INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 02009

March 2021

Large-scale school meal programs and student health

Evidence from rural China

Jingxi Wang

Manuel A. Hernandez

Guoying Deng

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

AUTHORS

Jingxi Wang (wangjingxi211@hotmail.com) Guangdong University of Finance, China.

Manuel A. Hernandez (m.a.hernandez@cgiar.org), is a Senior Research Fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute (IFPRI), Washington, DC.

Guoying Deng (dengguoying@scu.edu.cn), Sichuan University, Chengdu, China.

Abstract

Reducing urban-rural gaps in child health and nutrition is one of the most difficult challenges faced by many countries. This paper evaluates the impact of the Nutrition Improvement Program (NIP), a large-scale school meal program in rural China, on the health and nutritional status of compulsory education students aged 6-16. We use data from multiple rounds of the China Health and Nutrition Survey between 2004-2015 and implement a quasi-experimental approach exploiting cross-county variations in program implementation. We find that NIP participation is, on average, associated with a higher height-for-age z-score in the order of 0.22-0.42 standard deviations. The impacts are larger among students in a better health condition but small or not significant among the most disadvantaged. We do not observe heterogeneous effects across several individual and household characteristics. We also do not find significant effects on Body Mass Index-for-age and weight-for-age z scores. The results suggest that NIP partially improved students' health over the first years of implementation, but more support is needed to achieve broader impacts that effectively reach all vulnerable students. Several robustness checks support our findings.

Keywords: School meal program; student health; nutrition; program evaluation; rural China
JEL Codes: I18; O15; H51; I38; R28

Acknowledgments

Guoying Deng gratefully acknowledges financial support from the Youth Outstanding Talent Training Program of Sichuan University (SKSYL201812 and 2019hhf-08) and National Natural Science Foundation of China (71773081).

1. Introduction

In recent decades, school feeding programs (SFPs) have been extensively implemented worldwide and are the most predominant social protection and assistance program (The World Bank, 2018). Roughly one in every two school children receive food at school (World Food Programme-WFP, 2019), although the number of children covered varies significantly across low-, middle-, and high-income countries. SFPs are intended to serve multiple purposes, including poverty relief, reduce vulnerability, promote educational achievement, as well as to improve the health and nutrition of school-age children (Bundy et al., 2009; Drake et al., 2017). On the latter objective, while the first years of life are the most critical window of growth and development, improving or maintaining an adequate nutritional status at childhood and early adolescence is key for the continuum of development (Best, 2010; Bundy et al., 2017). School nutrition programs can thereby have important implications for student health, cognition, and human capital accumulation (see, e.g., Kristjansson et al., 2007; Glewwe and Miguel, 2008; Jomaa et al., 2011; Krishnaratne et al., 2013; Watkins et al., 2015; Wang and Fawzi, 2020).

This paper examines the impact of the Nutrition Improvement Program (NIP) on the health and nutritional status of compulsory education students in rural China. NIP was officially launched in November 2011 by the Chinese government to improve the nutrition of children living in mainly deprived rural areas and enrolled in compulsory education (elementary and junior high schools).[1] The concentration of the program on impoverished rural locations responds to the important gaps in health and nutrition outcomes between urban and rural China, particularly for children in poor rural households.[2] A national survey conducted by the China Development Research Foundation (CDRF), reported in Zhaoyu (2011), reveals that 12% of poor students are stunted and the height of female and male boarding students in poverty-stricken rural areas is 9 and 11 centimeters shorter, respectively, than the average rural student (and weigh 7 and 10 kilograms less). School aged children in rural areas are also shorter and thinner than urban children (Zong and Li, 2014; Xu and Hang, 2017; Ao et al., 2019).[3]

Compared to national SFPs in other countries, NIP was launched relatively late. It started as a gradual pilot program that initially targeted compulsory education students in rural areas of 699 counties (also referred to as national pilot counties), relying on funds from the central government. The program then underwent rapid expansion, using both central and local funds, such that it is currently the third largest national SFP in the world (after the programs in

[1] Compulsory education in China comprises six years of primary education and three years of junior secondary education. Senior secondary education, which is not mandatory, takes three additional years.

[2] 99% of the poor in China live in rural areas (The World Bank, 2009).

[3] The stunting rate of children under five years old is more than four times larger in poor rural areas compared to urban areas (WFP, 2017).

India and Brazil). According to China's State Council, as of 2017 NIP had benefited more than 36 million Chinese rural students across 134,000 schools in close to 1,600 counties.[4] Two special features of NIP are that schools directly receive the funds equivalent to four yuan (65 US cents) per student per school day to improve or rebuild their cafeterias and provide free meals (mainly lunch), and the program includes a nutrition education component for students, parents, school staff, and caterers.

