documents the benefits to alleviating underweight among preschoolers. Estimates show a larger effect in rural areas, suggesting that the reform narrows rural-urban disparities in education access and undernutrition prevention. I also explore the impact on caregivers' health consciousness and find improved health-seeking behavior when children get sick.

Keywords: Early childhood education, preschool, health, underweight, health-seeking behavior, China

JEL classifications: H52, H75, J13, I12

*I am grateful to Darren Lubotsky and Ben Feigenberg for their continuing guidance and support. I also thank Joanna Lahey, So Yoon Ahn, Ben Ost, Javaeria Qureshi, Yiqun Chen, Katherine McElroy, Shogher Ohannessian, Vinitha Varghese, and Jianhao Chen for their helpful discussions and comments. Declarations of interest: none.

†Correspondence: Department of Economics, University of Illinois at Chicago, 601 South Morgan, Chicago, IL 60607, USA. Email: mren8@uic.edu.

1 Introduction

Although China has experienced remarkable economic growth in recent years, the provision of quality early childhood education (ECE) remains a major issue of public concern. According to the Ministry of Education of the People's Republic of China (MOE), the national enrollment rate in pre-primary education was only 50.9% in 2009, ranking China behind not only developed countries (e.g., Australia 81.3%, the United Kingdom 79.8%, Finland 66.1%) but also developing countries (e.g., Vietnam 69.4%, Sri Lanka 58.7%, Algeria 80.3%) (UNESCO, 2009). With the growing realization that child care and early interventions are powerful and necessary mechanisms for alleviating poverty and overcoming disparities, China's government issued two key policy documents in 2010, ushering in the nation's first universal child care reform with the aim of increasing the accessibility, affordability, and accountability of preschools.

This paper studies the causal effects of the first universal child care reform in China on preschool attendance rates and health-related outcomes. In particular, this reform improves access to affordable preschools that provide full-day center-based education with meals provided and teachers, caretakers, and healthcare personnel on campus. I utilize this policy experiment to investigate whether preschool reform with the goal of universalizing preschool education produces any short-term effects on health-related outcomes of preschoolers (3-6 years old). I measure the health of children with four indicators, including stunting, underweight, overweight, and being sick in the previous month, all of which are broadly used in health-related research. This paper also looks at a measure of caregivers' health consciousness to examine any potential impact on their health-seeking behaviors: namely, whether the child is sent to the doctor when they become sick. The data comes from the China Family Panel Studies (CFPS), a national representative dataset available from 2010, supplemented with provincial statistics data from the Educational Statistics Yearbook.

This study follows Duflo [2001] and defines the policy intensity as the provincial number of newly established preschools between 2010 and 2016 per 1,000 children in 2010. To evaluate the effect on preschool attendance, I implement a difference-in-differences (DID) strategy that exploits variation across birth cohorts and provinces in policy intensity. It combines differences across provinces in the number of subsidized preschools constructed with differences across cohorts induced by the timing of the reform (i.e., age in 2010). My estimates show that the reform is effective in promoting preschool attendance, with a greater effect on rural children. The reform led to an increase of approximately 300 seats per 1,000 preschool-aged children and an increase of 25.6 percentage points in the preschool attendance rate from 2010 to 2016. Conclusions remain consistent using both a binary (high- and low-intensity via a median split) and a continuous measure of policy intensity. I also conduct placebo tests and find it reassuring that the unexposed cohorts reveal no change in preschool attendance.

After establishing the reform's effectiveness in increasing preschool attendance, this paper subsequently examines the causal effects on preschoolers' health-related outcomes. The outcomes of interest consist of dummy indicators including stunting, underweight, overweight, ever sick last month, and seeing a doctor when children get sick. The focus of this study is to explore the short-run effects on children's health, highlighting the potential for interventions to enhance early childhood development during their preschool years. Thus, instead of exploiting variation across birth cohorts, I leverage indicators of survey year as well as child age in each survey year to determine treatment status in this section. Given that each province started to develop its 3-year Action Plan in 2011, the year 2010 serves as the year before the expansion of preschool facilities, and the years 2012, 2014, and 2016 are treated as the years after the expansion. Utilizing a DID strategy, I track the health indicators for preschoolers in high-intensity provinces before and after the policy implementation and compare these changes with the corresponding changes for preschoolers in low-intensity provinces. Considering that several issues (e.g., pre-reform differential trends) might bias my DID estimates, I further develop a triple-difference (DDD) model adding another dimension of differencing between preschoolers (aged 3-6) and toddlers (aged 1-2). When I compare the change in preschoolers' health-related outcomes before and after the introduction of the reform, in high-intensity provinces and low-intensity provinces, the DDD estimates reflect a strong and robust impact on reducing the rate of underweight among preschoolers relative to that of toddlers. Specifically, preschoolers from provinces with a large expansion in subsidized preschools are 4.7 percentage points (33%) less likely to be underweight than their counterparts from provinces with only a small change in access to preschool education. My estimates are robust to different estimation approaches (DID and DDD), different control variables, different sample restrictions (e.g., children without inter-province migration and siblings), and different choices of intensity measure (binary and continuous measures of policy intensity).[1] In addition, I conduct several placebo tests with respect to unrelated samples and find null effects. Further investigation provides some suggestive evidence of a larger effect on children residing in rural areas than urban areas. Results show that stunting, overweight, and the probability of being sick are unaffected.

I also detect some evidence of caregivers' health-seeking behavioral response to the reform when handling the illness of preschoolers. Although caregivers' health-seeking behavior response can serve as a potential mechanism for child's physical health, the former itself is an important outcome worth attention, given that poor or delayed care-seeking accounts for up to 70% of all under-five child deaths worldwide (WHO). Child health and parental investment both affect each other, and parental behaviors are acknowledged to be major determinants of child development (Attanasio et al. [2020a]). Results show that the reform has encouraged caregivers of preschoolers to refer sick children to a doctor instead of finding medicines to treat the children themselves.

The universal child care reform may have beneficial effects on preschoolers' health and caregivers' medical-seeking behavior through several potential mechanisms, such as nutritional meals provided in preschool; outdoor activities and healthy eating habits established in school; interactions between school health personnel and children's families; and increased parental employment and disposable income (e.g., Baker et al. [2008]; Bauernschuster and Schlotter [2015]). Food provided by schools can have a direct effect on reducing children's underweight that may come from food deficiencies and imbalances of nutrients. At the same time, this universal preschool reform has been found to have a strong positive effect on promoting mothers' entrepreneurial activities

[1]The DDD strategy for health-related outcomes is my preferred specification in this analysis. More details are discussed in Section 4.3.

and grandmothers' labor force participation (Lin and Wang [2019]; Wang and Lin [2019]). Therefore, the preschool reform may indirectly affect children's prevalence of underweight and caregivers' medical-seeking behavior through increased employment and purchasing power in food consumption and healthcare services. However, my analysis of potential mechanisms shows a negative and insignificant effect on household food expenditure, suggesting that the main mechanism behind the decline in underweight may be nutritional meals and general child care in school, instead of the increased parental employment and income. In this study, I do not find any significant impact on the probability of sickness, indicating that the observed rise in doctor visits is not driven by increased actual sickness among preschoolers.[2]

Early childhood education (ECE) programs have been found to provide wide-ranging benefits to children, families, and society as a whole. Numerous studies have documented that ECE programs can be effective in promoting children's social and cognitive development (e.g., Burchinal et al. [2000];Peisner-Feinberg et al. [2001]; Anderson et al. [2003];Burger [2010]; Rao et al. [2019]; Cascio [2023]). Studies in China find that preschool education has positive impacts on children's readiness, mathematics, self-regulation, and language development in the first several years of primary school as well as long-term benefits to social skills (Luo et al. [2012]; Rao et al. [2012]; Wu et al. [2012]; Liu et al. [2013]; Gong et al. [2016]). However, less research has investigated the health impacts associated with ECE programs, with most evidence focusing on the increased risk of infectious diseases and injury in the short term (Ball et al. [2002]; D'Onise et al. [2010]). A few recent studies investigate the effects on various measures of child health with a focus on programs targeted at disadvantaged groups (e.g., Perry Preschool, Abecedarian, and Head Start reforms in the United States). For instance, Lumeng et al. [2015] and Ansari et al. [2015] have examined the health impacts of the Head Start program, the largest federally funded preschool program in the United States for children up to age five from low-income families, and they find that participation in the Head Start program is associated with improvements in health, such as healthier BMI, reduced overweight, and

[2] There is still a possibility that the observed increase in doctor visits when dealing with a child's sickness captures the severity of a child's illness. In other words, children may experience more severe diseases while at school, causing caregivers to seek medical help. However, I rule out this pathway by demonstrating null effects on children's hospitalization prevalence. More details are discussed in Section 4.3.4.

reduced underweight, especially for preschoolers with an unhealthy weight status. Similarly, India's Integrated Childhood Development Services (ICDS), a government program that provides nutritional supplements, healthcare services (e.g., vaccinations and health checkups), and non-formal preschool education to targeted children younger than six from the most vulnerable and economically disadvantaged households, has been found to reduce prevalence of malnutrition (e.g., underweight and stunting) (Kandpal [2011]; Jain [2015]; Ravindran [2021]; Ganimian et al. [2021]).

Despite the sparse evidence pointing to potential health impacts, findings from targeted interventions have limited applicability to universal ECE reforms across larger populations. Instead of examining targeted or small-scale interventions, only a handful of studies make use of universal ECE reforms as natural experiments to explore their potential impacts on child health. Study by van den Berg and Siflinger [2022] estimates the health effects of a day-care reform in Sweden that reduced daycare fees. Their paper shows an immediate adverse effect on physical health (e.g., an increase in infections, ear problems, and medical visits) at younger age, which fades out as children's immunity develops. Similarly, Bosque-Mercader [2022] explores the long-term health effects associated with a Spanish universal preschool program, which entailed the expansion of public preschool education for three-year-olds in 1991. She analyses young adults between ages 11 and 27 and finds null effects from universal preschool on most of measured outcomes in health and healthcare use. With mixed findings in developed countries, the health impacts of universal child care in developing countries have not been well explored in the literature. This study seeks to remedy this gap, which is particularly important because children in developing countries are more likely to be exposed to malnutrition, poor access to safe playgrounds, and child neglect.

This paper contributes to the literature in three ways. First, it extends the broad literature on the relationship between preschool and child development by examining the impact on preschoolers' health-related outcomes. My paper highlights the important role of child health, recognizing that poor health at early ages can have detrimental effects on future cognitive development and human capital formation (Attanasio et al. [2017];Attanasio et al. [2020b]). Child health is particularly relevant in the context of low-income and middle-income countries where poor health at a young age is more prevalent and promoting child development with limited resources remains a key policy issue. Closely related to my study, Nguyen [2022] documents that longer exposure to preschool in Vietnam brings health benefits to children, reflected in reduced prevalence of underweight and stunting. And he concludes that the main channel behind his findings is due to nutrition provided by preschools. My study adds to the evidence by bringing in evidence from China, which helps identify common patterns in developing countries. In China, Wu et al. [2012] find a positive correlation between preschool enrollment and preschoolers' weight, height, and cognitive development based on a household survey conducted in 2010 in Hunan Province of China. However, their findings from Hunan are only suggestive, while this paper conducts a nationally representative analysis. Moreover, the natural experiment of the universal preschool reform in China allows me to investigate differential response patterns of children from both disadvantaged and advantaged backgrounds, in contrast to targeted reforms with limited applicability.

Second, this study is the first to rigorously examine the causal effects of this universal preschool reform in China on preschool attendance. Although previous literature has documented that this reform has promoted mothers' entrepreneurial activities and grandmothers' labor force participation (Lin and Wang [2019]; Wang and Lin [2019]), it is unknown at the first stage how effective this child care reform is in promoting preschool attendance and how it affects children heterogeneously across demographic groups. Finally, this is the first paper to examine caregivers' health-seeking behavioral response to preschool education. My study highlights the important role of child healthcare, recognizing that child health and parental investment in healthcare can have an impact on one another (Attanasio et al. [2017]; Attanasio et al. [2020a]; Attanasio et al. [2020b]).

This paper introduces background information about this reform in Section 2 and data sources used to analyze it in Section 3. Next, Section 4 provides an overview of the empirical strategies used in this analysis and presents the main results. Finally, Section 5 concludes with a discussion of future research.

2 Background

2.1 Preschool education in China

In this paper, the term "preschool" refers to all early childhood education (ECE) programs for children between the ages of 3 and 6 in China.[3] There have been two major center-based ECE programs in China since the 1980s: kindergartens and pre-primary classes. Kindergartens have independent premises and provide full-day care and education to children between 3 and 6, or sometimes 7 years of age before formal schooling. Pre-primary classes, which are usually attached to rural primary schools, offer a one-year pre-primary program for literacy and numeracy preparation before children start the first grade in primary school (Rao et al. [2012]). It caters to the needs of children from 5 to 6, or sometimes 7 before formal schooling. Preschool attendance is voluntary, followed by six years of compulsory education in primary school, typically starting at the age of 6 or 7, and another three years of compulsory education in junior secondary education (junior middle school). In this paper, I will focus on preschool attendance generally without distinguishing the differences between kindergarten and pre-primary classes.

There are mainly three types of preschool providers, including the government, private providers, and work units (state-owned and public enterprises). Work units usually provide child care to a certain group of parents that work and reside in work-related comprehensive communities.

In China, preschool was underdeveloped before the launch of the reform. Why did so few children participate in preschools in China? Both the availability and affordability of services proved challenging. First, insufficient numbers of facilities have hindered access to educational resources for preschool children. In 1978, China launched the transition from a centrally planned to a market economy, which unintentionally led to a dramatic decline in publicly funded child care reforms. The number of preschools decreased by 28.5% between 1997 and 2006 due to the economic transition (Du and Dong [2013]). Many public child care facilities previously run by state-owned enterprises were

[3] This is a definition that is in line with recent studies, such as Rao et al. [2012].

shut down, but some of them had been changed to private service providers, inducing the increase of the private service provision in preschools. As a result, China's ECE system became more privatized and market-oriented before the preschool reform in 2010. However, private markets for child care pose cost barriers for disadvantaged families since the tuition and fees can be relatively expensive and government subsidies had been largely missing before 2010 (Wang and Lin [2019]). This is the second reason why many families, especially low-income ones, would choose to forego preschool education for their children and use informal child care at home instead.

2.2 China's universal child care reform

2010 is a milestone year in the history of ECE development in China. With the growing realization that child care and early interventions are powerful and necessary mechanisms for alleviating poverty and overcoming disparities, China's government made great progress in preschool education policy development by issuing two key policy documents. It was then that the Chinese government decided to devote greater efforts to the development of ECE. In July 2010, China announced the "Outline of China's National Plan for Medium and Long-term Education Reform and Development" (hereinafter referred to as 'Plan'). This Plan provided a blueprint for promoting child development and the modernization of education over the next 10 years. It set forth the goals of universalizing preschool education, reassuring nine-years of compulsory education at primary schools and junior secondary schools, and improving the enrollment rate in senior high school and higher education. To better implement preschool strategic goals in the Plan, in November 2010, China's State Council (State Council, 2010) published "Issues Regarding Current Development of Early Childhood Education" (hereinafter referred to as 'Issues'). Issues required every county in the nation to develop one or more phases of a three-year action plan for the development of preschool services, offering a balance between public and private providers. Thus, in 2011 each province began developing its 3-year Action Plan to implement the policy in several phases, beginning with the first 3-year action plan in 2011-2013, followed by the second phase from 2014-2016 and the third phase from 2017-2020.

This is the first universal child care reform in China. One major development mission in the Plan was to realize universal preschool education by the year 2020. In particular, by 2020, the target enrollment rate for those enrolled in 1-year, 2-year, and 3-year preschools was planned to be 95%, 80%, and 70%, respectively (State Council, 2010). Issues encouraged the establishment of public preschools but also provided financial support to private ones. Financially supplemented by the central government, local governments were expected to devote greater efforts to fund, plan, and manage preschools. In general, the goal was to promote the accessibility, affordability, and accountability of preschools. Accordingly, there have been three important strategies implemented under the guidelines of the Plan and Issues.

First, the government pledged a substantial increase in funding to strengthen the preschool service provision. Governmental funding for preschool development has rapidly increased since 2010, accounting for 7.2% of the entire education budget in 2016, compared with 1.48% in 2009 and 3.7% in 2010. As illustrated in Figure 1, the number of preschools in mainland China increased from 150,420 in 2010 to 266,677 in 2018, a 77% rise in eight years. The number of educational personnel in preschools experienced a 145% rise from 1,849,301 in 2010 to 4,531,454 in 2018. The increase in funding from the central government was more intense in provinces with underdeveloped child care systems. Figure 2 shows that the policy intensity was negatively related to the baseline preschool enrollment in 2009. This negative relationship reflects the reform's provision that more preschools were built in provinces where baseline enrollment rates were low. In addition, the policy intensity was greater in less developed provinces and lesser in more developed provinces, as shown in Figure 3, which reveals a negative relationship between the policy intensity and pre-reform provincial GDP per capita. The universal preschool reform also reduced the urban-rural gaps in preschool supply by increasing school construction to a larger extent in rural areas, as illustrated in Figure 4. Therefore, provinces with a less developed child care system, a lower level of economic development, and a greater proportion of rural residents experienced a more substantial increase in access to affordable child care. As a result of the school expansion reform, preschools have become more accessible to the public.

Second, school enrollment is subsidized, enhancing the affordability of preschool education. Before 1997, most kindergartens in China were publicly run by the government and government institutions, state-owned and public enterprises, and even neighborhoods and rural collectives (Xiaodong [2008]). Due to the economic transition in the early 1990s, a substantial number of private preschools have been established. Under the guidelines of the Plan and Issues, the Chinese government supports the development of each type of "universal preschool". The government has vigorously increased funding in supporting the establishment of public-run preschools with subsidized prices. Some provinces, such as Shaanxi, have begun offering one-year and three-year free public preschools since 2011 (Shaanxi Government, 2011). In addition, the Chinese government uses a variety of strategies to support private preschools so that they are run by the market but charge low tuition fees that are close to those of public- or enterprise-operated facilities (Wang and Lin [2019]). For instance, Beijing's government provides private schools with subsidies for rent, educational toys, and facility reconstructions, and each enrolled student in Beijing receives 1,200 CNY per year to cover tuition and fees (Beijing Bureau of Finance, 2011).