We use data from multiple rounds of the China Health and Nutrition Survey (CHNS) between 2004-2015 and exploit cross-county differences in program implementation. We implement a differences-in-differences (DID) and changes-in-changes (CIC) approach to compare variations in anthropometric outcomes of compulsory education students aged 6-16 located in rural areas of counties that were part of the NIP national pilot, relative to akin students located in rural areas not covered by NIP. We pay special attention to height-for-age as it is generally considered an overall (longer-term) indicator of nutritional status (de Onis and Branca, 2016).[5] The prevalence of low height-for-age among poor children in rural China has also been of a larger concern over past years compared to, for example, obesity or underweight (Li et al., 2009). It is thus worth assessing whether NIP has positive impacts on student height or whether it prevents further damage among those already stunted. We additionally examine impacts on Body Mass Index (BMI) and weight, which are two other relevant anthropometric measures for school-age children that can reflect short- as well as long-term nutritional status.

The estimation results show that program participation is associated with an average positive increase in the height-for-age z-score (HAZ) of 0.22-0.42 standard deviations. We observe larger impacts among students in a better health condition but small or no impacts among the most deprived and likely stunted students, such that the effects do not translate into a lower stunting rate. We do not find heterogeneous effects across numerous individual and household characteristics. Similarly, NIP participation is not associated with significant changes in BMI-for-age and weight-for-age z-scores (BMIz and WAZ). Our findings suggest that NIP partially improved students' health over the first years of implementation, but more targeted and intensive support is needed to achieve wider impacts, especially among the most vulnerable that also require earlier interventions to effectively fight malnutrition and food insecurity.

The study contributes to the literature that evaluate the health and nutrition impacts of SFPs in developing countries. Jomaa et al. (2011) perform an extensive review of SFPs across several countries and find that school feeding seems to be effective on helping students receive

---

[4] http://english.www.gov.cn/news/top_news/2017/06/03/content_281475675232760.htm (accessed January 2021).

[5] Height-for-age during early childhood is also regarded as a good predictor of human capital accumulation (Victora et al., 2008; Alderman et al., 2017).

sufficient nutrients and improve their micronutrient status, but the positive impacts on growth are less conclusive. It is difficult, however, to draw general conclusions on SFPs impacts as programs in each country may vary on policy design, objectives and targeting, program components (i.e., in-school meals or take-home rations; meals combined or not with micronutrient supplementation or food fortification; accompanying nutrition education interventions), coordination and implementation aspects, and households' response, such that health and nutritional impacts are expected to vary across programs. Gelli and Daryanani (2013) and Kristjansson et al. (2016) further show that SFPs costs per child differ significantly across countries; for low- and middle-income countries these range from nine to 270 US dollars per year. Alderman and Bundy (2012) note that differences in evaluation methods across studies could be another source of variation in the results obtained.

Relatively recent studies examining the impacts of large-scale SFPs on students' anthropometric outcomes include Buttenheim et al. (2011), Nkhoma et al. (2013), Singh et al. (2014), and Gelli et al. (2019), which rely on experimental or quasi-experimental evaluation designs.[6] Buttenheim et al. (2011) focus on SFPs in Lao PDR, which include on-site feeding, take-home rations and a combination of both, and do not find conclusive evidence that the feeding programs improved children's height or weight, with varying results by feeding modality. Nkhoma et al. (2013) find that provision of a daily ration of corn-soy blend porridge in the Malawian SFP accelerated the growth of lean muscle in children but there are no effects on height and weight. Singh et al. (2014) examine the national SFP in India (Midday Meal Scheme) that offers a daily cooked meal and obtain important positive effects on both the height and weight of children whose households self-report having suffered from drought. Gelli et al. (2019) show that a large-scale school meals program in Ghana, which provide a daily hot meal, do not have overall effects on children's height or BMI but do have localized effects on height among girls and younger children (especially among children living in poor households).

Our study also adds to the ongoing discussion in China about the impacts of NIP on the health and nutritional status of students in targeted rural areas. Previous related studies have either solely documented changes over time in height or weight outcomes among participating students in selected provinces and find varying patterns (e.g., Deng et al., 2016; Zhan et al., 2019) or assess immediate impacts focusing on a specific school grade with mixed results (e.g., Wang et al., 2019).[7] About the latter, Wang et al. (2019) use 2013-2014 data from the China

[6] The literature on the links between SFPs and health and nutrition outcomes is extent. Other related studies include Grillenberger et al. (2003), van Stuijvenberg (2005), Neumann et al. (2007), Afridi (2010), Kazianga et al. (2014), Adelman et al. (2019), and Berry et al. (2020).

[7] Deng et al (2016) compare changes between 2012 and 2015 in the nutritional status of NIP students aged 6-14 in the rural area of Hunan province, and do not find that their height or weight in 2015 is necessarily higher than the average rural student; Zhan et al. (2019) analyze changes in health and nutrition indicators between 2012 and

Education Panel Survey (CEPS) for seventh grade students and find a positive effect of NIP on height and no effect on weight. Their analysis relies on a DID model with propensity score matching and concentrates on locations where NIP started after 2013, which were mainly local initiatives.[8]

Overall, the paper aims to generate evidence regarding the impacts of NIP national pilot on anthropometric measures of elementary and junior high school students, after three years of program implementation. The data sample period, including four pre-treatment survey rounds, and wide geographic coverage allows us to implement several analyses around model identification and construct different plausible counterfactuals. We also discuss potential selection issues given the dataset and program nature, and perform multiple robustness checks to assess the reliability of our results. The derivation of distributional treatment effects, in addition to average effects, is intended to better inform policy regarding the impacts of NIP on both more and less disadvantaged students within targeted areas.