Last, the universal preschool reform emphasizes that early childhood education is critical for young children's health, habit formation, and intellectual development. Under the guidelines of the Plan and Issues, the Chinese government called for actions to issue standards and guidelines for the utilization of space, healthcare delivery and education, and qualification requirements of preschool educational personnel. The Chinese government also launched short-term national training projects for preschool heads in 2011 and 2013 and required all local governments to enhance training for preschool teachers before 2015 to ensure their professionalization. However, there were minimal trend changes in quality indicators (i.e., teacher qualifications measured by their education level, school quality measured by the student/teacher ratio, as well as school quality measured by the student/staff ratio) before and after the implementation of the reform, as illustrated in Figures 5, 6, and 7. These findings suggest that the reform's accountability aspect may not effectively lead to notable improvements in the quality of preschools. In line with previous papers studying the same reform in China (e.g., Lin and Wang [2019] and Wang and Lin [2019]), my paper seeks to assess the effects of the scale of preschool construction that aims to increase access to affordable preschool education, capturing the key policy component of this reform.

2.3 Child's health development

The nutritional status of children has been one of the most important global public health issues. Despite the impressive and sustained socioeconomic development in China over the past two decades, inadequate nutrition remains a serious problem in China's poor households due to poverty and food insecurity, especially in rural areas and in western areas. The prevalence of under-five stunting in China has significantly declined from 32.3% in 1990 to 17.8% in 2000 (The World Bank). As of 2013, the national prevalence of under-five stunting was 8.1%, which is less than the developing country average of 25% but still much higher than that of developed countries such as the United States (2.1% in 2012). Research by Shen et al. [2015] suggests that the prevalence of food insecurity was 6.1% among rural elementary school students aged 6-14 years, with 16.3% experiencing severe malnutrition. Generally, undernutrition remains a public health challenge in China and shows a significant difference separating Western and Eastern provinces and rural and urban areas. At the same time, China has also experienced a shift toward increased obesity among children, bearing the dual burden of undernutrition and overnutrition. From 1985 to 2015, the possible risk of being overweight increased from 6.51% to 12.57% among children between the ages of 2 and 7 (Zong et al. [2019]). For older children, the prevalence of overweight and obesity was even higher (19.4% in 2014 among 7- to 18-year-old Chinese children) (Wang et al. [2017]). A potential cause behind this increase is that Chinese children have easier access to fast food and are less physically active, along with the rapid urbanization in China (Ji and Cheng [2008]; Van de Poel et al. [2009]). Given that overweight in childhood can be carried into adulthood and increases the risk of type 2 diabetes, cardiovascular disease, and some cancers (Pulgaron and Delamater [2014]), the burden of overweight and obesity needs special attention, signaling the importance of strong nutritional governance.

The universal preschool reform under the Plans and Issues can potentially serve as a strategy to improve children's nutrition status and physical health through nutritious meals, outdoor play, and healthy eating habits in school. The typical daily schedule of preschoolers starts at around eight o'clock when they arrive at school. After breakfast on campus, children alternate between class sessions and free playtime, have a nutritious lunch at the scheduled time, take a long nap, have some snacks and fruit, and then go back to class sessions and playtime. At around five or six o'clock, families, usually parents or grandparents, pick the children up from school. In this process, a variety of educational personnel takes the responsibility to provide a safe and hygienic environment, ensure a healthy and regular diet and work to help children develop in a healthy and happy manner.

Center-based preschools often have health care teachers, caretakers, and sometimes school doctors (medical college-trained doctors) on campus to provide health-related services. Their duties include, but are not limited to, health screenings, maintaining a hygienic environment, planning nutritious meals, and detecting sick or injured children as early as possible. When a child gets sick or injured in school, teachers and caretakers will notify the child's families immediately and suggest further treatment, such as visiting a doctor and taking sick leave. Guided by an action plan brief report from China's Ministry of Education (MOE), they also cooperate with local medical workers to deliver health knowledge lectures to children's families occasionally during preschool publicity month (May 20 to June 20, MOE). Common interactions between school health personnel and children's families suggest a potential mechanism that attending preschool may yield some impacts on families' health-seeking behavior when their children get sick. Under the guidelines of the Plans and Issues, the number of school health personnel and others (e.g., security guards) nationwide has more than tripled after the initiation of the reform, increasing from 543,990 in 2010 to 1,657,945 in 2018 (MOE, 2018).

3 Data

3.1 Data source

My research draws on a nationally representative data set, the China Family Panel Survey (the CFPS), supplemented with provincial statistics data from the Educational Statistics Yearbook of China. Designed and launched by Peking University, CFPS is a longitudinal social survey conducted every two years since 2010. Its baseline sample covers 15,717 households, 33,600 adult respondents (age 16 or above), and 8999 children (aged 0–15) in 25 provinces, representing 94.5% of the total population in China. The statistics for this study are derived from the 2010, 2012, 2014, and 2016 waves.[4] The CFPS collects individual-, family-, and community-level longitudinal data biannually to keep track of China's economic development and social change. All household members older than 15 had in-person questionnaires. There are direct interviews with children between 10 and 15 years old, and data from caregiver reports are also available for all children below 15 years old.

In addition to the CFPS individual dataset, information about provincial school constructions is obtained from the Educational Statistics Yearbook of China published by the Ministry of Education of China. To exploit the variation in policy intensity across provinces and over the years, this study follows Duflo [2001] and defines the policy intensity as the provincial number of newly established preschools between 2010 and 2016 per 1,000 children in 2010. It is easier to interpret the findings if the policy intensity is measured by dividing the number of newly constructed preschools by the number of preschool-aged children (3-6 years old) in each province. Due to data limitations, the number of preschool-aged children in each province is unavailable. Therefore, this paper instead utilizes the number of children aged 0-14 in each province at the baseline (2010) from the Educational

[4]Thus far, six waves of the CFPS (2010, 2012, 2014, 2016, 2018, and 2020) have been made available. This study focuses on the time period until 2016, allowing me to evaluate the policy effect at a 6-year interval following the first phase (2011-2013) and the second phase (2014-2016) of 3-year action plans. It helps ensure consistency across waves since CFPS made changes to its weighting variables from population weights to normalized weights from the 2018 wave. More importantly, analysis from the 2010–2016 waves of CFPS helps avoid excessive post-treatment years that might introduce potential biases from external factors (e.g., China's one-child policy ended in 2016 and COVID-19 outbreak in 2019), given that there is only one pre-treatment period prior to 2011, with 2010 being the first wave available in the CFPS survey dataset.

Statistics Yearbook of China to construct the intensity measure. In the empirical analysis, I create a binary indicator of policy intensity to allow for clear intuitions of the estimated effects. Provinces are classified into high-intensity and low-intensity provinces depending on whether the provincial policy intensity is larger than the national median. To assess the sensitivity of my findings to the choice of intensity measure, I also conduct additional analyses using a continuous intensity measure and explore whether conclusions remain consistent.[5]

I analyze the effects of the universal preschool reform employing four waves of the CFPS data (2010, 2012, 2014, 2016). Children aged 0–15 are studied in the paper, with preschool attendance, physical health indicators, and a measure of medical-seeking behavior being the dependent variables. The unit of the analysis is a child.

3.2 Outcome measures

3.2.1 Preschool attendance

In the CFPS survey, the caregiver of children reports whether a child is currently attending preschool or had ever attended preschool. As indicated above, neither the CFPS survey nor my research design distinguishes between kindergartens and pre-primary classes. Thus, my dependent variable of interest is a dummy variable reflecting whether this child is attending/has attended any type of preschool.[6] When assessing the impact of the reform on preschool attendance, I restrict my analysis to children with non-missing information about preschool attendance, which leaves an effective sample of 10,297 unique participants between 3-15 years old.

[5] In the empirical analysis, I use both binary and continuous measures of policy intensity. The binary measure, which equals 1 if the intensity is larger than the national median and 0 otherwise, enables a simple comparison in means between children from the high- and low-intensity provinces (discussed in Section 4.1). When I use the continuous measure of policy intensity, a value of 1 in intensity indicates that 1 more subsidized preschool per 1,000 children (0-14 years old) was established after the reform.

[6] As previously discussed, one major development mission of the preschool reform in China is to realize the target enrollment rate (95%, 80%, and 70% for those taking 1-year, 2-year, and 3-year preschools, respectively) by 2020. The preschool measure in this paper is similar to a 1-year preschool enrollment indicator because the preschool measure is equal to 1 if a child had ever attended any type of preschool, no matter the length of the enrollment.

3.2.2 Physical health

Physical health is the foundation of children's comprehensive development, and it influences their long-term well-being. Physical health refers to the biological status of a child, including his/her overall health and functioning, BMI, and involvement in healthy lifestyle behaviors. I construct four indicators from the CFPS survey data, including stunting, underweight, overweight, and the incidence of sickness, to examine the effects of preschool reform on children's physical health. Based on age in months, gender, height, and weight measurements, I calculate the height-for-age, weight-for-age, and BMI-for-age z-scores (standard deviation scores) according to the World Health Organization (WHO) 2006 growth standards for children aged 0-5 years and WHO 2007 growth reference for children aged 5-19 years.[7] These standards enable us to assess child nutrition status relative to the reference population of the same age and sex. Undernutrition, widely measured by stunting and underweight indicators, has adverse functional consequences on children in several domains, such as morbidity risk and cognitive capacity. Stunting, an indicator of chronic malnutrition, is defined as somebody with a height-for-age z-score less than -2 standard deviations (SD) compared with the reference median of same age and sex (HAZ<-2 SD). Based on WHO standards, weight-for-age z-score (WAZ)<-2 SD is termed as underweight. BMI-for-age z-score (BAZ), an indicator for body fat, is useful to screen for overweight and obesity, a global trend that has been extensively documented (Popkin et al. [2006]). Therefore, I construct the overweight measure according to BAZ to evaluate the reform effects on the spreading of overweight in China. Based on references developed by the WHO, school-aged children and adolescents aged 5-15 are defined as overweight if their BAZ is greater than 1 (BAZ>+1 SD). However, WHO defines overweight for young children with a different standard so that children below 5 years old are overweight if their BAZ is more than two standard deviations above the median (BAZ>+2 SD). [8]

[7] I make use of publicly available R-Shiny tools provided by the Canadian Pediatric Endocrine Group (CPEG) to calculate the above-mentioned anthropometric z-scores. For more details on the use of Shiny, see https://cpeg-gcep.net/content/shiny-apps.

[8] In the CFPS survey, information about height and weight is reported by caregivers, which has the potential to introduce measurement errors into z-score calculations. Although implausible z-scores should have limited effect on the stunting, underweight, and overweight indicators, I also explore estimates when observations flagged for biologically implausible z-scores based on WHO standards are dropped and find consistent results. For example, WAZ<-6 SD or

Additionally, caregivers are asked whether their child became sick within the last month. This self-reported measure of individual morbidity, an important health indicator, refers to the incidence rate of sickness and helps to explore children's health status by measuring, quantifying, and comparing illness and disease in individuals (Murray and Chen [1992]; Choi et al. [2019]). Thus, I add the indicator of being sick last month to my analysis to explore the policy effects on a child's individual morbidity.

3.2.3 Caregiver's health-seeking behavior

Under the reform, preschools are guided to acknowledge children's development and use scientific teaching methods to ensure that children develop in a healthy and happy manner. With an increased number of school health personnel providing health-related services on campus, preschool may introduce health benefits to preschoolers and also gradually change their caregivers' health-seeking behavior. The CFPS survey for caregivers asked how they dealt with their child's illnesses. Options include seeing a doctor right away, finding/buying medicine by oneself, or choosing other options such as asking for religious help or taking no action and waiting for a natural recovery. In the baseline year 2010, seeing a doctor right away is the first preferred choice, reported by nearly 70% caregivers of preschoolers, while finding/buying medicine by themselves constitutes an alternative with 27% voters. To understand the impact of China's preschool reforms on children's health, the last variable of interest in my analysis is preschoolers' health-seeking behaviors, which reflects the health consciousness of caregivers. It is constructed as a dummy variable indicating whether a caregiver takes their children to see a doctor right away when the children get sick.

To highlight the importance of the health-seeking behaviors of caregivers of young children, it is important to provide context for the medical development and environment in China. Thanks to improved healthcare conditions, big countries like Brazil and China have reduced their child mortality rates 10-fold over the last 4 decades. In China, the under-five child mortality rate has fallen from 15.8

WAZ>5 SD are defined as biologically implausible values for children ages 0–5 years, and my estimates are similar when these observations (less than 1% in my sample) are dropped. Results can be provided upon request.

deaths per 1,000 live births in 2010 to 8.6 in 2018, compared with a decrease from 7.3 in 2010 to 6.5 in 2018 in the US (UNICEF). Noting that poor or delayed care-seeking contributes to up to 70% of all under-five child deaths worldwide (WHO), the motivation for care-seeking delivers significant benefits to the child. This is especially relevant in China, where mistrust of the healthcare system has become a prominent issue.

There are three main issues that discourage Chinese people from seeking healthcare. First, China has been experiencing a widespread and prominent crisis of mistrust in healthcare services; the number of conflicts between patients and physicians in China has increased by 12,000 cases from 2002 to 2010, more than tripled in 8 years (Chan [2018]). The second issue is the lack of health literacy in China, which is highly correlated with lower levels of education and income in rural areas. Nevertheless, if people with sickness decide to find medicines by themselves, such as buying over-the-counter drugs without consulting a healthcare professional or using folk medicine, then they must rely on their own understanding of the illness. This can be dangerous for preschool-aged children if their caregivers incorrectly judge the cause of the sickness and provide incorrect treatment. Last but not least, expensive medical bills and the difficulty of accessing quality medical services build barriers to care seeking. Many disadvantaged residents in China are not only uncovered by healthcare insurance but also have difficulties finding high-quality health facilities near their residences. According to the World Health Organization Global Health Expenditure database, the out-of-pocket expenditure (% of current health expenditure) in China was 40.8% in 2010, higher than many countries like the United States (12.4%) and France (10.2%). This suggests a serious impediment for families with financial constraints to utilize medical services where they may not be able to receive even the most basic healthcare for their children.

The following mechanisms may impact health-seeking behavior when preschoolers get sick under the universal preschool reform. First, school health personnel on campus tend to suggest a doctor visit when a child gets sick, promoting the child's caregiver's decision to see a doctor when his/her children are ill. Second, health knowledge lectures held on campus for caregivers occasionally during preschool publicity month (May 20 to June 20, MOE) may improve their health knowledge gradually, leading to more careful decision-making when dealing with a child's sickness. Another mechanism is through any potential impact on children's health, as evidenced by preschoolers' nutrition status and morbidity. On the one hand, nutrition improvement can reduce child morbidity from infectious diseases. On the other hand, children's health status also depends upon environmental hygiene as well as exposure to infections. Considering this factor, the impact of preschool reform may be reflected in increased children morbidity. Given the two sides of potential effects, this paper examines the net impact of the reform on children's health status, which may trigger their caregivers' health-seeking behavior response. Finally, this universal preschool reform has a strong positive effect on promoting mothers' entrepreneurial activities and grandmothers' labor force participation (Lin and Wang [2019]; Wang and Lin [2019]). Income effects from increased labor supply can lead to increased purchasing power in healthcare services, while substitution effects tend to reduce doctor visits, given escalated time costs. Thus, it is unclear how caregivers' labor supply response has changed their health-seeking decisions when children get sick.

Seeing a doctor right away is not necessarily a better choice than caregivers finding medicine on their own. Self-medication or recuperating at home can also be a rational choice when children manifest only mild symptoms and families with adequate health knowledge are familiar with the medical conditions. However, self-medication has the potential to bring negative consequences such as incorrect diagnosis, inappropriate treatment and dosage, and delayed seeking of proper care. For example, the abuse of antibiotics, a widespread issue associated with self-medication, has become a global public health problem, particularly in developing countries with high levels of antibiotic use and rising drug resistance (Radyowijati and Haak [2003]). China consumes the most antibiotics in the world, and acquiring antibiotics from community pharmacies without a prescription is still common in China, despite being outlawed by China's Food and Drug Administration regulations since 2004 (Fang [2014]; Quan-Cheng et al. [2016]; Zhu et al. [2021]). A recent study by Zhu et al. [2021] documents that 40% of infants and young children under 5 years had used antibiotics for common cough, based on a 2019 survey in a highly populated city in Zhejiang Province of China. According to a survey conducted in Hefei City, China, 31.8% and 34.5% of children with common colds and diarrhoea, respectively, were self-treated by their caregivers with antibiotics, indicating widespread antibiotics abuse (Bi et al. [2000]). Generally, self-medicating behavior by guardians for their children can be dangerous without a reasonable basis for successful self-medication. There are substantial health benefits when caregivers, mostly parents and grandparents, consult health professionals when they have any unanswered questions such as the cause of illness and medication dosages. This is of particular importance with respect to rural families in China with lower levels of education and inadequate health literacy.

4 Methods and results

This section presents my methodology and main empirical results. First, I provide a descriptive picture of the sample used in this paper. Next, I study the effect of establishing subsidized preschools on preschool attendance.[9] I then examine changes in preschoolers' health-related outcomes and how these changes vary by demographic group. Last, I explore the potential mechanisms behind my findings.

4.1 Summary statistics

Table 1 presents summary statistics for the sampled children in the baseline wave before the policy (the 2010 wave) and across all three waves after the policy (2012, 2014, 2016 waves), respectively. I code the policy as being implemented from 2011 because each province started to develop its 3-year Action Plan in 2011. Table 1 includes the mean and the standard deviation of the overall samples and subsamples of children by the policy intensity. The results are weighted with individual-level national sampling weights. In the 2010 wave, there are 7,624 children aged 0-15 with no missing data for the outcome variables, with an average of 43% urban and 55% male.