The remainder of the paper is organized as follows. Section 2 provides background information on nutrition assistance initiatives from the Chinese government in rural areas and discusses in more detail the evaluated SFP. Section 3 describes the data and empirical methodology followed to assess the association between NIP participation and students' health and nutritional status. Section 4 presents and discusses the estimation results, while Section 5 concludes.

2. Background

China has undertaken numerous nutrition assistance initiatives over the past decades targeting poor children, which are mainly concentrated in rural areas. In collaboration with the World Food Programme (WFP), the government has been combating malnutrition in deprived areas since 1979, both in the form of direct food aid and resources to improve local production capacities (WFP, 2009).[9] In the 1990s, an Outline of China's Child Development Plan was promulgated to integrate child development into national economic and social development plans, while in 2001 the central government promised to establish a nutrition security system for children and adolescents in poverty-stricken areas. In 2007, the China Development Research Foundation (CDRF), a public entity established to advance good governance and policy to promote economic and social development, launched the Nutrition Improvement

---

2016 among NIP students in grades 1-8 in Guizhou, Heilongjiang, and Hubei provinces, and find that malnutrition decreased but obesity increased.

[8] Prior to NIP, Qi and Zhao (2012) evaluated an experimental (local) pilot nutrition improvement program for elementary school students in Hebei and Guangxi provinces using a DID approach, and do not find statistically significant effects on height and weight.

[9] Between 1979 and 2005 the Chinese government received more than one billion US dollars from WFP.

Program for Boarding Primary School Students in two poor counties in Hebei and Guangxi provinces. This three-year pilot program was the first experimental policy project in China to rely on existing school infrastructure to implement nutrition interventions and study their effects on students.[10] Other analogous initiatives, including infant and child nutrition programs in specific disadvantaged regions, were implemented in parallel during the second half of the 2000s, but more comprehensive nutrition programs covering a wider range of rural areas across the country remained missing.

Towards the end of 2011, the government launched the Nutrition Improvement Program (NIP), a nationwide school meal program for compulsory education students in mainly impoverished rural areas. NIP initially targeted 699 national pilot counties across 21 provinces in the country, which were gradually incorporated into the program during its first two years and were exclusively funded by the central government;[11] by 2017, the program had been extended (including many locally funded initiatives) to 1,590 counties across 29 provinces.

As noted by Zhang et al. (2014), NIP aims to improve the nutritional status of rural students and reduce the gap between urban and rural populations. The program consists in entitling all enrolled students across the selected counties to a school meal allowance equivalent to four yuan per school day.[12] Considering that there are 200 school days per year, participating students receive school meal subsidies totaling 800 yuan per year. This is equivalent to 7.6% of the annual per capita average disposable income for rural Chinese residents in 2014. The financial aid can be used to outsource the student's food supply and build or improve school cafeterias. The program is complemented with the provision of nutritional information and promotion of healthy diets. By 2017, the central government had allocated around 160 billion yuan to the program since its launch (CDRF, 2017).

NIP has certain features that are different from many other national SFPs implemented across the world. First, the subsidy is not directly transferred to students; rather, the central government provides the funds to primary and secondary rural schools in the selected counties based on their number of students. Schools are then responsible for procuring, preparing, and distributing the food among their students with particular emphasis on the provision of lunch; the progress report from CDRF (2017) indicates that 56% of the monitored counties offer free lunch, 36% offer free lunch and breakfast (or free lunch and snacks), and the remaining 8% only offer free breakfast. Second, schools are assisted by the local and central government to improve or rebuild their cafeterias. Third, the program has a nutrition education component.

[10] The selected counties were Chongli (10 elementary schools) in Hebei, which is in the north of China, and Du'an (3 elementary schools) in Guangxi, which is in the south.

[11] Most of the national pilot counties started in 2012.

[12] The amount was three yuan when the program started and was then increased to four yuan in November 2014.