[9] In this paper, the terms "preschools" and "subsidized preschools" are used interchangeably when discussing newly constructed preschools following the reform. This is because different types of preschools (e.g., private and public preschools) were all supported to offer subsidized prices, as discussed in Section 2.2.

Family members in the CFPS are defined as financially related household members, and the average family size is 4.9 in 2010. Differencing samples in the high-expansion provinces and low-expansion provinces depending on whether the provincial policy intensity is larger than the national median, I find that children in high-intensity provinces tend to be less likely to live in an urban area, have lower household income, and most importantly have a lower preschool attendance rate, which is consistent with the policy design.

The provincial policy intensity is defined as the level change in the number of preschools from 2010 to 2016 per 1,000 children aged 0 to 14 in each province.[10] A value of 1 in intensity indicates that 1 more subsidized preschool per 1,000 children was established after the reform. Assuming that the number of children is evenly distributed by age, a value of 1 in intensity can be interpreted as 1 more preschool per 267 (1,000/15*4) preschool-aged children (3-6 years old). Given that a Chinese preschool, on average, had around 200 enrolled children in 2010,[11] this magnitude is equivalent to offering 200 more seats in preschool per 267 preschool-aged children, or around 750 more seats in preschool per 1,000 preschool-aged children. Based on this, I define a high-intensity dummy to equal 1 if the provincial policy intensity is greater than the national median and 0 otherwise. Thus, by definition, 12 provinces are high-expansion provinces, and 13 provinces are low-expansion ones. As shown in Table 1, an average value of 0.4 in policy intensity in the sample indicates that 0.4 more preschools were newly built per 1,000 children aged 0 to 14 from 2010 to 2016. Moving from low-expansion to high-expansion provinces reflects a 0.29 increase in policy intensity, corresponding to obtaining 217 more seats in preschool per 1,000 preschool-aged children.[12] In the regression analysis, I explore the results of representing the policy intensity as a dummy variable and a continuous variable.

[10] As discussed earlier, the number of preschool-aged children in each province is unavailable. Due to this data limitation, my analysis instead utilizes the number of children aged 0-14 in each province from the Educational Statistics Yearbook of China to construct the intensity measure.

[11] The size of preschools in China varies from 97 enrolled children to 522 enrolled children across different provinces (Educational Statistics Yearbook of China, 2010).

[12] A 0.29 increase in policy intensity means that provinces that experienced a high level of expansion after the reform newly constructed 0.29 more preschools per 1,000 children than provinces that experienced a low level of expansion during 2010-2016. In other words, moving from low-expansion to high-expansion provinces corresponds to obtaining 217 (0.29*200/267*1000=217) more seats in preschool per 1,000 preschool-aged children.

Table 1 also displays that in the 2010 wave, 23% of children (aged 0-15) were stunted, 30% of children were overweight, 17% of were underweight, 29% were sick last month, and 68% of caregivers would take children to see a doctor when they got sick instead of finding medicines by themselves without consulting healthcare professionals.[13]

4.2 Effects on preschool attendance

4.2.1 Empirical strategy

This section displays the empirical strategy used in this study to identify the effect of increasing accessibility to formal child care facilities on children's preschool attendance. The date of birth and the region of education jointly determine an individual's exposure to the reform. Chinese children normally attend preschool between the ages of 3 and 6, while delayed school entry could lead some 7-year-old children to stay in preschool.[14] All children born in 2002 or before were 8 or older in 2010, the year immediately before the expansion of child care facilities. Thus, they generally did not benefit from the reform since they should have left preschools before the implementation of the reform. For younger children, who were between the ages of 2 and 7 in 2010, the exposure is a decreasing function of their age in 2010. Children born in 2009 or after (the calculated age below 1 in 2010) were fully exposed; thus, the reform has potentially imposed a similar effect on these youngest cohorts. In summary, the effect of the reform should be close to 0 for children 8 or older in 2010, the year immediately before the expansion of child care facilities, and increasing for younger children until they reach full exposure.

The region of education is the second dimension of variation in the intensity of the reform.

[13] Chen and Meng [2015] used the CFPS 2012 wave and found that 30% of adult respondents were overweight or obese, and 11% were underweight. Their results are based on the standard specifically designed for China, in which underweight is defined as BMI$<$18.5 and overweight or obese as BMI $\geq$ 24.

[14] The official enrollment age of primary school entry varies across regions and is set at 6 or 7 years old (Rao et al. [2012]). In many areas of China, especially rural ones, children generally start primary school between 6 and 8 years of age (e.g., He et al. [2007]; Chen [2015]; Chen [2017]). Therefore, some children aged 7 in 2010 with delayed school entry could also be partially affected. This study aims to assess the impact of China's child care reform on preschool-aged children, which are generally defined as those aged 3-6 years old following the China ECE literature (e.g., Rao et al. [2012]; Wu et al. [2012]; Gong et al. [2016]).

However, it could be endogenous if families moved to benefit from the reform, which would lead to bias in the reform effect (Rosenzweig and Wolpin [1988]). Selective location is less of an issue in this study because strong institutional constraints imposed by the household registration system (hukou) make interprovincial movement in China difficult.[15] Even when parents move for opportunities, many migrant parents leave their children at home (approximately 60% in 2007) considering financial barriers and the difficulty of enrolling in public preschools for migrant children without a local hukou (Démurger and Xu [2015]). Around 0.9% of children in my sample experienced interprovincial migration during the sample period. To eliminate potential bias from endogenous migration, I code their residing provinces as per their location in 2010 if they appear in the 2010 wave.[16]

I evaluate the effect of this reform on preschool attendance by combining differences across provinces in the number of subsidized preschools constructed with differences across cohorts induced by the timing of the reform. To exploit the variation in treatment intensity across provinces and cohorts, I implement a DID model that yields the following:

$$Attendance_{iskt} = C + \alpha_s + \lambda_k + \omega_t + \{P_s * T_i\}\gamma_1 + \{E_s * \lambda_k\}\delta_1 + X_{iskt}\eta_1 + \varepsilon_{iskt} \quad (1)$$

$Attendance_{iskt}$ indicates whether child i in province s from cohort k is attending/has attended preschool as observed in survey year t. P_s denotes a high-intensity dummy indicating whether this child resided in a province s that experienced a high level of preschool expansion after the reform. The policy intensity is defined as the level change in the number of preschools from 2010 to 2016 per 1,000 children aged 0 to 14. P_s equals 1 if the intensity is larger than the national median and 0 otherwise. T_i is a young-cohort dummy indicating whether the child i belongs to the "young" cohort in the subsample who were young enough to be affected by the preschool reform. The "young" cohort refers to individuals with age 0-6 in 2010 and individuals born in 2011 and 2012 with calculated age -2 and -1 in 2010. They are those who were exposed to the preschool reform between 2010 and 2016. C is

[15] The Chinese government uses the household registration system (hukou) to control internal migration and regulate resource allocation. The rigidity of hukou imposes strong migration restrictions, especially for people who want to relocate from the countryside to a city.

[16] Additionally, I test whether migration responds to the preschool reform. More details are discussed in Section 4.3.

a constant. α_s is a province fixed effect, λ_k is a cohort of birth fixed effect, and ω_t is a survey-year fixed effect. E_s is the pre-reform enrollment rate in 2009, and X_{iskt} is a vector of various individual and household characteristics that may relate to school choice, including age, gender, urban-rural residence status, and family size.[17]

A DID approach generally assumes that in the absence of the reform, the increase in attendance rate between cohorts would not have been systematically different across provinces. It will be violated if the pattern of increased attendance would have been observed even if the reform had no effect. As illustrated, the policy intensity is negatively correlated with the initial level of preschool enrollment. If provinces with a low baseline enrollment level would experience a substantial rise in enrollment in the absence of the reform, the estimated differences in differences will be biased upward due to mean reversion. To alleviate this potential concern, I present specifications that control for the interactions between cohort dummies and the provincial preschool enrollment rate in 2009. To further probe the parallel trends assumption of the DID framework, I perform placebo tests on older cohorts not likely to be affected. In addition, I compute robust standard errors clustered at the province level to account for serial correlation within provinces over time. One concern might arise given the relatively small number of clusters (25 provinces) in my analysis. Specifically, a number of studies have discussed that one can over-reject null hypotheses with few clusters (e.g., Cameron et al. [2008]; Cameron and Miller [2015]; MacKinnon and Webb [2017]; Canay et al. [2021]). To correct for potentially downward biased cluster-robust standard errors in the presence of small number of clusters, I also conduct the wild bootstrap procedure, as proposed by Cameron et al. [2008]. I report p-values obtained with wild cluster bootstrap with 500 replications and Rademacher weights implemented, following Cameron et al. [2008].[18]

[17]It is preferred not to control for variables that could be potentially affected by the reform in my main specifications, such as household income. However, regressions that include household income as controls exert similar estimates. This result can be provided upon request.

[18]Rademacher weights are used in accordance with the best practices for practitioners outlined in Canay et al. [2021]. Their study demonstrates the reliability of wild bootstrap with a few small clusters when the clusters are large.

4.2.2 Results

Table 2 presents the estimated effects on preschool attendance. I compare the attendance history of children who had no exposure to the reform (aged 9 to 13 in 2010) to those of individuals who were exposed (aged -2 to 6 in 2010) in high-intensity provinces and low-intensity provinces. Estimates in Table 2 are positive and statistically significant, indicating a positive impact on preschool attendance under the universal preschool reform.[19] The estimates for the specifications taking into account mean reversions and controlling for other control factors in columns 2-3 are slightly smaller in magnitude than that in column 1 but are still robust and significant. The suggested effect in column 3 of Table 2 is that moving from low-intensity provinces to high-intensity provinces would have increased the preschool attendance rate by 14.3 percentage points (20% compared with the baseline mean) if the child was young enough to experience the reform. The p-values for wild bootstrapped standard errors are reported in squared brackets. My results are robust to wild bootstrapping as the statistical significance level for my main coefficients remains similar.

To assess the sensitivity of my results, I introduce additional control variables as a robustness check and present estimates in Table A1. Although my main specifications control for a range of variables that may influence school choice, there may still be endogeneity problems if provincial intensity remains correlated with pre-existing trends that also influence child education. It raises concerns when areas with a higher demand for female labor might be more likely to invest in child care facilities in order to counteract a negative trend in maternal employment. To mitigate this concern, I conduct additional robustness checks controlling for the interactions between cohort dummies and women's baseline employment rate in each province. In doing so, I allow for different underlying trends in women's employment, depending on the pre-reform employment environment characteristics of the province. After introducing this additional control, my estimates hold steady, as shown in

[19] As discussed, this study aims to assess the impact of China's child care reform on preschool-aged children, which are generally defined as those aged 3-6 years old following the China ECE literature (e.g., Rao et al. [2012]; Wu et al. [2012]; Gong et al. [2016]). Given that some children aged 7 and even a few aged 8 in 2010 with late enrollment in primary school could still potentially benefit from the reform, I choose to exclude children aged 7 to 8 in 2010 from my control group here. My estimates remain positive and significant when those children are included in either the treatment group or the control group, despite a slightly smaller magnitude. Results can be provided upon request.

column 1. To further test for robustness, I present specifications that include the interactions between cohort dummies and the baseline number of children in each province, as in Duflo [2001]. To capture the provincial-level dynamic factors that may affect children's school choices, I include provincial characteristics, including provincial GDP per capita and the population each year. To further control for variations in observed or unobserved village-level characteristics, I incorporate village fixed effects via 631 dummy variables reflecting the 631 villages/communities in my sample, which enables better control of time-invariant characteristics specific to each village/community that may be correlated with both the preschool expansion and children's preschool attendance. With better control of spatial heterogeneity, as shown in column 4, my estimates remain positive and significant, despite a slightly smaller magnitude. Overall, my results remain robust after running the battery of specification checks.

To obtain better statistical power, I measure intensity as a continuous variable in the model and run the following regression:

$$Attendance_{iskt} = C + \alpha_s + \lambda_k + \omega_t + \{Intensity_s * T_i\}\gamma_1 + \{E_s * \lambda_k\}\delta_1 + X_{iskt}\eta_1 + \varepsilon_{iskt} \qquad (2)$$

$Intensity_s$ denotes the intensity of the reform in province s. It is defined as the level change in the number of preschools from 2010 to 2016 per 1,000 children aged 0 to 14. C is a constant. α_s is a province fixed effect, λ_k is a cohort of birth fixed effect, and ω_t is a survey-year fixed effect. E_s is the baseline enrollment rate in 2009, and X_{iskt} is a vector of various individual and household characteristics that may relate to school choice.

In Panel A of Table 3, the estimated coefficients reflect that one additional preschool per 1,000 children contributes to an increase in the preschool attendance rate by 64 percentage points. To state it differently, 1 more preschool per 267 preschool-aged children, which is equivalent to offering around 750 more seats in preschool per 1,000 preschoolers, is associated with a 64 percentage point rise in preschool attendance. Since on average the policy intensity is around 0.4, as shown in Table 1, the reform introduced 300 (0.4*750) more seats in subsidized preschools per 1,000 preschool-aged children nationwide. Correspondingly, the estimated effect is a 25.6 (64*0.4) percentage point increase in the preschool attendance rate (defined as attending at least some preschool) in China from 2010 to 2016. Flexibly controlling for interactions of pre-reform enrollment levels with cohort dummies does not change the estimates, as shown in column 2 of Table 3, suggesting that my results are not driven by mean reversion.

Even though it is not feasible to directly test the parallel trends assumption of the DID framework, it is reassuring that one implication of the identification assumption can be tested using a control experiment. Given that children older than 8 in 2010 were not exposed to the reform, I should not expect the increase in preschool attendance between cohorts in this age-group to differ systematically across provinces. I consider two cohorts in the control experiment. Children aged 9 to 13 in 2010 constitute the "young" cohort, and those aged 14 to 15 constitute the "old" cohort. I would expect to see some effects if the pattern of increase in preschool attendance varies systematically across provinces. As suggested in Panel B of Table 3, no significant effect is detected. This helps validate the parallel assumption that treatment and control groups show a nonsignificant difference in preschool attendance in the absence of the reform. The results are robust under different model specifications.

To explore the estimated effects on attendance for each birth cohort respectively, I also show results from an interaction terms analysis. Consider the following:

$$Attendance_{iskt} = C + \alpha_s + \lambda_k + \omega_t + \{Intensity_s * \lambda_k\}\gamma_1 + \{E_s * \lambda_k\}\delta_1 + X_{iskt}\eta_1 + \varepsilon_{iskt} \tag{3}$$

In this specification, I measure the time dimension of exposure to the reform with 18 cohort-of-birth dummies (aged -2 to 15 in 2010). The omitted dummy is the dummy for age 15 in 2010. After controlling for province fixed effects, cohort fixed effects, and survey-year fixed effects, interactions between dummy variables indicating the age of the child in 2010 and the provincial policy intensity are plausibly exogenous. Figure 8 plots the γ_1. Each dot is the coefficient of the interaction between a dummy for being a given age in 2010 and the number of preschools constructed per 1,000 children.

A 95-percent confidence interval is shown in line widths. As expected, the reform had almost zero effect on the attendance of cohorts not exposed to it, and it had a positive effect on the attendance of younger cohorts.[20] The coefficients generally decrease with age for exposed cohorts. Overall, Figure 8 validates that the identification strategy is plausible and that the reform positively influenced preschool attendance.

Estimations from Equation 3 are useful to examine whether γ_1 equals 0 for older cohorts not likely to be affected. To obtain more efficient estimates for each birth cohort, I can instead impose this restriction and run a restricted version of Equation 3. I set the omitted group (the control group) to be children aged 7 to 15 in 2010 so that I force a zero effect on those children, which produces more conservative estimates. Table A2 shows the estimates for each birth cohort. The estimates in column 3 suggest that one preschool per 1,000 children aged 0-14 increases the preschool attendance rate of the cohort aged 3 in 2010 by 52 percentage points. Since on average the policy intensity is around 0.4, as shown in Table 1, the estimated effect on the cohort aged 3 in 2010 is a 20.8 (52*0.4) percentage point increase in the preschool attendance rate. All coefficients are significantly different from zero, and the coefficients generally decrease with age. The results are robust under different model specifications, as shown in Table A2.

4.2.3 Heterogeneity analysis

Rural residents usually face greater financial barriers regarding the decision to send children to preschool than urban residents. Considering this budget constraint, it may be the case that children who reside in rural areas experience a larger effect under the preschool reform. To examine the heterogeneity of my estimated effects by urban-rural residence status (defined based on the administrative district), I estimate Equation 2 for urban and rural children, respectively. As demonstrated in the first two columns of Table 4, the positive effects on preschool are mainly driven by rural chil-

[20] As discussed, the official enrollment age of primary school entry varies across regions and is set at 6 or 7 years old (Rao et al. [2012]). In many areas of China, especially rural ones, children generally start primary school between 6 and 8 years of age (e.g., He et al. [2007]; Chen [2015]; Chen [2017]). Therefore, it is not surprising that children aged 7 in 2010 also experienced significant effects since children between the ages of 7 and 8 can also choose to attend preschool before beginning their first grade in primary school, despite this being less common.

dren. In particular, one preschool per 1,000 children aged 0-14 contributed to an increase in the preschool attendance rate by 58 percentage points in rural areas, while it contributed to an increase of 38 percentage points in urban areas. I also report the p-value of the test of difference between the coefficients for urban and rural groups, and the difference across them is significant at the 5% level. In addition, the effect of child care expansion on preschool attendance may be heterogeneous to the child's gender. Because preference for sons is common in China (and in some other Asian countries like the Republic of Korea and India) (Das Gupta et al. [2003]), there may be differential impacts for girls as compared to boys due to gender-based discrimination. The last two columns of Table 4 present the results of boy and girl subsamples and reveal a slightly greater effect on boys than girls. However, I cannot reject the equality of the coefficients at the conventional levels.

4.3 Effects on health-related outcomes

4.3.1 Empirical strategy

To estimate the causal effects of the universal preschool reform on preschoolers' health-related outcomes, I first utilize a DID strategy. Slightly different from the DID model in the prior section, I track the health indicators for preschoolers in highly-treated provinces before and after the policy implementation and compare these changes with the corresponding changes for preschoolers in provinces that experienced a small change in the supply of subsidized preschools. As discussed before, the focus of this study is to explore the short-run effects on children's health, highlighting the potential for preschool interventions to enhance early childhood development during their preschool years. Emphasizing the short-term effects of the preschool reform helps policy makers to understand the immediate benefits and potential pathways through which preschool interventions can influence children's long-run well-being. Thus, instead of using cohort of birth dummies in previous sections, I leverage indicators of survey year as well as age in each survey year to determine treatment status in this section. Given that each province started to develop its 3-year Action Plan in 2011, the year 2010 serves as the year before the expansion of preschool facilities, and the years 2012, 2014, and

2016 are treated as the years after the expansion. This approach enables me to monitor outcomes measured right after the reform, expecting that intervention effects on health indicators (e.g., underweight and sickness last month) manifest relatively rapidly. The outcomes of interest consist of dummy indicators including stunting, underweight, overweight, ever sick last month, and see a doctor when children get sick. Regression analyses are weighted with CFPS individual-level weights.

I first estimate a DID model of the following form:

$$Y_{ist} = C + \alpha_1 High_s + \alpha_2 Post_t + \gamma\{High_s * Post_t\} + \delta_1 \sum\{E_s * \lambda_t\} + \eta_1 X_{ist} + \varepsilon_{ist} \tag{4}$$

The variable Y_{ist} represents a health-related outcome for child i (a preschooler aged 3-6) residing in province s in year t. $High_s$ denotes a high-intensity dummy indicating whether province s experienced a high level of preschool expansion after the reform. The policy intensity is again defined as the level change in the number of preschools from 2010 to 2016 per 1,000 children aged 0 to 14. $High_s$ again equals one if the intensity is larger than the national median and zero otherwise. $Post_t$ is a post-reform dummy which equals one for the years after 2010 and zero for the year 2010. As mentioned earlier, 2010 is the first wave available in my dataset and is prior to the implementation of the preschool reform. λ_t denotes year dummy, and E_s is the baseline enrollment rate in 2009. X_{isk} is a vector of characteristics that may be related to the child's health outcomes, including age, gender, urban-rural residence status, and family size.

The key coefficient of interest is γ, which is the simple DID estimate of the effect of subsidized preschool expansions. This method takes advantage of the child care reform that focused on expanding educational infrastructure rather than addressing children's health issues. In other words, the allocation of newly constructed preschools is governed by the historical distribution of the child care facilities in a certain province (Wang and Lin [2019]), which is arguably unrelated to a child's current health. The key identifying assumption in the DID model is that provinces with varying levels of policy intensity would have similar trends in outcomes in the absence of treatment. This assumption indicates that differences in children's health-related outcomes observed before and after the reform are (conditionally) attributed to the effect of the preschool reform. This assumption encompasses that families do not anticipate the policy change, in the sense that their pre-reform behavior should not be affected by awareness of the forthcoming reform, which appears plausible in China's political context. Although DID's common trend assumption is fundamentally not testable, I can test if high- and low-intensity provinces had the same trends prior to 2010 as compelling evidence. Ideally, I want to probe pre-existing time trends in the outcomes of interest. However, 2010 is the first wave available in the CFPS survey dataset, so I explored other data sources to show pre-existing trends. Specifically, I obtained moderate to severe malnutrition rates in each province for children below 5 years old from the China's Health and Family Planning Statistics Book. Figure 9 shows that before 2010, the malnutrition rate in provinces with a higher level of established subsidized preschools moved together with provinces that experienced less intense treatment. To probe the validity of DID assumptions, I also check if the estimated impact of the reform on toddlers (aged 1-2) yields a zero coefficient. An estimate close to zero is the most compelling evidence since toddlers are too young to be affected by the preschool reform in the current year.

One may be concerned that several issues can bias my DID estimates. Besides pre-reform differential trends, concerns about biased estimates remain if other confounding factors and policies happen at the same time while affecting provinces differently. For instance, this situation might occur in cases where the government grants funding to rural regions to improve health outcomes; in such cases, the intervention's impact on individuals might be more significant in our treatment provinces than in the control provinces since the treatment provinces are comprised of more rural areas. To address such possible confounding factors, I develop a DDD model adding another dimension of differencing between preschoolers (aged 3-6) and toddlers (aged 1-2).

The DDD regression model takes the following form:

$$\begin{aligned} Y_{ist} = \alpha_0 + \alpha_1 Post + \alpha_2 High + \alpha_3 Presch + \alpha_4 \{Post * High\} + \alpha_5 \{Presch * High\} + \\ \alpha_6 \{Post * Presch\} + \gamma Post * High * Presch + \delta_1 \sum \{E_s * \lambda_t\} + \eta_1 X_{ist} + \varepsilon_{ist} \end{aligned} \tag{5}$$

The variable Y_{ist} represents the health-related outcomes of interest for child i (either a preschooler aged 3-6 or a toddler aged 1-2) in province s in year t. *Presch* is a dummy variable indicating whether this child was a preschooler (aged 3-6) in that year, with 1 denoting age group 3-6 and 0 denoting age group 1-2. The key coefficient of interest is γ, which is the DDD estimate of the effect of the universal preschool reform on children's health-related outcomes.

Unlike the DID model, I add another dimension of differencing between preschoolers (aged 3-6) and toddlers (aged 1-2). Considering that toddlers in the current year should not be impacted by the preschool reform but may be exposed to "other" changes that occur at the same time in treated provinces, my identifying assumption is that the difference in trend between high-intensity and low-intensity provinces is the same for preschoolers as it is for toddlers in the absence of the preschool reform. The DDD model does not assume that provinces have similar trends in outcomes in the absence of treatment anymore. Instead, it is assumed that preschoolers and toddlers were equally impacted by other contemporaneous shocks, and the difference in the trends between high-expansion and low-expansion provinces would be the same for preschoolers or toddlers without the preschool reform. This method can potentially account for the unobserved trajectories in health outcomes across provinces and the outcome changes of both preschoolers and toddlers in the treatment provinces.

4.3.2 Results

Panel A in Table 5 presents the estimation results from the DID framework. The first column in Panel A suggests no significant effect on stunting. Stunting is influenced by a variety of factors (e.g., genetics) besides nutrition alone and typically begins in utero (Dewey and Begum [2011]). Nutritional studies indicate that once a child reaches the age of two, it is difficult to rectify any stunting that has occurred previously (e.g., Martorell et al. [1994]; Victora et al. [2010]). As a result, addressing stunting, which is less responsive to short-term interventions and requires interventions to be implemented early enough to prevent stunting, may be more challenging than addressing underweight, which can be influenced by dietary changes and exercise. This may explain why my estimates indicate a negligible and nonsignificant effect on stunting of preschoolers.[21] The second column in Panel A shows strong evidence that the preschool expansion reform had significant effects on reducing underweight. Specifically, a preschool-aged child is 5 percentage points less likely to be categorized as underweight if the child resides in a high-expansion province. The results indicate that attending preschool contributes to alleviating undernutrition among preschool-aged children. In addition, results from Panel A show that the reform led to improved health-seeking behavior by showing that caregivers of preschoolers were more likely to see a doctor when children got sick instead of finding medicines by themselves. Panel B in Table 5 shows the results of the control experiment using toddlers. All estimated coefficients are close to zero and statistically insignificant. This is reassuring because toddlers are too young to be affected by the preschool reform.

Panel C in Table 5 presents my preferred specification's estimates from the DDD regression model. Similar to findings from the DID approach discussed above, my DDD estimates suggest that the preschool reform contributes to alleviating underweight among preschoolers. Also, I find no significant effect on stunting and overweight indicators and individual morbidity. Based on the DDD estimate in column 2 of Panel C in Table 5, preschoolers from provinces with a large expansion in subsidized preschools are 4.7 percentage points (33%) less likely to be underweight than their counterparts from provinces with only a small change in access to preschool education. When children get sick, caregivers residing in high-intensity provinces are 12.3 percentage points (18%) more likely to seek professional help from doctors, under the impact of preschool expansion. Since I conduct regressions for multiple outcomes simultaneously, there is a probability that some of the coefficients become statistically significant by chance. To address this concern, I evaluate the robustness of my estimates to multiple hypothesis testing adjustments. Specifically, I adjust for simultaneous inferences and correct the p-values using Romano-Wolf correction (Romano and Wolf [2005a]; Romano

[21] As discussed, this paper focuses on evaluating the short-term effects of the preschool reform and monitors health-related outcomes measured right after the reform. However, it's less likely for the stunting outcome to exhibit an immediate response to interventions compared with other health-related outcomes, as any potential effects on height and stunting are anticipated to manifest slowly. Therefore, one alternative method to estimate the effects on stunting is to exploit the variation across birth cohorts and provinces in policy intensity, a method similar to the one employed when studying any potential effects on preschool attendance. The corresponding results are presented in Table A3 and again show insignificant effects on stunting.

and Wolf [2005b]; Romano and Wolf [2016]), which protects against potential Type I errors and controls the familywise error rate (FWER), the probability of even one false rejection. The Romano-Wolf approach is considerably more powerful than earlier multiple hypothesis testing methods, such as the Bonferroni correction that is commonly used when the individual tests are independent of each other (Bonferroni [1936]; Holm [1979]; Clarke et al. [2020]).[22] This is because the Romano-Wolf approach uses resampling and step-down procedures to take into account the underlying dependence structure of the test statistics with correlated outcomes, which is more appropriate in my study. The Romano-Wolf p-values are calculated using the rwolf package in Stata (Clarke et al. [2020]) and reported in Panel C in Table 5. Despite slight changes to significance levels comparing my reported p-values in parentheses and the Romano-Wolf p-values, the negative estimated effects on being underweight and the positive estimated effects on seeking behavior remain significant at the 10% and 5% level, respectively, with the Romano-Wolf correction. Those results are also robust to several specification checks, as displayed in Table 6.

As noted earlier, the validity of my DDD model relies on several assumptions. Although it is not feasible to directly test these assumptions, I conduct parallel analyses to alleviate potential concerns and support the soundness of my DDD model design. The triple-difference design contains three dimensions of differencing: high-intensity versus low-intensity provinces, pre-reform versus post-reform, and treated group (preschooler) versus untreated group (toddler). As a control experiment, I change the third differencing to toddlers (aged 1-2) versus infants (aged 0) and children aged 11-14 versus children aged 9-10, respectively. Those cohort comparisons are adjacent to those used in the main specifications. Theoretically, the outcomes of infants, toddlers, and children aged 9-14 should not be influenced by the universal preschool reform, so the absence of effects would support the plausibility of my identification. Panel B and Panel C in Table 7 present the triple-difference results in two control experiments individually. Specifically, Panel B in Table 7 explores outcomes of young children (aged 0-2) between 2010 and 2016, while Panel C in Table 7 explores outcomes of children aged 9-14 between 2010 and 2014, considering those children are never exposed to the

[22] An increasing number of recent papers have implemented the Romano-Wolf approach in their studies, such as Gertler et al. [2014], Liu et al. [2015], and Chowdhury et al. [2022].

preschool reform. As expected, each of these coefficients is small and not significantly different from zero across all specifications. Overall, I implement two types of falsification checks to support my triple-difference methodology, and convincing evidence is provided.

To ensure the robustness of my findings, I introduce additional controls as a robustness check. As shown in columns 1 and 6 of Table A4, I incorporate year fixed effects which allows for more flexibility in capturing time trends, instead of a "Post" dummy. In column 2, I further control for the interactions between year dummies and women's baseline employment rate in each province to allow for different underlying trends in women's employment, depending on the pre-reform employment environment characteristics of the province. To capture the provincial-level dynamic factors that may affect children's health, I include provincial characteristics, including provincial GDP per capita and the population each year. To further control for variations in observed or unobserved village-level characteristics, I incorporate village fixed effects via 631 dummy variables reflecting the 631 villages/communities in my sample, which enables better control of time-invariant characteristics specific to each village/community that may be correlated with both the preschool expansion and children's health. With better control of spatial heterogeneity, as shown in columns 5 and 10, my estimates remain consistent. Overall, my results are robust after running the battery of specification checks.

In addition, I explore whether my estimates are sensitive to two sample restrictions. Specifically, I limit my sample to children without siblings or inter-province migration. The effects on child health and their caregiver's medical-seeking behavior are direct for preschoolers. However, within-family spillovers may indirectly affect the siblings of preschoolers via any reallocation of resources within a family in response to preschool reforms (e.g., see Ravindran [2021]). For example, preschool may reshape caregivers' health knowledge and habits, which impress upon them the importance of professional consultation when their children (e.g., preschoolers and their siblings) get ill. To rule out any confounding effects from siblings residing together, I explore estimates limiting my sample to children without siblings who would not experience within-family spillovers. The results remain similar, as shown in Table A5.[23] Besides the sample restriction of children without siblings, I also explore my estimates limiting my sample to children without inter-province migration. I show in columns 1 and 3 of Table A5 that my estimates hold steady when I exclude children who experienced interprovincial migration (1% in my sample). Meanwhile, I test whether migration responds to the preschool reform using both the DID and DDD approaches and show results in Table A6.[24] Null effects on migration relieve the concerns of endogenous migration.

To assess the sensitivity of my findings to the choice of intensity measure, I conduct additional analyses using a continuous intensity measure. As discussed before, a binary indicator of policy intensity is employed to allow for clear intuitions and straightforward explanations of the estimated effects. Provinces are classified into high-intensity and low-intensity provinces depending on whether the provincial policy intensity is larger than the national median. To ensure my findings are robust to the choice of intensity measure, I then use a continuous intensity measure to estimate the effects on health-related outcomes and present results in Tables A7 and A8. Reassuringly, both the DID and DDD estimates reveal significant effects of reduced underweight and improved seeking behavior, as shown in Table A7. And as expected, all coefficients are small and not significantly different from zero in two control experiments, as shown in Panels B and C in Table A8.

4.3.3 Heterogeneity analysis

The effect of the universal preschool reform on children's health-related outcomes may be heterogeneous to individuals' characteristics. Given the common issue of gender inequality worldwide, it is important to investigate whether the effect shows some heterogeneity by children's gender. In addition, a large gap between China's rural and urban areas has existed for a long time in various dimensions, such as income, education, and health.[25] Thus, it is of great importance to examine the

[23]Because of China's one-child policy, which was repealed in 2016, many families had only one child before then to avoid penalties and fines. When I limit my sample to include children without siblings, around 11% of the observations are dropped.

[24]I define migration as whether the current province of residence differs from the previous survey year (1 = interprovincial migration, 0 = no migration).

[25]The underweight ratio among children in rural areas is 18.6%, which is 10 percentage points larger than that in urban areas based on my sample in the CFPS 2010 survey.

differential effects by children's urban-rural residential status.

Heterogeneous effects on the underweight indicator and caregiver's health-seeking behavior are displayed in Table 8. Columns 1 and 2 in Panel A of Table 8 indicates that alleviation of underweight is solely driven by children living in rural areas. Rural families are more economically disadvantaged and are more likely to suffer from childhood malnutrition and food insecurity (Chen et al. [2007]; Hannum et al. [2014]). This implies that nutritious meals, outdoor play, and healthy eating habits in preschool may reduce undernutrition more effectively in rural areas, consistent with the findings in this paper. However, despite the fact that the magnitude is substantially larger among rural children, I cannot reject the equality of the coefficients in columns 1 and 2. Columns 3 and 4 in Panel A of Table 8 show that the preschool expansion leads to a greater impact on urban areas' health-seeking behavior (15.5 percentage points) than it does on rural areas (6.8 percentage points). This may result from a low baseline level of health-seeking behavior, a larger positive effect on mother's entrepreneurship, and better access to healthcare services in urban areas.[26] However, these estimates are imprecise, and the heterogeneity across rural and urban areas is insignificant. Panel B of Table 8 presents the heterogeneous effects by gender. Results in columns 1 and 2 of Panel B provide suggestive evidence that preschool expansion has a greater effect for girls than boys on preventing underweight. One possible explanation can be the root of preferences for sons in China (Das Gupta et al. [2003]); some girls may be poorly taken care of in the absence of preschool attendance compared with boys, especially in rural areas with greater scope for improvement. This differential effect by gender suggests that preschool attendance, which provides meals and essential health services at school, contributes to narrowing the gender gap in preventing nutritional deprivation of children. However, this finding should be interpreted with caution, given an insignificant difference across gender. In contrast, results in columns 3 and 4 of Panel B suggest a larger influence on boys regarding caregivers' medical-seeking behavior (p-value of equality = 0.083). Considering that China has the cultural stigma of having female children as well as the historical preference for sons, this might imply such gender inequality in child healthcare.

[26] Wang and Lin [2019] find suggestive evidence that China's subsidized child care reform has a larger effect on entrepreneurial activities of mothers in urban areas.

4.3.4 Mechanisms

In this section, I first attempt to investigate numerous ways by which the preschool reform affects the prevalence of underweight among preschoolers. Because of the nutritious meals offered at school and healthy eating habits established on campus, improved access to preschool education can directly enhance children's food consumption and reduce the prevalence of underweight. Meanwhile, preschool may bring an indirect effect by increasing parental labor force participation and disposable income. Preschool attendance has been well documented to be associated with higher rates of maternal employment in many studies (e.g., Baker et al. [2008]; Bauernschuster and Schlotter [2015]). Most related to my work, Lin and Wang [2019] and Wang and Lin [2019] have shown that this preschool reform in China has promoted mothers' entrepreneurial activities and grandmothers' labor force participation. If preschool attendance translates into more disposable income as a result of improved employment, this enables higher spending on improving children's food security and nutrition at home. To test this spending channel, I examine the impact on food expenditure at the household level. As shown in Table 9, the estimated effects on three measures of food expenditure, including food expenditure level in column 1, food expenditure as a share of total family expenditure in column 2, and food expenditure per capita (i.e., per family member) in column 3 are all negative and insignificant. These results demonstrate that the mechanism of increased spending on food and nutrition at home may not be the main factor behind the decline in underweight. Although this finding should be interpreted with caution since household food expenditure may not precisely reflect the expenditure specifically for children, it suggests that meals provided, healthy eating habits, and better caring in preschool might be the major mechanisms driving the decline in underweight.