While the food prepared does not have to follow specific nutrient-based standards, meat, eggs, milk, and other highly nutritious foods are explicitly recommended; among the monitored counties, the main food ingredients identified include rice or noodles, meat, oil, eggs, and the four most common vegetables are carrots, potatoes, tomatoes, and cabbage. Similarly, a Dietary Nutrition Guidelines for Students in Rural Areas is distributed to promote increased knowledge of nutrition and healthy eating among students, parents, teachers, school administrators, and caterers.[13]

We exploit cross-county variations in the implementation of NIP to assess the impact of the program on students' health and nutritional status. We compare before-after changes in anthropometric measures of rural students from a subset of counties that were included in the initial national pilot program, relative to students in other counties that were not part of NIP. We account in the analysis for individual and household characteristics as well as by location characteristics at the village level, considering the additional variation in socioeconomic development and thereby available resources across localities. While NIP is directly funded by the central government, local governments are largely responsible for the program implementation in their area as well as for financing compulsory education, health, and nutrition services in general, which also affect the health and nutritional status of children. We also evaluate potential heterogenous effects across several individual and household characteristics.

3. Empirical approach

3.1 Data

The data used in the study are from the China Health and Nutrition Survey (CHNS), which was designed to examine the extent to which social, economic, and demographic changes in the country affect the health and nutritional status of the population.[14] The first round of CHNS data was collected in 1989 and nine additional rounds were collected between 1991 and 2015. We use data from seven provinces and one autonomous region that have been continuously followed by the CHNS in 2004, 2006, 2009, 2011, and 2015. These include Guizhou, Henan, Hubei, Hunan, Jiangsu, Liaoning, Shandong, and Guangxi. The survey uses a multistage, random cluster process to draw samples from each province (region). Counties in these regions are stratified by income (low, medium, and high), and four counties are randomly selected in

[13] CDRF started a dedicated monitoring platform (Sunshine School Meals) in 2015, commissioned by the National Student Nutrition Office of the Ministry of Education, to monitor the program implementation in 100 selected counties across 13 provinces.

[14] The CHNS is an international collaborative project launched jointly by the Carolina Population Center at the University of North Carolina at Chapel Hill and the Chinese Center for Disease Control and Prevention (CCDC). For further details see https://www.cpc.unc.edu/projects/china (accessed January 2021).

each region using a weighted sampling scheme; towns and villages within each county are then randomly selected. The CHNS is longitudinal in nature where the same counties, towns, and villages are generally visited, but not necessarily the same households are continuously followed over time due to the high level of attrition across waves (see, e.g., Popkin et al., 2010). We return to this discussion below, in the context of our empirical approach and sample of interest.

Since CHNS only publishes province names, we follow the location identification approach implemented by Chyi and Zhou (2014) to identify the specific counties in the data. This implies comparing total county area and population (as reported in the CHNS community data), with area and population details from various Province Statistical Yearbooks, which is the source of information in the CHNS sample design. Among the 36 counties identified in the seven provinces and autonomous region surveyed between 2004-2015, 32 counties include rural areas and eight of them were part of the 699 national pilot counties covered by NIP during its initial years (verified in the 2018 Bulletin of the Ministry of Education). The villages in these eight counties (24 communities) constitute our treatment area, while the villages across the remaining 24 counties (72 communities) serve for comparison purposes.[15] We consider different alternative control groups as the areas included in the initial national pilot are not necessarily directly comparable to all excluded areas.

Our relevant sample are compulsory education students that are enrolled into elementary and junior high schools in rural areas (covered and not covered by NIP). We further restrict the sample to students aged 6 to 16, which is the common age range for compulsory education in rural areas given school entry delays and regular grade repetition.[16] The total working sample comprises 2,949 observations between 2004 and 2015, where 34.5% correspond to NIP treatment areas, and 65.5% to non-treatment areas used for comparison purposes.

The key outcome of interest is the height-for-age z-score (HAZ) that measures the distance in standard deviations of the height of the student from the World Health Organization (WHO) growth reference for a child or adolescent of the same age and sex.[17] A negative HAZ

[15] Of the 24 treatment villages in our sample, 22 are observed across all five survey rounds and 2 in four rounds; of the 72 non-treatment villages, 54 are observed in five rounds, 9 in four rounds, 6 in three rounds, and 3 in two rounds.

[16] The results are qualitatively similar if we alternatively restrict the sample age to 6-15 years old, which is the expected age range for compulsory education across the country. China's Compulsory Education Law stipulates that all children who have reached the age of six (in areas where conditions are less favorable, they may wait until the age of seven) should be sent to school by their parents or legal guardians to receive and complete compulsory education.

[17] https://www.who.int/toolkits/growth-reference-data-for-5to19-years (accessed January 2021). See also de Onis et al. (2007).

value reveals that the student has a low height for their age, and values less than -2 are indicative of stunting. As a general indicator of overall (longer-term) nutritional status, we are particularly interested in assessing the impact of NIP on HAZ. We also analyze below the effects of the program on the probability of being stunted (if HAZ < -2) as well as on two other anthropometric measures relevant for children in school age (BMI-for-age z-score or BMIz and weight-for-age z-score or WAZ).[18]

Regarding control variables, the CHNS dataset contains several demographic and socioeconomic variables at the individual, household, and village level, which are potentially correlated with the health and nutritional status of students. Individual characteristics include gender, age, and if single child. Household characteristics include father's and mother's education level, age of mother at childbirth, household size, and per capita income. Village characteristics include scores (on a scale 0-10) for the overall economic environment, quality of health services, and availability of social services in the village directly calculated by the CHNS team and provided in the dataset.