The universal child care reform can improve caregivers' medical-seeking behavior through several potential mechanisms, such as more severe illness in school, the interactions between school health personnel and children's families, and the increased parental employment and purchasing power. First, caregivers may be more inclined to see a doctor if the likelihood of sickness among preschool-aged children increases following the preschool reform. However, I do not see such evidence, as discussed in Section 4.3.2. Second, the observed increase in doctor visits when dealing with a child's sickness may capture the severity of illness. Although I do not find any significant impact on the probability of sickness, there is still a chance that children experience more severe illness while at school, causing caregivers to seek medical help. The estimated null effects on children's hospitalization prevalence reported in column 4 of Table 9 suggest that this may not be the main channel. Overall, the resultant findings could serve as suggestive evidence that the interactions between school health personnel and children's families as well as the increased parental employment and purchasing power may be major mechanisms behind the improved health-seeking behavior.

5 Conclusion

I set out to answer the following question using a quasi-experiment of China's first universal child care reform: Can the expansion of child care facilities improve health-related outcomes for preschoolers? This paper first confirms the positive effects on attendance rates that result from providing access to affordable preschool education. On average, the reform has led to an increase of approximately 300 seats per 1,000 preschool-aged children and an increase of 25.6 percentage points in the preschool attendance rate from 2010 to 2016. Further investigation shows a larger effect on children residing in rural areas than those in urban areas. This trend suggests a potential narrowing of the current gap in access to education between rural and urban areas.

Then I examine the impact on preschoolers' health-related outcomes and find that the reform has the potential to address undernutrition. My estimates suggest that the reform is effective in reducing underweight, which significantly enhances the welfare of low-income families. Specifically, preschoolers from provinces with a large expansion in subsidized preschools are 4.7 percentage points (33%) less likely to be underweight than their counterparts from provinces with only a small change in access to preschool education. The test of heterogeneity shows alleviating underweight mainly occurs in rural areas, which could lead to a reduction in health disparities between China's rural and urban areas. On the other hand, results suggest that the probability of being stunted, overweight, or sick among preschoolers is unaffected.

This paper also detects some evidence of caregivers' health-seeking behavioral response to the reform. When children get sick, caregivers residing in high-expansion provinces are, on average, 12.3 percentage points (18%) more likely to seek professional help from doctors instead of finding treatments themselves. Combining the reform's effects on preschool attendance and health-related outcomes demonstrates that attending preschool at subsidized prices with school meals provided is a powerful tool to reduce children's underweight that may come from food deficiencies and imbalances of nutrients. These results are robust to various changes in model and sample specifications.

In closing, future research should prioritize identifying the direct welfare effects of such a reform, considering the underlying financial costs and welfare benefits. Given that more than 10 years have passed since the initiation of China's universal preschool reform, it is imperative for future research to study the long-term effects of it on child development.

References

Anderson, Laurie M, Carolynne Shinn, Mindy T Fullilove et al. (2003) "The effectiveness of early childhood development programs: A systematic review," *American Journal of Preventive Medicine*, 24 (3), 32–46.

Ansari, Arya, Kierra Pettit, and Elizabeth Gershoff (2015) "Combating obesity in head start: outdoor play and change in children's BMI," *Journal of developmental and behavioral pediatrics: JDBP*, 36 (8), 605.

Attanasio, Orazio, Raquel Bernal, Michele Giannola, and Milagros Nores (2020a) "Child development in the early years: Parental investment and the changing dynamics of different dimensions,"Technical report, National Bureau of Economic Research.

Attanasio, Orazio, Costas Meghir, and Emily Nix (2020b) "Human capital development and parental investment in India," *The Review of Economic Studies*, 87 (6), 2511–2541.

Attanasio, Orazio, Costas Meghir, Emily Nix, and Francesca Salvati (2017) "Human capital growth and poverty: Evidence from Ethiopia and Peru," *Review of Economic Dynamics*, 25, 234–259.

Baker, Michael, Jonathan Gruber, and Kevin Milligan (2008) "Universal child care, maternal labor supply, and family well-being," *Journal of Political Economy*, 116 (4), 709–745.

Ball, Thomas M, Catharine J Holberg, Michael B Aldous, Fernando D Martinez, and Anne L Wright (2002) "Influence of attendance at day care on the common cold from birth through 13 years of age," *Archives of pediatrics & adolescent medicine*, 156 (2), 121–126.

Bauernschuster, Stefan and Martin Schlotter (2015) "Public child care and mothers' labor supply—Evidence from two quasi-experiments," *Journal of Public Economics*, 123, 1–16.

van den Berg, Gerard J and Bettina M Siflinger (2022) "The effects of a daycare reform on health in childhood–Evidence from Sweden," *Journal of Health Economics*, 81, 102577.

Bi, Peng, Shilu Tong, and Kevin A Parton (2000) "Family self-medication and antibiotics abuse for children and juveniles in a Chinese city," *Social Science & Medicine*, 50 (10), 1445–1450.

Bonferroni, Carlo (1936) "Teoria statistica delle classi e calcolo delle probabilita," *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, 8, 3–62.

Bosque-Mercader, Laia (2022) "The effect of a universal preschool programme on long-term health outcomes: Evidence from spain,"Technical report, HEDG, c/o Department of Economics, University of York.

Burchinal, Margaret R, Joanne E Roberts, Rhodus Riggins, Jr, Susan A Zeisel, Eloise Neebe, and Donna Bryant (2000) "Relating quality of center-based child care to early cognitive and language development longitudinally," *Child development*, 71 (2), 339–357.

Burger, Kaspar (2010) "How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds," *Early childhood research quarterly*, 25 (2), 140–165.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2008) "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 90 (3), 414–427.

Cameron, A Colin and Douglas L Miller (2015) "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, 50 (2), 317–372.

Canay, Ivan A, Andres Santos, and Azeem M Shaikh (2021) "The wild bootstrap with a "small" number of "large" clusters," *The Review of Economics and Statistics*, 103 (2), 346–363.

Cascio, Elizabeth U (2023) "Does universal preschool hit the target? Program access and preschool impacts," *Journal of Human Resources*, 58 (1), 1–42.

Chan, Cheris Shun-ching (2018) "Mistrust of physicians in China: society, institution, and interaction as root causes," *Developing World Bioethics*, 18 (1), 16–25.

Chen, He and Tianguang Meng (2015) "Bonding, bridging, and linking social capital and self-rated health among Chinese adults: Use of the anchoring vignettes technique," *PloS one*, 10 (11), e0142300.

Chen, Qihui (2015) "Ready for school? Impacts of delayed primary school enrollment on children's educational outcomes in rural China," *International Journal of Educational Development*, 45, 112–128.

——— (2017) "Impacts of Late School Entry on Children's Cognitive Development in Rural Northwestern China—Does Preprimary Education Matter?" *Asia & the Pacific Policy Studies*, 4 (3), 586–601.

Chen, Zhuo, David B Eastwood, and Steven T Yen (2007) "A decade's story of childhood malnutrition inequality in China: Where you live does matter," *China Economic Review*, 18 (2), 139–154.

Choi, Jeoungbin, Moran Ki, Ho Jang Kwon et al. (2019) "Health indicators related to disease, death, and reproduction," *Journal of Preventive Medicine and Public Health*, 52 (1), 14.

Chowdhury, Shyamal, Matthias Sutter, and Klaus F Zimmermann (2022) "Economic preferences across generations and family clusters: A large-scale experiment in a developing country," *Journal of Political Economy*, 130 (9), 2361–2410.

Clarke, Damian, Joseph P Romano, and Michael Wolf (2020) "The Romano–Wolf multiple-hypothesis correction in Stata," *The Stata Journal*, 20 (4), 812–843.

Das Gupta, Monica, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok (2003) "Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea," *The Journal of Development Studies*, 40 (2), 153–187.

Démurger, Sylvie and Hui Xu (2015) "Left-behind children and return migration in China," *IZA Journal of Migration*, 4 (1), 1–21.

Dewey, Kathryn G and Khadija Begum (2011) "Long-term consequences of stunting in early life," *Maternal & Child Nutrition*, 7, 5–18.

D'Onise, Katina, John W Lynch, Michael G Sawyer, and Robyn A McDermott (2010) "Can preschool improve child health outcomes? A systematic review," *Social science & medicine*, 70 (9), 1423–1440.

Du, Fenglian and Xiao-yuan Dong (2013) "Women's employment and child care choices in urban China during the economic transition," *Economic Development and Cultural Change*, 62 (1), 131–155.

Duflo, Esther (2001) "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *American economic review*, 91 (4), 795–813.

Fang, Yu (2014) "China should curb non-prescription use of antibiotics in the community," *BMJ*, 348.

Ganimian, Alejandro J, Karthik Muralidharan, and Christopher R Walters (2021) "Augmenting State Capacity for Child Development: Experimental Evidence from India," Working Paper 28780, National Bureau of Economic Research, 10.3386/w28780.

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.

Gong, Xin, Di Xu, and Wen-Jui Han (2016) "The effects of preschool attendance on adolescent outcomes in rural China," *Early Childhood Research Quarterly*, 37, 140–152.

Hannum, Emily, Jihong Liu, and Edward A Frongillo (2014) "Poverty, food insecurity and nutritional deprivation in rural China: Implications for children's literacy achievement," *International Journal of Educational Development*, 34, 90–97.

He, Mingguang, Wenyong Huang, Yingfeng Zheng, Li Huang, and Leon B Ellwein (2007) "Refractive error and visual impairment in school children in rural southern China," *Ophthalmology*, 114 (2), 374–382.

Holm, Sture (1979) "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, 65–70.

Jain, Monica (2015) "India's struggle against malnutrition—is the ICDS program the answer?" *World Development*, 67, 72–89.

Ji, Cheng Ye and Tsung O Cheng (2008) "Prevalence and geographic distribution of childhood obesity in China in 2005."

Kandpal, Eeshani (2011) "Beyond average treatment effects: distribution of child nutrition outcomes and program placement in India's ICDS," *World Development*, 39 (8), 1410–1421.

Lin, Mengyun and Qing Wang (2019) "Center-based childcare expansion and grandparents' employment and well-being," *Social Science & Medicine*, 240, 112547.

Liu, Lily Y, Andrew J Patton, and Kevin Sheppard (2015) "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, 187 (1), 293–311.

Liu, Yan, JH Zhao, and Li Zhang (2013) "Value added analysis of teaching effectiveness of one-year preschool education," *Journal of Educational Studies*, 9 (3), 56–66.

Lumeng, Julie C, Niko Kaciroti, Julie Sturza, Allison M Krusky, Alison L Miller, Karen E Peterson, Robert Lipton, and Thomas M Reischl (2015) "Changes in body mass index associated with head start participation," *Pediatrics*, 135 (2), e449–e456.

Luo, Renfu, Linxiu Zhang, Chengfang Liu, Qiran Zhao, Yaojiang Shi, Scott Rozelle, and Brian Sharbono (2012) "Behind before they begin: The challenge of early childhood education in rural China," *Australasian Journal of Early Childhood*, 37 (1), 55–64.

MacKinnon, James G and Matthew D Webb (2017) "Wild bootstrap inference for wildly different cluster sizes," *Journal of Applied Econometrics*, 32 (2), 233–254.

Martorell, Reynaldo, L Kettel Khan, and Dirk G Schroeder (1994) "Reversibility of stunting: epidemiological findings in children from developing countries.," *European Journal of Clinical Nutrition*, 48, S45–57.

Murray, Christopher JL and Lincoln C Chen (1992) "Understanding morbidity change," *The Population and Development Review*, 481–503.

Nguyen, Cuong Viet (2022) "The effect of preschool attendance on Children's health: Evidence from a lower middle-income country," *Health Economics*, 31 (8), 1558–1589.

Peisner-Feinberg, Ellen S, Margaret R Burchinal, Richard M Clifford, Mary L Culkin, Carollee Howes, Sharon Lynn Kagan, and Noreen Yazejian (2001) "The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade," *Child development*, 72 (5), 1534–1553.

Van de Poel, Ellen, Owen O'Donnell, and Eddy Van Doorslaer (2009) "Urbanization and the spread of diseases of affluence in China," *Economics & Human Biology*, 7 (2), 200–216.

Popkin, Barry M, Wolney Conde, Ningqi Hou, and Carlos Monteiro (2006) "Is there a lag globally in overweight trends for children compared with adults?" *Obesity*, 14 (10), 1846–1853.

Pulgaron, Elizabeth R and Alan M Delamater (2014) "Obesity and type 2 diabetes in children: epidemiology and treatment," *Current diabetes reports*, 14 (8), 1–12.

Quan-Cheng, Kan, Wen Jian-Guo, Liu Xiang-Hua, and Li Zhen-Zhen (2016) "Inappropriate use of antibiotics in children in China," *The Lancet*, 387 (10025), 1273–1274.

Radyowijati, Aryanti and Hilbrand Haak (2003) "Improving antibiotic use in low-income countries: an overview of evidence on determinants," *Social Science & Medicine*, 57 (4), 733–744.

Rao, Nirmala, Ben Richards, Jin Sun, Ann Weber, and Alanna Sincovich (2019) "Early childhood education and child development in four countries in East Asia and the Pacific," *Early Childhood Research Quarterly*, 47, 169–181.

Rao, Nirmala, Jin Sun, Jing Zhou, and Li Zhang (2012) "Early achievement in rural China: The role of preschool experience," *Early Childhood Research Quarterly*, 27 (1), 66–76.

Ravindran, Saravana (2021) "Parental investments and early childhood development: Short and long run evidence from India," *Available at SSRN 3928352*.

Romano, Joseph P and Michael Wolf (2005a) "Exact and approximate stepdown methods for multiple hypothesis testing," *Journal of the American Statistical Association*, 100 (469), 94–108.

——— (2005b) "Stepwise multiple testing as formalized data snooping," *Econometrica*, 73 (4), 1237–1282.

——— (2016) "Efficient computation of adjusted p-values for resampling-based stepdown multiple testing," *Statistics & Probability Letters*, 113, 38–40.

Rosenzweig, Mark R and Kenneth I Wolpin (1988) "Migration selectivity and the effects of public programs," *Journal of Public Economics*, 37 (3), 265–289.

Shen, Xiuhua, Xiang Gao, Wenjing Tang, Xuanxia Mao, Jingyan Huang, and Wei Cai (2015) "Food insecurity and malnutrition in Chinese elementary school students," *British Journal of Nutrition*, 114 (6), 952–958.

Victora, Cesar Gomes, Mercedes De Onis, Pedro Curi Hallal, Monika Blössner, and Roger Shrimpton (2010) "Worldwide timing of growth faltering: revisiting implications for interventions," *Pediatrics*, 125 (3), e473–e480.

Wang, Qing and Mengyun Lin (2019) "Work-family policy and female entrepreneurship: Evidence from China's subsidized child care program," *China Economic Review*, 54 (C), 256–270.

Wang, Shuo, Yan Hui Dong, Zheng He Wang, ZY Zou, and Jun Ma (2017) "Trends in overweight and obesity among Chinese children of 7-18 years old during 1985-2014," *Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine]*, 51 (4), 300–305.

Wu, Kin Bing, Mary Eming Young, and Jianhua Cai (2012) *Early child development in China: Breaking the cycle of poverty and improving future competitiveness*: World Bank Publications.

Xiaodong, Zeng (2008) "A design of an appropriate early childhood education funding system in China," *Chinese Education & Society*, 41 (2), 8–19.

Zhu, Yao, Xuewen Tang, Rui Yan, Zhujun Shao, Yang Zhou, Xuan Deng, Shuying Luo, and Hanqing He (2021) "Non-prescription antibiotic use for cough among Chinese children under 5 years of age: a community-based cross-sectional study," *BMJ Open*, 11 (12), e051372.

Zong, Xin-Nan, Hui Li, Ya-Qin Zhang, and Hua-Hong Wu (2019) "Child nutrition to new stage in China: evidence from a series of national surveys, 1985–2015," *BMC public health*, 19 (1), 1–13.

Table 1: Summary Statistics

	2010 Before the reform			2012-2016 After the reform		
	(1) All	(2) High Intensity	(3) Low Intensity	(4) All	(5) High Intensity	(6) Low Intensity
Individual and Family Characteristics						
Age	8.19	8.03	8.43	8.17	8.12	8.24
	(4.40)	(4.39)	(4.40)	(4.26)	(4.28)	(4.23)
Male	0.55	0.55	0.55	0.52	0.51	0.53
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Urban	0.43	0.39	0.49	0.49	0.45	0.55
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)
Family size	4.87	4.91	4.81	5.02	5.03	5.00
	(1.58)	(1.57)	(1.59)	(1.66)	(1.65)	(1.67)
Household income per capita	5,640.59	5,226.67	6,259.40	8,678.12	8,034.53	9,573.81
	(4,059.25)	(3,911.95)	(4,195.13)	(6,260.66)	(5,869.13)	(6,665.73)
School Expansion						
Preschool policy intensity	0.43	0.55	0.26	0.42	0.54	0.25
	(0.18)	(0.11)	(0.13)	(0.18)	(0.10)	(0.13)
Outcome Variables						
Ever attended preschool (Age$\geqslant$ 3)	0.72	0.67	0.78	0.84	0.82	0.87
	(0.45)	(0.47)	(0.41)	(0.37)	(0.38)	(0.33)
Stunting	0.23	0.24	0.21	0.17	0.18	0.16
	(0.42)	(0.43)	(0.41)	(0.38)	(0.38)	(0.37)
Overweight	0.30	0.31	0.30	0.30	0.30	0.29
	(0.46)	(0.46)	(0.46)	(0.46)	(0.46)	(0.46)
Underweight	0.17	0.18	0.16	0.11	0.11	0.12
	(0.38)	(0.38)	(0.36)	(0.32)	(0.31)	(0.32)
Kids ever sick last month	0.29	0.29	0.28	0.27	0.26	0.28
	(0.45)	(0.45)	(0.45)	(0.44)	(0.44)	(0.45)
See a doctor when kids get sick	0.68	0.66	0.70	0.60	0.59	0.61
	(0.47)	(0.47)	(0.46)	(0.49)	(0.49)	(0.49)
Observations	7,624	4,455	3,169	18,433	11,174	7,259

Notes: Sample includes children aged 0-15. Columns (1)–(3) report summary statistics for the year 2010 (before the reform), (4)–(6) for 2012, 2014, and 2016 (after the reform). Statistics are weighted with CFPS individual-level national weights. Standard deviations are reported in parentheses.

Table 2: Effects of the reform on preschool attendance

	Attendance	Attendance	Attendance
(Aged below 6 in 2010)*High Intensity	0.178**	0.139**	0.143**
	(0.071)	(0.054)	(0.053)
	[0.040]**	[0.020]**	[0.013]**
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
Baseline Mean	0.700	0.700	0.700
R-squared	0.2441	0.2552	0.2811
Obs	8,001	8,001	8,001

Notes: This table presents the DID estimates of the reform on preschool attendance. It compares the attendance history of children who had no exposure to the reform (aged 9 to 13 in 2010) to those of individuals who were exposed (either below age 6 in 2010 or born between 2010 to 2012) in high-intensity provinces and low-intensity provinces. The dependent variable is a dummy variable reflecting whether a child is attending/has attended any preschools. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 3: Effects on preschool attendance: experiment of interest and control experiment

Panel A: Experiment of Interest: Individuals aged below 6 or aged 9-13 in 2010			
(Youngest cohort: Individuals aged -2 to 6 in 2010)			
	Attendance	Attendance	Attendance
(Aged below 6 in 2010)*Policy intensity in that province	0.667***	0.644***	0.643***
	(0.074)	(0.103)	(0.101)
	[0.000]***	[0.000]***	[0.000]***
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
Baseline Mean	0.700	0.700	0.700
R-squared	0.265	0.268	0.294
Obs	8,001	8,001	8,001
Panel B: Control Experiment: Individuals aged 9 to 15 in 2010			
(Youngest cohort: Individuals aged 9 to 13 in 2010)			
	Attendance	Attendance	Attendance
(Aged 9-13 in 2010)*Policy intensity in that province	0.006	0.083	0.060
	(0.044)	(0.055)	(0.052)
	[0.953]	[0.127]	[0.307]
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
Baseline Mean	0.700	0.700	0.700
R-squared	0.255	0.256	0.333
Obs	4,084	4,084	4,084

Notes: This table presents the estimated effects on preschool attendance. Panel A shows the results of the experiment of interest (comparing the cohort aged below 6 in 2010 to the cohort aged 9 to 13 in 2010). Panel B shows the results of the control experiment (comparing the cohort aged 9 to 13 in 2010 to the cohort aged 14 to 15 in 2010). The dependent variable is a dummy variable reflecting whether a child is attending/has attended any preschools. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 4: Heterogeneous effects by urban-rural residence status and gender

	Attendance			
	(1)		(2)	
	Urban	Rural	Boy	Girl
(Aged below 6 in 2010)*Policy intensity in that province	0.383***	0.577***	0.704***	0.592***
	(0.125)	(0.112)	(0.105)	(0.106)
	[0.013]**	[0.000]***	[0.000]***	[0.000]***
Coef. comp. p-value	0.028		0.209	
Province FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Baseline Mean	0.896	0.587	0.704	0.697
R-squared	0.140	0.339	0.301	0.295
Obs	3,085	4,916	4,345	3,656

Notes: This table presents the estimation results on sub-samples by urban-rural residence status and gender. The dependent variable is a dummy variable reflecting whether a child is attending/has attended any preschools. Controls include child age, family size, and gender or urban-rural residence status. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 5: Effects on health-related outcomes: DID and DDD estimates

Panel A: DID Estimates - Experiment of Interest					
Sample: Preschoolers (aged 3-6)					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity	0.014	-0.052**	-0.010	-0.042	0.067*
	(0.027)	(0.022)	(0.025)	(0.033)	(0.037)
	[0.647]	[0.047]**	[0.720]	[0.260]	[0.040]**
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.306	0.143	0.414	0.408	0.701
R-squared	0.048	0.027	0.021	0.019	0.027
Obs	8,248	8,248	8,248	8,248	8,248
Panel B: DID Estimates - Control Experiment					
Sample: Toddlers (aged 1-2)					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity	-0.023	-0.006	-0.046	-0.027	-0.058
	(0.034)	(0.016)	(0.040)	(0.045)	(0.051)
	[0.193]	[0.760]	[0.220]	[0.553]	[0.287]
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Baseline Mean (1-2)	0.391	0.086	0.477	0.517	0.734
R-squared	0.059	0.024	0.031	0.011	0.025
Obs	4,228	4,228	4,228	4,228	4,228
Panel C: Preferred Specification - DDD Estimates					
Sample: Preschoolers (aged 3-6) and Toddlers (aged 1-2)					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity*Preschooler	0.039	-0.047**	0.035	-0.021	0.123***
	(0.042)	(0.023)	(0.047)	(0.053)	(0.042)
	[0.367]	[0.067]*	[0.400]	[0.700]	[0.013]**
Romano-Wolf p-value	0.327	0.069*	0.416	0.673	0.010**
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.306	0.143	0.414	0.408	0.701
R-squared	0.056	0.026	0.024	0.024	0.026
Obs	12,476	12,476	12,476	12,476	12,476

Notes: This table presents the DID and DDD estimates on health-related outcomes. Panel A shows the DID results of the experiment of interest, and the sample consists of preschoolers (aged 3-6) in each year from 2010 to 2016. Panel B shows the DID results of the control experiment, and the sample consists of toddlers (aged 1-2) in each year from 2010 to 2016. Panel C shows the DDD estimates for health-related outcomes, and the sample consists of preschoolers (aged 3-6) and toddlers (aged 1-2) in each year from 2010 to 2016. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 6: Robustness checks: Specifications with different controls

Outcome	Underweight			See a doctor		
	(1)	(2)	(3)	(1)	(2)	(3)
Post*High intensity*Preschooler	-0.042*	-0.041*	-0.047**	0.125***	0.131***	0.123***
	(0.023)	(0.023)	(0.023)	(0.043)	(0.042)	(0.042)
	[0.090]*	[0.090]*	[0.067]*	[0.000]***	[0.000]***	[0.013]**
Year Dummy*Pre-reform Enrollment	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
R-squared	0.006	0.008	0.026	0.007	0.015	0.026
Obs	12,476	12,476	12,476	12,476	12,476	12,476

Notes: This table presents the estimated effects from the DDD model under different specifications. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 7: Experiment of interest and control experiment: DDD estimates

Panel A: Experiment of Interest					
Preschooler (aged 3-6) vs Toddlers (aged 1-2)					
Sample: Children aged 1-6, 2010-2016					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity*Preschooler	0.039	-0.047**	0.035	-0.021	0.123***
	(0.042)	(0.023)	(0.047)	(0.053)	(0.042)
	[0.367]	[0.067]*	[0.400]	[0.700]	[0.013]**
R-squared	0.056	0.026	0.024	0.024	0.026
Obs	12,476	12,476	12,476	12,476	12,476
Panel B: Control Experiment					
Toddlers (aged 1-2) vs Infants (aged 0)					
Sample: Children aged 0-2, 2010-2016					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity*Aged 1-2	0.071	0.007	0.029	0.048	0.005
	(0.106)	(0.086)	(0.059)	(0.072)	(0.119)
	[0.560]	[0.900]	[0.613]	[0.540]	[0.933]
R-squared	0.067	0.045	0.041	0.034	0.024
Obs	5,264	5,264	5,264	5,264	5,264
Panel C: Control Experiment					
Children aged 11-14 vs Children aged 9-10					
Sample: Children aged 9-14, 2010-2014					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*High intensity*Age11-14	0.044	0.022	-0.008	-0.027	0.106
	(0.050)	(0.044)	(0.049)	(0.056)	(0.065)
	[0.367]	[0.667]	[0.813]	[0.660]	[0.160]
R-squared	0.077	0.026	0.054	0.006	0.040
Obs	8,471	8,471	8,471	8,471	8,471

Notes: This table compares the DDD estimates for health-related outcomes in the experiment of interest (Panel A) and two control experiments (Panels B and C). Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 8: Heterogeneous effects by urban-rural residence status and gender

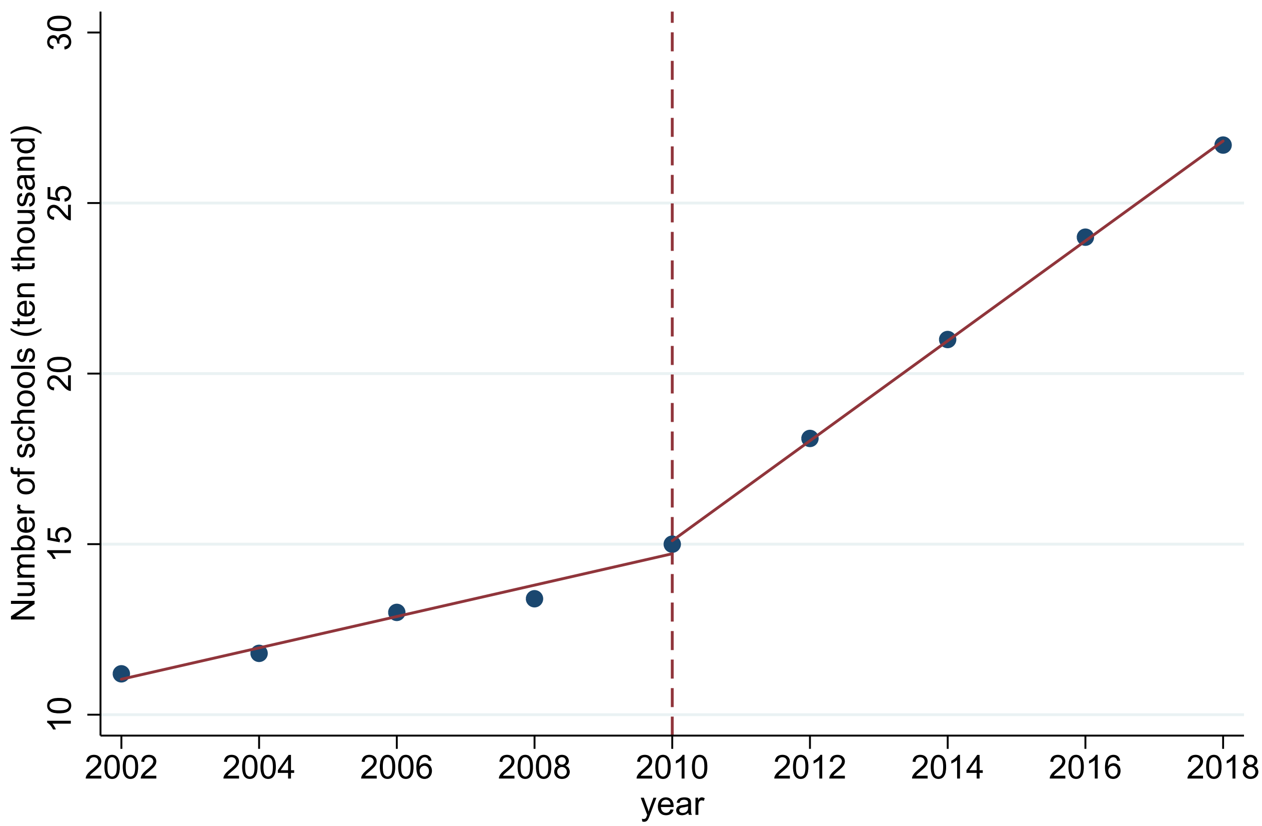
	Underweight		See a doctor	
Panel A: Heterogeneous effects by urban-rural residence status				
	(1)	(2)	(3)	(4)
	Urban	Rural	Urban	Rural
Post*High intensity*Preschooler	0.002	-0.088**	0.155**	0.068
	(0.031)	(0.041)	(0.068)	(0.045)
	[0.933]	[0.070]*	[0.047]**	[0.133]
Coef. comp. p-value	0.117		0.266	
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.084	0.186	0.656	0.733
R-squared	0.010	0.020	0.020	0.016
Obs	4,907	7,569	4,907	7,569
Panel B: Heterogeneous effects by gender				
	Underweight		See a doctor	
	(1)	(2)	(3)	(4)
	Boy	Girl	Boy	Girl
Post*High intensity*Preschooler	-0.010	-0.093*	0.198***	0.017
	(0.034)	(0.052)	(0.067)	(0.056)
	[0.867]	[0.127]	[0.013]**	[0.767]
Coef. comp. p-value	0.182		0.083	
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.142	0.144	0.715	0.683
R-squared	0.026	0.028	0.024	0.029
Obs	6,890	5,586	6,890	5,586

Notes: This table presents the estimation results on sub-samples by child urban-rural residence status in Panel A and by child gender in Panel B, respectively. Controls include child age, gender, and family size in Panel A. Controls include child age, urban-rural residence status, and family size in Panel B. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table 9: Mechanism investigation

	(1)	(2)	(3)	(4)
Outcome	Food expenditure	Food/Total expenditure	Food expenditure per capita	Hospitalized in the past year
Post*High intensity*Preschooler	-874.766	-0.013	-116.368	0.018
	(1422.332)	(0.022)	(320.141)	(0.056)
	[0.553]	[0.673]	[0.673]	[0.733]
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	9308.524	0.360	2020.816	0.107
R-squared	0.143	0.045	0.195	0.027
Obs	12,085	11,591	12,085	12,355

Notes: This table presents the DDD estimates for four additional outcomes, including three measures of food expenditure (food expenditure in CNY, food expenditure as a share of total family expenditure, and food expenditure in CNY per family member) and an indicator of whether the child was ever hospitalized in the past year. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.



Notes: The solid lines are the linear fits during pre- and post-reform periods. Data come from Educational Statistics Yearbook of China