Table 1 presents descriptive statistics of the outcome variables and covariates for the full sample period and by year. We observe an important improvement in health and nutritional indicators among the sampled rural students during the period of analysis. The average HAZ shows a monotonic increase and reversal from -0.44 in 2004 to 0.27 in 2015, while the stunting rate decreased from 8.2% to 1.7%. BMIz exhibit an increase from -0.33 to -0.06 and WAZ from -0.38 to 0.2.[19] These improvements are generally in line with the national trends reported in the UNICEF/WHO/World Bank joint child malnutrition estimates for children under 5 years old in China.[20] Similar to rural areas nationwide, we find a continuous increase in the per capita household income and the socioeconomic environment at the community (village) level captured by the economic, health quality, and social services' scores.[21] Overall, we have a balanced number of female and male students, which are generally a single child and whose parents have 7-8 years of schooling.

---

[18] BMIz and WAZ capture, respectively, the student's BMI and weight relative to WHO's reference (median) values for someone of the same age and sex. Similar to HAZ, BMIz is applicable for all students in our sample, while WAZ only applies for students up to 10 years old.

[19] BMIz is only available for 2,659 observations in the full sample, while WAZ for 1,392 observations (out of 1,519 observations in the full sample corresponding to students aged 6-10).

[20] https://www.who.int/news/item/31-03-2020-unicef-who-wb-jme-group-new-data (accessed January 2021).

[21] The Gross Domestic Product (GDP) in China increased around 9.8% on an annual basis between 2004 and 2015 based on official estimates from the National Bureau of Statistics, while the GDP in rural areas is estimated to have annually increased by 3.8% according to the Green Books on China's rural areas, compiled by the Rural Development Institute of Chinese Academy of Social Sciences and the Department of Rural Surveys of National Bureau of Statistics.

3.2 Methodology

We follow a quasi-experimental approach exploiting cross-county differences in NIP implementation across our sample to assess the impact of the program on the health and nutritional status of compulsory education students. We compare changes in anthropometric measures among rural students located in counties covered by NIP, before and after the program implementation, relative to rural students located in counties not covered by NIP. We implement both a standard differences-in-differences (DID) model to derive an average treatment effect on the treated (ATT) and a changes-in-changes (CIC) model to derive distributional treatment effects on the treated (DTT).

The DID regression model is given by,

$$z_{ijt} = \beta_0 + \beta_1 NIP_j + \beta_2 t_{NIP} + \beta_3 NIP_j * t_{NIP} + \gamma_1 X_{ijt} + \gamma_2 V_{jt} + u_{ijt} \quad (1)$$

where z_{ijt} is the HAZ or height-for-age z-score of student i located in rural village j at survey year (time) t; NIP_j is the treatment variable equal to one if the student is in a village where NIP was implemented, and zero otherwise; t_{NIP} is a dummy variable equal to one for the treatment period (i.e., year=2015, which is three years after the program started in our treatment areas), and zero otherwise;[22] X_{ijt} and V_{jt} are, respectively, vectors of individual and household covariates and time-varying village controls (mentioned in the previous section) that likely influence a students' health; and u_{ijt} is an error term defined as $u_{ijt} = \tau_j + \theta_t + e_{ijt}$, which includes village fixed effects (τ_j) to account for unobserved time-invariant differences across locations, year fixed effects (θ_t) to control for cohort health trends that are common across all areas,[23] and a white noise term e_{ijt}.

The parameter of interest in equation (1) is β_3, which is the DID estimator measuring the average effect of NIP on students' height-for-age standardized scores. This estimator is equivalent to the ATT of the program under two assumptions (Ashenfelter and Card, 1985; Abadie, 2005),[24] and thereby approximates $\Delta_{\text{ATT}} \equiv \text{E}\left[HAZ^T - HAZ^{NT} | NIP_j = 1, t_{NIP} = 1\right]$, where HAZ^T is the height-for-age z-score of rural students in NIP areas after the program

---

[22] NIP was officially launched in November 2011, but the program started in 2012 across the eight treatment counties in our sample.

[23] The results are robust to alternatively including a time trend and squared term or village-specific time trends.

[24] The first assumption is the common trend assumption (explored below) that assumes that in the absence of the program, the difference between the treatment and control group would remain constant over time. The second assumption is the full compliance assumption that implies a full treatment rate among the treated and a zero-treatment rate among the control group (and prior to the program), which is satisfied by the universal nature of NIP to all compulsory education students across the intervention areas.

started and HAZ^{NT} is the height-for-age z-score of these students had the program not been implemented.

The second approach followed is the CIC model proposed by Athey and Imbens (2006). This model permits us to assess changes over the entire distribution of students' height-for-age z-scores in NIP versus non-NIP areas, before and after the program implementation. In particular, we approximate the DTT of the program defined as $\Delta_{DTT}(q) \equiv F^{-1}_{HAZ^T|NIP_j=1,t_{NIP}=1}(q) - F^{-1}_{HAZ^{NT}|NIP_j=1,t_{NIP}=1}(q)$, where $q \in [0,1]$ is a given percentile and F is the cumulative distribution function (cdf) of HAZ. We can examine with this model if NIP has a differentiated impact over students with varying z-scores.

Compared to the DID framework where we are interested in deriving average counterfactual outcomes (z-scores) in the absence of the program, in the CIC model we estimate the whole counterfactual distribution of outcomes.[25] Athey and Imbens (2006) show that under certain assumptions the counterfactual distribution of HAZ can be described as $F\left(HAZ^{NT}|NIP_j = 1, t_{NIP} = 1\right) \equiv F_{HAZ^{NT},11}(z) = F_{HAZ^{NT},10}\left(F^{-1}_{HAZ^{NT},00}\left(F_{HAZ^{NT},01}(z)\right)\right)$.[26] By inverse transformation, we can calculate the effect of NIP for a given percentile q as,

$$\delta^{CIC}_q = F^{-1}_{HAZ^T,11}(q) - F^{-1}_{HAZ^{NT},01}\left(F_{HAZ^{NT},00}\left(F^{-1}_{HAZ^{NT},10}(q)\right)\right). \tag{2}$$

Function F and its inverse are estimated using the standard plug-in method to obtain an empirical cdf. As δ^{CIC}_q has an asymptotically normal distribution, we use bootstrapping to derive its variance. We also account for covariates (as in the DID model) by first regressing a z-score model with all the controls specified in equation (1) and then implementing the CIC method on the corresponding residuals.

3.2.1 Identification

A central aspect to our empirical approach is the assumption that, in the absence of treatment, the difference between the treatment and control group would remain constant such that the estimation of the trend in one group can assist in eliminating the trend in the other group. To assess the plausibility of this common (parallel) trend assumption, Panel A of Figure 1 plots the average HAZ for rural students in NIP areas (treatment group) and non-NIP areas (control

---

[25] The CIC model is more general than the DID model in that the distribution of unobservables may vary across the treatment and control groups in arbitrary means (e.g., groups may differ in terms of the distribution of outcomes in the program's absence as well as on the effects of the program).

[26] The conditions to derive the counterfactual distribution of HAZ are four: i) in the program's absence, z-scores satisfy the following relationship $HAZ^{NT} = k(e, t_{NIP})$, where *e* represents the set of unobservables; ii) HAZ^{NT} is a monotone increasing function of *e*; iii) any differences across groups remain stable over time, i.e., $e \perp t_{NIP}|NIP$; and iv) $\Gamma_1 \subseteq \Gamma_0$, where Γ_1 and Γ_0 are the domains of $F(e|NIP = 1)$ and $F(e|NIP = 0)$.

group) in each survey year. We effectively observe that students' HAZ exhibit a parallel upward trend across the treatment and control groups over multiple years prior to the program implementation. While the health status of students in NIP areas is lower than those in non-NIP areas, the difference in the average z-score remains very stable across the two groups between 2004 and 2011 (0.57-0.59 standard deviations); after the program started, however, the difference considerably narrows down (0.27 standard deviations in 2015).

A standard event-study analysis on the evolution of HAZ across the treatment and comparison cohorts provides additional support to the parallel trend assumption. We regress HAZ on time indicators for each survey year (base category is 2004), interactions of the year indicators with the treatment dummy variable, and the set of control variables used in equation (1). Panel B of Figure 1 reports the estimated interaction terms between the year indicators and the treatment variable that capture the differential changes over time in HAZ between NIP and non-NIP areas. Up to 2011, we find no statistically significant differences in the evolution of the z-score across areas, as opposed to 2015 when NIP is already underway.

Considering that NIP mainly targets poverty-stricken counties (at least during the first years of its implementation) and that the areas included in the initial national pilot in our sample are not fully comparable to all rural areas not included, we consider two alternative control groups to model the counterfactual state of the treatment group. Following Blundell and Costa Dias (2000) and Blundell et al. (2004), matching treatment and non-treatment areas based on observable characteristics prior to the implementation of the DID (and CIC) model can aid to better account for potential unobservable confounders in the analysis.[27] We accordingly pre-balance our data by matching each treatment village to the closest non-treatment village based on average student, household, and village characteristics using Mahalanobis distance matching.[28] We implement matching without and with replacement where the 24 treatment villages are, respectively, matched with 24 and 19 comparison villages.

Table A.1 in the Appendix reports average differences in observable characteristics between the treatment and control areas, for the full and matched village samples. While the treatment and control villages in the full sample do not exhibit major differences across most characteristics, the matching exercise helps to further reduce several of these discrepancies.[29] For comparability purposes, we report the results of the DID and CIC models for the full sample

[27] See also Ho et al. (2007) for a detailed discussion on implementing regression models using matched samples to reduce the dependence of the validity of the estimated treatment effect on a correct model specification.