Figure 1: Number of preschools in China

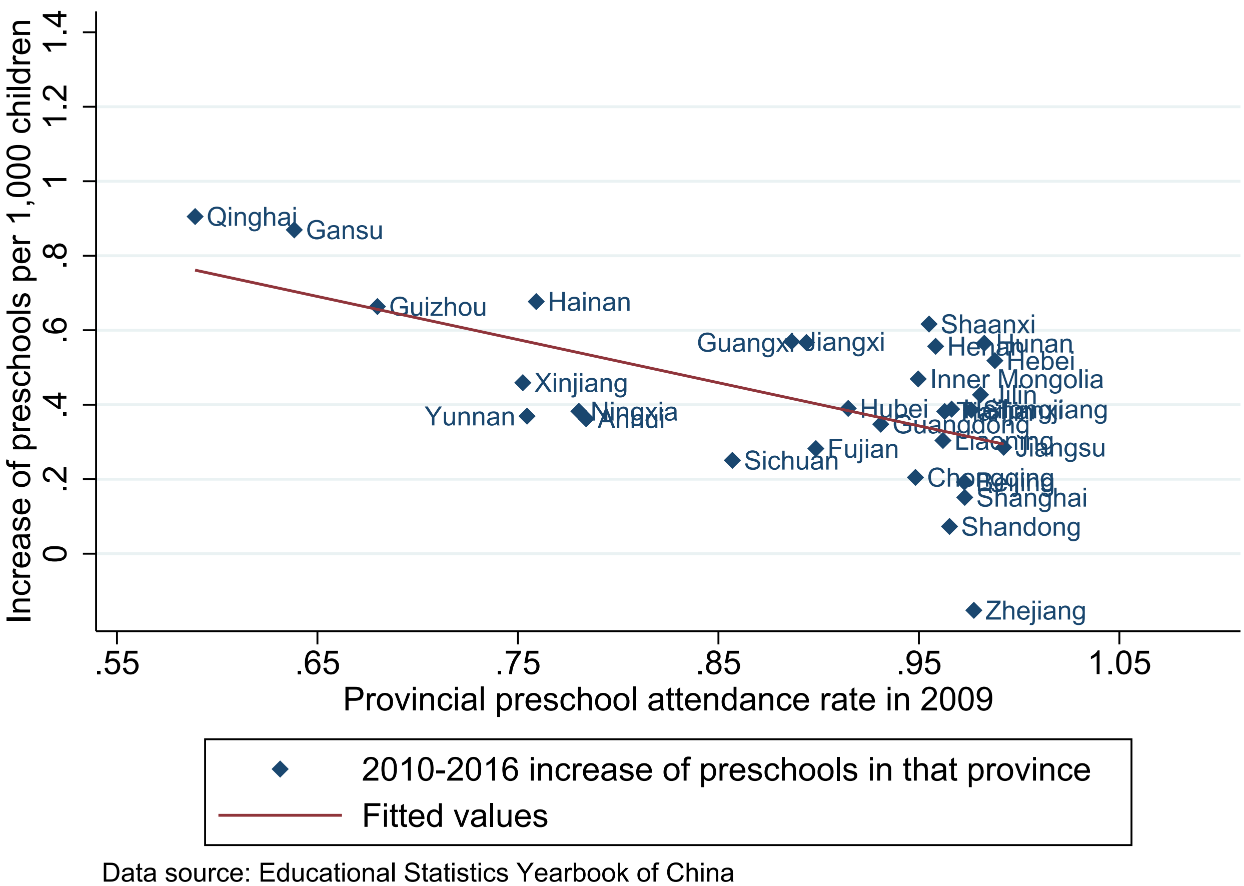
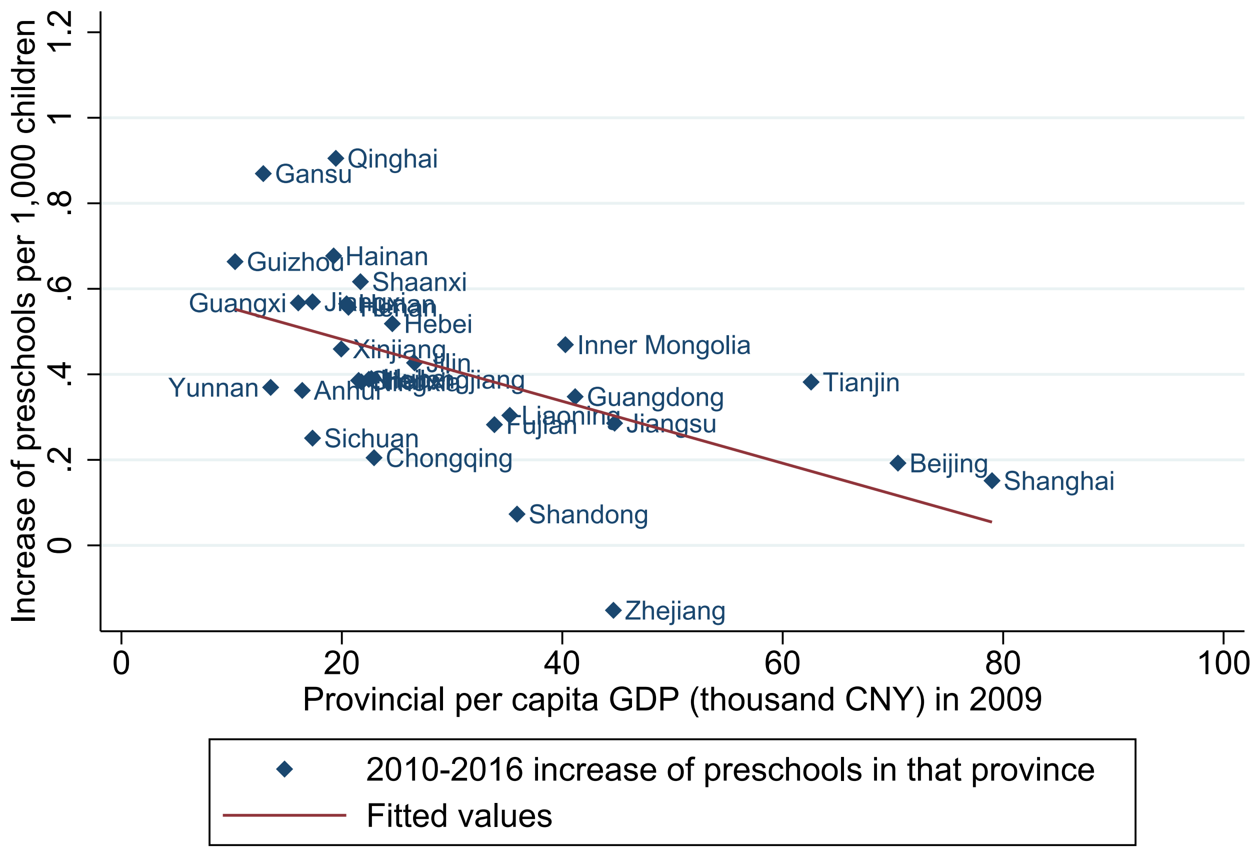
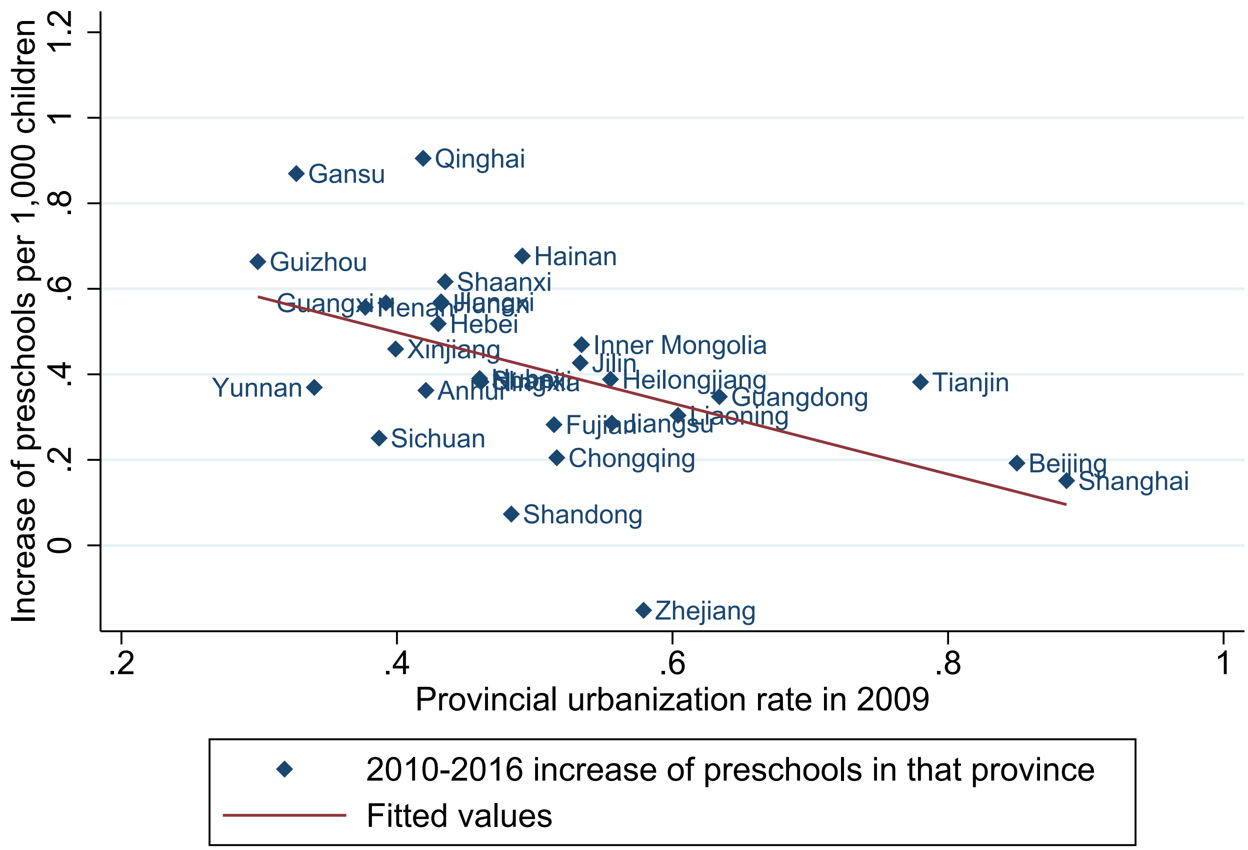


Figure 2: Correlation between pre-reform attendance rate and the increase of preschools (policy intensity)



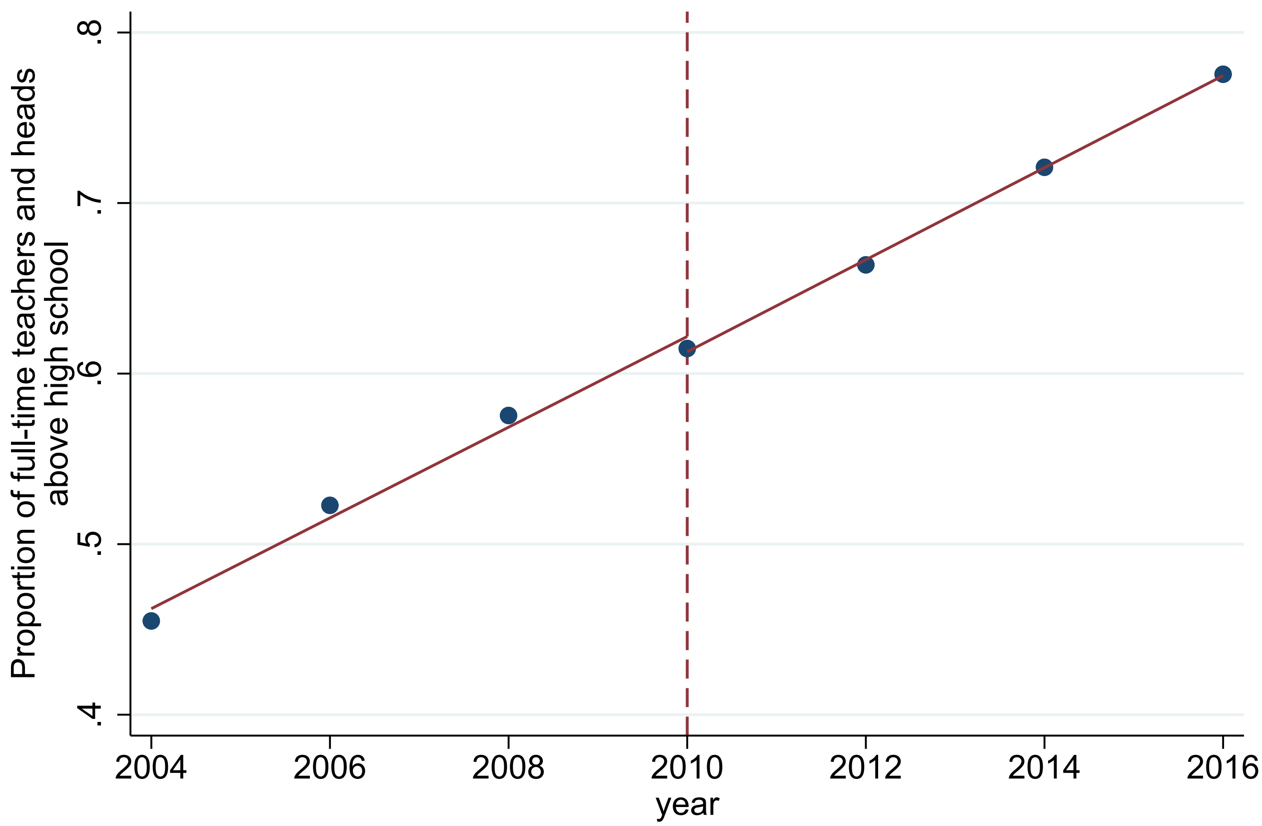
Data source: Educational Statistics Yearbook of China

Figure 3: Correlation between pre-reform per capita GDP and the increase of preschools (policy intensity)



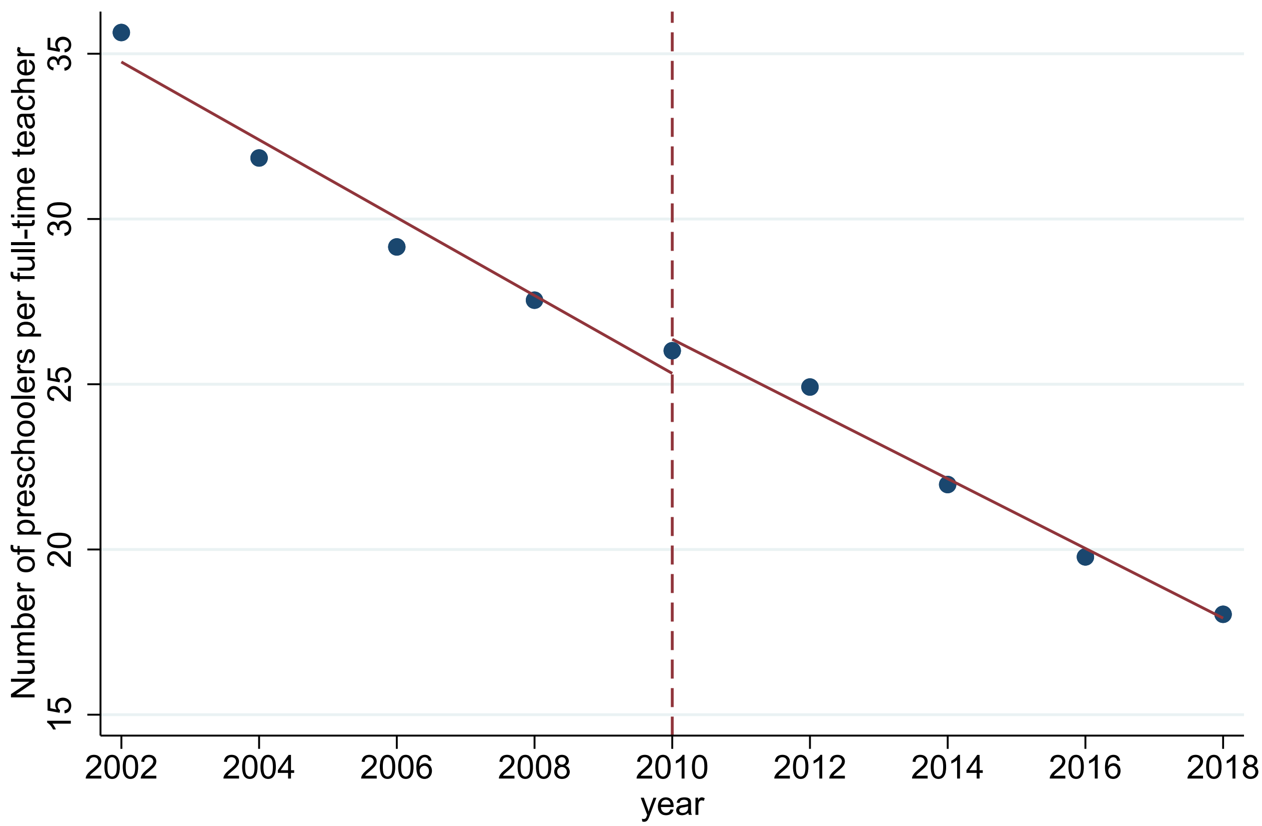
Data source: China Health and Family Planning Statistics Book

Figure 4: Correlation between pre-reform urbanization rate and the increase of preschools (policy intensity)



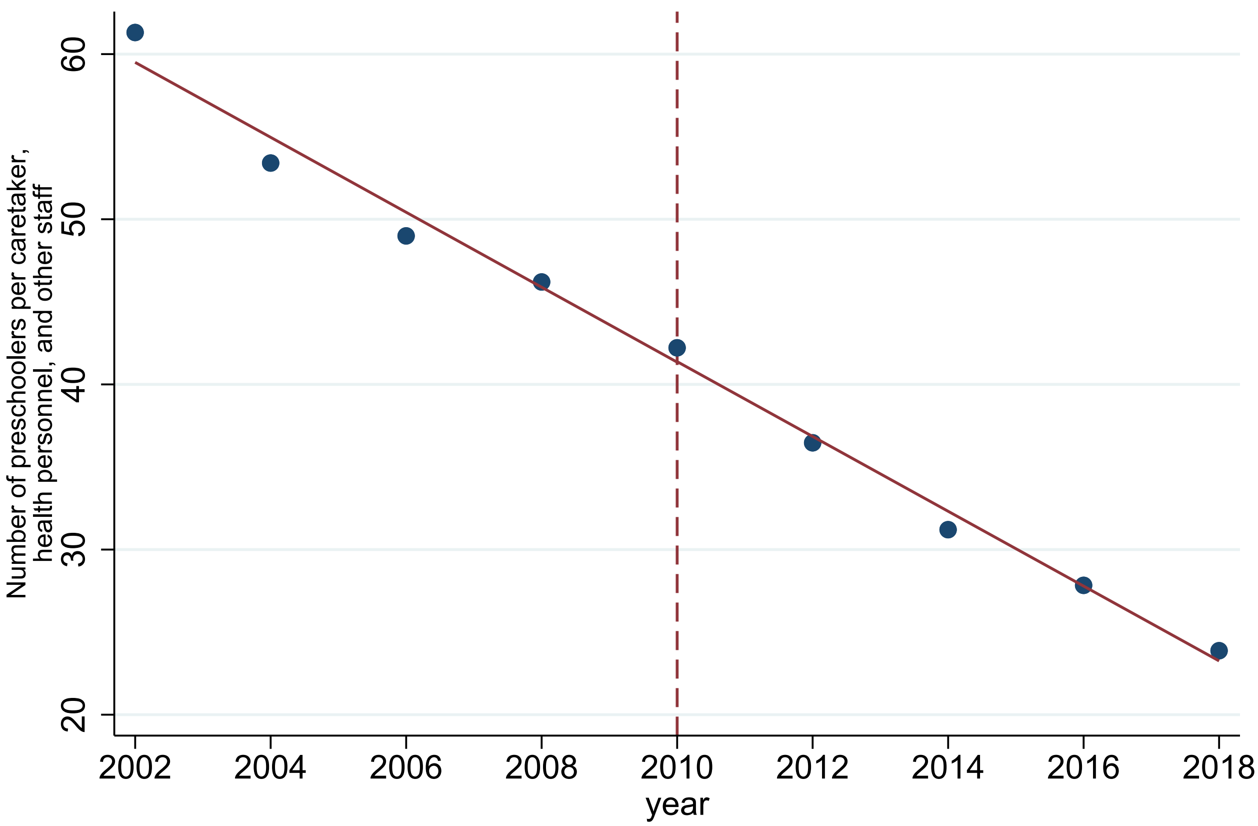
Notes: The solid lines are the linear fits during pre- and post-reform periods. Data come from Educational Statistics Yearbook of China

Figure 5: Proportion of full-time teachers and heads with educational attainment beyond high school



Notes: The solid lines are the linear fits during pre- and post-reform periods. Data come from Educational Statistics Yearbook of China

Figure 6: Number of preschoolers per full-time teacher in preschool



Notes: The solid lines are the linear fits during pre- and post-reform periods. Data come from Educational Statistics Yearbook of China

Figure 7: Number of preschoolers per caretaker, health personnel, and other staff in preschool