[28] We use Mahalanobis distance matching (instead of propensity score matching) given the relatively small number of villages to generate a propensity score for the subsequent pairing. Mahalanobis matching pairs areas that are closer to each other across all covariates.

[29] The pre-balancing basically helps to limit the comparison between NIP and non-NIP villages that share closer average individual and household characteristics.

and the matched samples at the village level. We also evaluate below the sensitivity of our findings to additional control groups.

Certainly, we cannot completely discard unobservable confounding factors not accounted for in the estimations that could be influencing our results, but we are unaware of other major rural programs or interventions that could have differentially affected student health outcomes between treatment and comparison areas during the period of study. The relevant policies we were able to identify are the Closures and Mergers of Rural Schools initiative launched in 2001 to better balance the teacher-student ratio, improve scale efficiency, and reduce the heavy financial burden of the rural compulsory education system; the Two Waivers and One Subsidy program launched in 2003 to cover part of tuition and other educational expenses of students in elementary and junior high school; and the New Rural Cooperative Medical Scheme launched in 2003 to subsidize medical insurance and coverage in rural areas.[30] All these programs, however, were fully operational well before 2011 and affected all rural students (households).

3.2.2 Survey attrition and school enrollment

As noted above, there is a high degree of household (individual) attrition across CHNS waves. In addition, our population of interest are students between 6 and 16 years old in each survey wave, which occurs every two to four years. Consequently, barely between 16-39% of individuals in our sample are observed across two waves that prevents us from fully exploiting the panel aspect of CHNS using an individual fixed effects model. We follow a repeated cross-section approach as the one described earlier with village (and time) fixed effects, which is in line with other recent studies that use data from multiple waves of CHNS.[31] We still evaluate below the sensitivity of our results when accounting for within-individual correlation and allowing for individual random effects. We also construct a pseudo panel at the village level to assess the impact of NIP on student health under this alternate estimation framework.

Considering the attrition patterns in our data, it is worth exploring if there are systematic differences between treatment and comparison counties in the number of instances a student is observed as these could potentially bias our results. To test whether attrition occurs with equal likelihood between areas, we regress an indicator variable equal to one if a child 6-

---

[30] For further details on these policies, see, e.g., Chyi and Zhou (2014), Liang and Wang (2020), and Zhang et al. (2020).

[31] E.g., Chyi and Zhou (2014), Fraumeni et al. (2019), Kerr (2019), Chen et al. (2020), and Yang and Bansak (2020). Kong et al. (2019) implement an individual fixed effects model, but they focus on urban areas using CHNS waves of 1997, 2000, and 2004. We tried implementing a child fixed effects model among the few students observed in both the 2011 and 2015 waves, but the estimated effects have a high standard error and are not statistically different from zero.

16 years old (whether enrolled or not to school) is only observed once on survey year indicators, interactions of year indicators with a dummy variable for NIP counties, and the set of covariates described in equation (1). As reported in column (1) of Table A.2 in the Appendix, the coefficients of the year-NIP interaction terms are not statistically significant at conventional levels, suggesting that attrition is equally likely among treatment and comparison rural areas over time.

These findings are further suggestive that any underlying migration dynamics, which could explain part of the observed attrition, is likely similar across areas and do not influence our estimations. Movement of an entire family from one place to another in China is structurally difficult as people are controlled by their registered residency (Hukou). In addition, it is highly unlikely that a family moved to a NIP area as the program mainly targets relatively poor rural areas and there are strong disincentives to move to such areas.

Another source of potential selection bias are eventual variations in the rate of school enrollment among treatment and comparison counties as NIP is applicable to all students attending elementary or junior high school. We should not expect though major differences across areas as one of the main factors explaining school enrollment decisions in China are schooling costs (Brown and Park, 2002; Connelly and Zheng, 2003), and providing affordable education at primary and junior high school levels across all rural areas have been a key government goal since the early 2000s. Not surprising, the rural enrollment rate for children aged 6-16 in our sample is over 92%.

Column (2) of Table A.2 in the Appendix reports the results of regressing an indicator variable equal to one if a child 6-16 years old is enrolled into school on year indicators, interactions of year indicators with a dummy variable for NIP counties, and the set of control variables defined in equation (1). The lack of significance of the interaction terms confirms no systematic differences in enrollment rates between treatment and control areas over time, including 2015 which also suggests that NIP did not affect the likelihood of enrollment.

4. Results

We now turn to the empirical results. We first present our base results regarding the impact of NIP on height-for-age using DID and CIC models. We then perform several robustness checks and additional estimations to assess the validity and sensitivity of our results. We also explore heterogeneous effects of the program along several individual and household dimensions. Finally, we examine NIP impacts on other anthropometric measures of interest.

4.1. Base results

Table 2 shows the estimation results of the DID model defined in equation (1), which allows us to approximate the ATT of the program on height-for-age of compulsory education students.