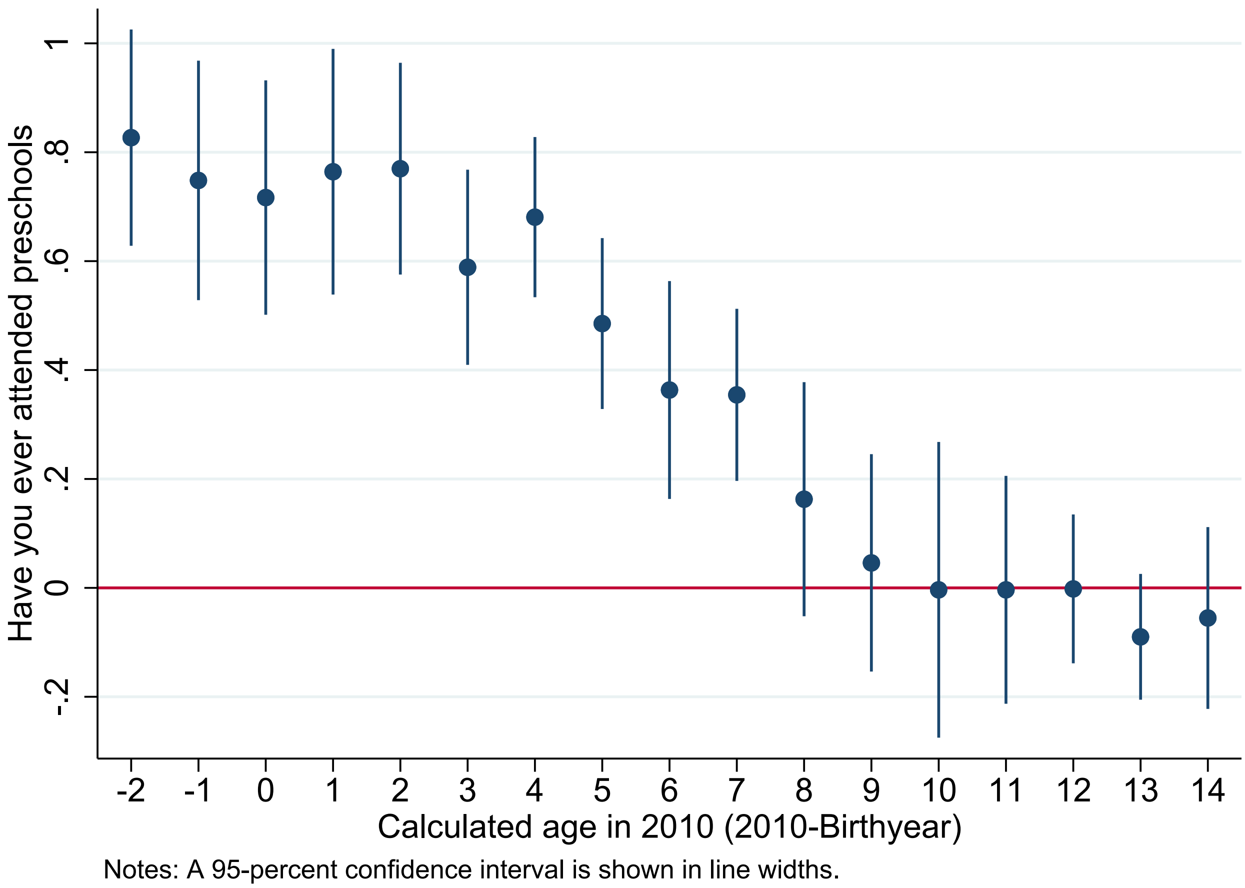


Figure 8: Coefficient of the Interactions Policy Intensity*Cohort Dummy

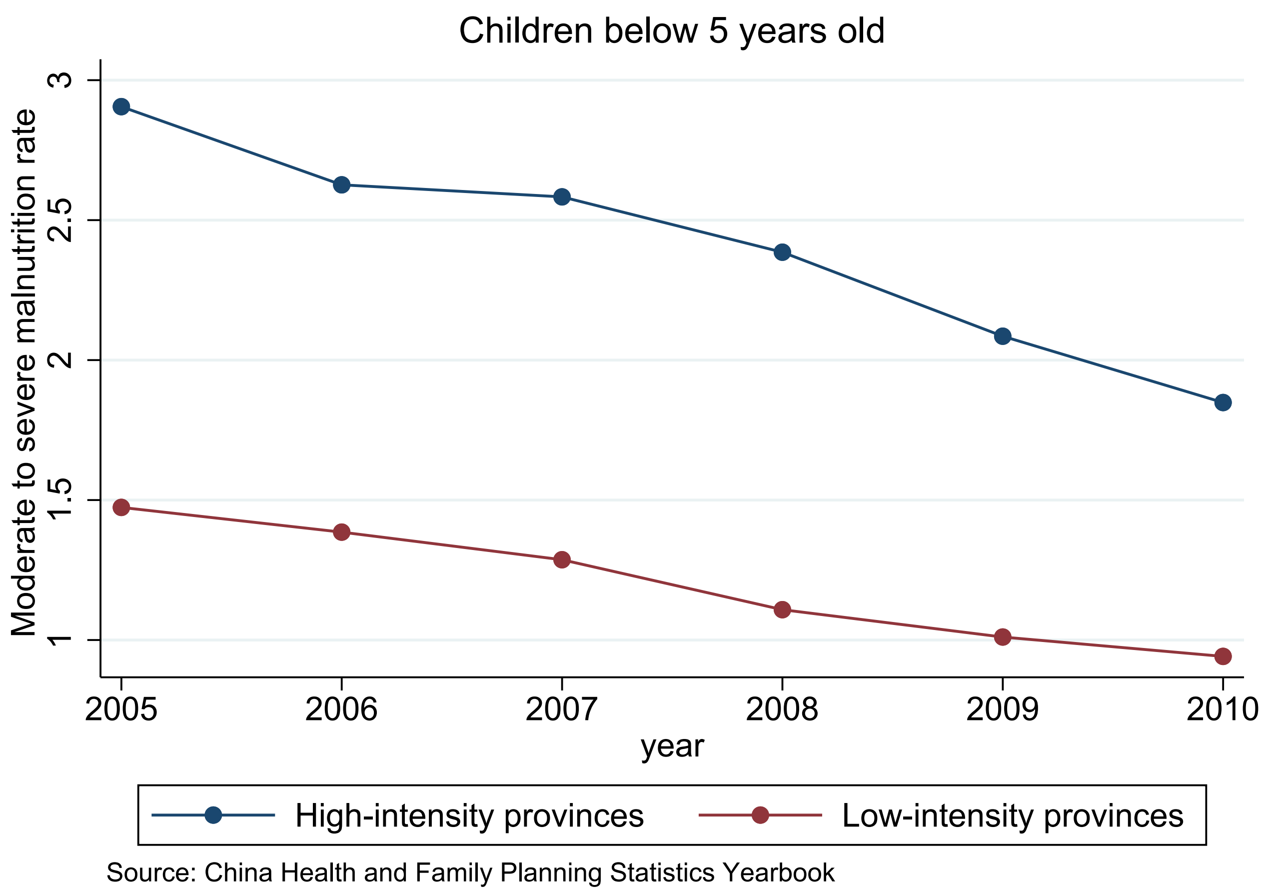


Figure 9: Moderate to severe malnutrition rate

6 Appendix

Table A1: Additional robustness checks: Effects on preschool attendance

	Attendance	Attendance	Attendance	Attendance
(Aged below 6 in 2010)*High Intensity	0.164***	0.164***	0.139**	0.101**
	(0.053)	(0.053)	(0.051)	(0.049)
	[0.000]***	[0.007]***	[0.020]**	[0.067]*
Province FE	Yes	Yes	Yes	No
Cohort FE	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Village FE	No	No	No	Yes
Cohort Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Cohort Dummy*Pre-reform Women Employment	Yes	Yes	Yes	No
Cohort Dummy*Pre-reform Number of Children	No	Yes	Yes	No
Individual Controls	Yes	Yes	Yes	Yes
Province Controls	No	No	Yes	No
Baseline Mean	0.700	0.700	0.700	0.700
R-squared	0.290	0.292	0.296	0.472
Obs	8,001	8,001	8,001	7,406

Notes: This table presents the DID estimates from additional robustness checks. Individual controls include child age, gender, urban-rural residence status, and family size. Province controls include per capita GDP and provincial population in each year. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A2: Restricted Estimation: Coefficients of the interactions between cohort dummies and the policy intensity

Age in 2010	Attendance	Attendance	Attendance
-2	0.797***	0.810***	0.781***
	(0.084)	(0.125)	(0.129)
-1	0.720***	0.670***	0.637***
	(0.129)	(0.150)	(0.145)
0	0.673***	0.578***	0.553***
	(0.115)	(0.152)	(0.138)
1	0.731***	0.668***	0.687***
	(0.126)	(0.150)	(0.148)
2	0.732***	0.664***	0.665***
	(0.115)	(0.143)	(0.135)
3	0.540***	0.505***	0.520***
	(0.068)	(0.107)	(0.105)
4	0.642***	0.640***	0.648***
	(0.062)	(0.109)	(0.108)
5	0.431***	0.406***	0.398***
	(0.071)	(0.094)	(0.093)
6	0.330***	0.429***	0.420***
	(0.054)	(0.053)	(0.058)
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
R-squared	0.289	0.290	0.322
Obs	10,297	10,297	10,297

Notes: This table presents the estimated effects on preschool attendance for each birth cohort. The omitted group is set to be children aged 7 to 15 in 2010. The dependent variable is a dummy variable reflecting whether a child is attending/has attended any preschools. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors are clustered at the province level and presented in parentheses. Significance levels: * 10%, ** 5%, ***1%.

Table A3: Effects on stunting: Specifications exploiting variation across birth cohorts

Panel A: Individuals aged below 6 or aged 9-13 in 2010, with binary intensity			
(Youngest cohort: Individuals aged -2 to 6 in 2010)			
	Stunting	Stunting	Stunting
(Aged below 6 in 2010)*High Intensity	0.004	0.005	0.006
	(0.027)	(0.027)	(0.028)
	[0.933]	[0.913]	[0.893]
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
Baseline Mean	0.233	0.233	0.233
R-squared	0.123	0.124	0.143
Obs	23,580	23,580	23,580
Panel B: Individuals aged below 6 or aged 9-13 in 2010, with continuous intensity			
(Youngest cohort: Individuals aged -2 to 6 in 2010)			
	Stunting	Stunting	Stunting
(Aged below 6 in 2010)*Policy intensity in that province	0.055	0.056	0.061
	(0.057)	(0.058)	(0.059)
	[0.393]	[0.387]	[0.373]
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Cohort Dummy*Pre-reform Enrollment	No	Yes	Yes
Controls	No	No	Yes
Baseline Mean	0.233	0.233	0.233
R-squared	0.123	0.124	0.144
Obs	23,580	23,580	23,580

Notes: This table presents the estimated effects on stunting by comparing the birth cohort aged below 6 in 2010 to the birth cohort aged 9 to 13 in 2010. Panel A uses a binary measure of policy intensity, and Panel B uses a continuous measure of policy intensity. The dependent variable is a dummy indicator of stunting. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A4: Additional robustness checks: Effects on underweight and see a doctor

Outcome	Underweight					See a doctor				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post*High intensity*Preschooler	-0.047**	-0.053**	-0.053**	-0.055**	-0.052*	0.124***	0.118**	0.107**	0.109**	0.089*
	(0.023)	(0.022)	(0.023)	(0.022)	(0.027)	(0.042)	(0.043)	(0.043)	(0.043)	(0.044)
	[0.073]*	[0.033]**	[0.060]*	[0.040]**		[0.013]**	[0.020]**	[0.033]**	[0.027]**	
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy*Pre-reform Women Employment	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No
Year Dummy*Pre-reform Number of Children	No	No	Yes	Yes	No	No	No	Yes	Yes	No
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province Controls	No	No	No	Yes	No	No	No	No	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village FE	No	No	No	No	Yes	No	No	No	No	Yes
R-squared	0.026	0.035	0.036	0.041	0.146	0.026	0.028	0.047	0.048	0.210
Obs	12,476	12,476	12,476	12,476	10,651	12,476	12,476	12,476	12,476	10,651

Notes: This table presents the DDD estimates from additional robustness checks. Individual controls include child age, gender, urban-rural residence status, and family size. Province controls include per capita GDP and provincial population in each year. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A5: Additional robustness checks: Sample without migrants or siblings

Outcome	Underweight		See a doctor	
	(1)	(2)	(3)	(4)
	Without Migrants	Without Siblings	Without Migrants	Without Siblings
Post*High intensity*Preschooler	-0.053**	-0.051*	0.128***	0.125***
	(0.021)	(0.025)	(0.043)	(0.043)
	[0.020]**	[0.067]*	[0.000]***	[0.007]***
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.143	0.145	0.701	0.705
R-squared	0.026	0.025	0.027	0.027
Obs	12,360	11,099	12,360	11,099

Notes: This table presents the DDD estimates from two additional robustness checks. In Columns (1) and (3), the sample comprises children without inter-province migration. In Columns (2) and (4), the sample comprises children without siblings. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A6: Effects on inter-province migration: DID vs DDD estimates

Panel A: Difference-in-difference estimates	
Sample: Preschooler (aged 3-6)	
Outcome	Migration
High intensity*Post	-0.014
	(0.011)
	[0.227]
Year Dummy*Pre-reform Enrollment	Yes
Controls	Yes
Sample Mean	0.008
R-squared	0.017
Obs	8,248
Panel B: Triple-difference estimates	
Sample: Preschooler (aged 3-6) and Toddlers (aged 1-2)	
Outcome	Migration
Post*High intensity*Preschooler	-0.012
	(0.013)
	[0.353]
Year Dummy*Pre-reform Enrollment	Yes
Controls	Yes
Sample Mean (3-6)	0.008
R-squared	0.012
Obs	12,476

Notes: This table presents the DID and DDD estimates for inter-province migration. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A7: Using continuous measures of policy intensity: DID vs DDD estimates

	Stunting	Underweight	Overweight	Sick last month	See a doctor
Panel A: Difference-in-difference estimates					
Sample: Preschooler (aged 3-6)					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*Intensity in that province	0.111	-0.154**	-0.017	-0.244*	0.240**
	(0.076)	(0.071)	(0.064)	(0.119)	(0.091)
	[0.153]	[0.060]*	[0.720]	[0.133]	[0.033]**
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.306	0.143	0.414	0.408	0.701
R-squared	0.052	0.027	0.021	0.023	0.027
Obs	8,248	8,248	8,248	8,248	8,248
Panel B: Triple-difference estimates					
Sample: Preschooler (aged 3-6) and Toddlers (aged 1-2)					
Outcome	Stunting	Underweight	Overweight	Sick last month	See a doctor
Post*Preschooler*Intensity in that province	0.156	-0.138*	-0.020	-0.185	0.341**
	(0.121)	(0.073)	(0.128)	(0.117)	(0.143)
	[0.167]	[0.092]*	[0.873]	[0.111]	[0.073]*
Year Dummy*Pre-reform Enrollment	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Baseline Mean (3-6)	0.306	0.143	0.414	0.408	0.701
R-squared	0.058	0.026	0.025	0.022	0.025
Obs	12,476	12,476	12,476	12,476	12,476

Notes: This table compares the DID estimates in Panel A and DDD estimates in Panel B when specifications measure the intensity as a continuous variable. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.

Table A8: Using continuous measures of policy intensity: DDD estimates from experiment of interest and control experiment

Panel A: Experiment of Interest						
Preschooler (aged 3-6) vs Toddlers (aged 1-2)						
Sample: Children aged 1-6, 2010-2016						
Outcome	Underweight	Underweight	Underweight	See a doctor	See a doctor	See a doctor
Post*Preschooler*Intensity in that province	-0.131*	-0.127*	-0.138*	0.347**	0.356**	0.341**
	(0.068)	(0.068)	(0.073)	(0.142)	(0.143)	(0.143)
	[0.090]*	[0.092]*	[0.092]*	[0.060]*	[0.067]*	[0.073]*
Year Dummy*Pre-reform Enrollment	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
R-squared	0.008	0.010	0.026	0.006	0.015	0.025
Obs	12,476	12,476	12,476	12,476	12,476	12,476
Panel B: Control Experiment						
Toddlers (aged 1-2) vs Infants (aged 0)						
Sample: Children aged 0-2, 2010-2016						
Outcome	Underweight	Underweight	Underweight	See a doctor	See a doctor	See a doctor
Post*Aged 1-2*Intensity in that province	0.056	0.057	0.049	0.085	0.077	0.069
	(0.221)	(0.221)	(0.224)	(0.459)	(0.447)	(0.434)
	[0.707]	[0.720]	[0.760]	[0.887]	[0.880]	[0.887]
Year Dummy*Pre-reform Enrollment	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
R-squared	0.014	0.017	0.045	0.008	0.017	0.023
Obs	5,264	5,264	5,264	5,264	5,264	5,264
Panel C: Control Experiment						
Children aged 11-14 vs Children aged 9-10						
Sample: Children aged 9-14, 2010-2014						
Outcome	Underweight	Underweight	Underweight	See a doctor	See a doctor	See a doctor
Post*Age11-14*Intensity in that province	0.152	0.158	0.187	0.194	0.193	0.256
	(0.183)	(0.183)	(0.166)	(0.184)	(0.187)	(0.172)
	[0.547]	[0.520]	[0.433]	[0.360]	[0.353]	[0.233]
Year Dummy*Pre-reform Enrollment	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
R-squared	0.017	0.018	0.033	0.006	0.007	0.037
Obs	8,471	8,471	8,471	8,471	8,471	8,471

Notes: This table compares the DDD estimates for health-related outcomes in the experiment of interest (Panel A) and two control experiments (Panels B and C). Specifications measure the intensity as a continuous variable. Controls include child age, gender, urban-rural residence status, and family size. Robust standard errors presented in parentheses are corrected for clustering at the province level. The p-values for wild bootstrapped standard errors are reported in squared brackets. Significance levels: * 10%, ** 5%, ***1%.