The first two columns report the results using the full sample and the remaining four columns using the matched samples (without and with replacement) at the village level. Columns (1), (3), and (5) include village and time fixed effects, which are omitted for ease of presentation, while in columns (2), (4), and (6) we add individual, household, and village (time-varying) controls.[32] The reported standard errors are robust and clustered by village and year. The statistical significance of the results is though not sensitive to alternatively clustering by village or by county-year.

The results indicate that NIP participation is associated with a positive increase in the average height-for-age of enrolled students. Based on the full sample, students in NIP treatment areas increased their HAZ by an additional 0.22-0.25 standard deviations (SDs) after the program implementation, relative to students in non-NIP areas. If we narrow down the comparison to treatment and non-treatment areas that share closer observable characteristics, the estimated increase is though larger and equal to 0.35-0.4 and 0.36-0.42 SDs for the village samples matched without and with replacement, respectively. Hence, the direct provision of school meals accompanied with nutrition education interventions seem to be correlated with a general improvement of student health in rural areas measured through HAZ. The effect is larger when comparing more similar areas as well as when accounting for individual covariates, which constitute our preferred specifications.

The estimated effects are larger than the one-year NIP impacts on HAZ derived by Wang et al. (2019) for seventh grade students using 2013-2014 CEPS data (0.11 SDs). Similarly, the results fall in the middle to upper range of the effects of certain large-scale school meal programs (modalities) in other countries. Buttenheim et al. (2011) find that take-home rations implemented by WFP in Lao PDR increase the HAZ of children aged 3-10 by 0.29 SDs, although on-site feeding and the combination of on-site feeding and take-home rations do not have any effect. Singh et al. (2014) show that the SFP in India (Midday Meal Scheme) increase HAZ of children in primary schools with an average age of 5.5 years by 0.27 SDs, but the effect is not statistically significant; however, when distinguishing between drought affected and non-affected students, they find a HAZ increase of about 0.81 SDs among the former, which more than compensates the negative drought impact. Gelli et al. (2019) find that while the provision of one meal in public primary schools in Ghana has no effect on HAZ across all children aged 5-15, it has a positive effect on girls (0.12 SDs), children aged 5-8 (0.12 SDs), and children aged 5-8 living in poverty conditions (0.22 SDs).

Regarding the covariates included in the regressions, we observe that girls have a 0.16-0.19 SDs lower HAZ than boys, which is indicative of a female-male child nutritional gap in

[32] The NIP dummy variable is naturally dropped due to the inclusion of location fixed effects and the treatment period indicator is absorbed by the time fixed effects.

rural areas (see Ren et al., 2014). We also find that HAZ is decreasing in age, students that are single child show a 0.12-0.14 SDs higher HAZ than those with siblings, and both parents' years of schooling is positively associated with the student height-for-age. The village economic environment score is though negatively correlated with student HAZ, although the coefficient is not statistically significant across all samples.[33]

Turning to the distributional effects of NIP based on the CIC model estimations, Figure 2 depicts the DTT of the program on student HAZ defined in equation (2). Panel A corresponds to the results using the full sample and Panels B and C to the results using the village samples matched without and with replacement. The horizontal axis in the figures indicate the percentile of the HAZ distribution (expressed in decimals from 0.05 or the 5th percentile to 0.95 or the 95th percentile) and the vertical axis is the estimated percentile treatment effect on the treated. The dashed lines are 95% confidence intervals obtained from 2,000 bootstrap replications.

The percentile treatment effect is an upward-sloping curve implying that NIP has a higher effect on students with a better anthropometric status (i.e., with a higher HAZ). When using the full sample (Panel A), the treatment effects for the 20%, 40%, 60% and 80% HAZ percentiles are -0.22, -0.06, 0.07, and 0.41 SDs, although we only observe statistically significant effects in the upper percentiles (above the 70% percentile). For the matched village samples, the corresponding treatment effects are 0.08, 0.28, 0.35, and 0.59 SDs in Panel B and 0.15, 0.31, 0.45, and 0.60 SDs in Panel C, and the impacts are generally significant from the 25% percentile onwards.[34] These patterns are naturally consistent with the higher ATT estimates from the DID model for the matched samples, compared to the full sample.

Overall, NIP appears to work or work better on less needy rural students. For more disadvantaged students, which are more likely to be stunted or suffer from malnutrition, providing school meals combined with school-based nutrition education does not seem to be enough to improve their health and nutritional status captured through HAZ. This finding is in line with the difficulties of attending chronic malnutrition at childhood or adolescence rather than at very early stages of life (The Lancet, 2008; SUN Movement, 2010; IFPRI, 2016).[35]

[33] A possible explanation for this apparent counterintuitive correlation could be that economic development may result in higher air (and water) pollution, which may affect children's health and growth, especially in China that heavily relies on coal for its energy needs (see, e.g., Millman et al., 2008).

[34] The slightly wider confidence intervals in Panel C (considering only villages matched with replacement) are due to the smaller number of observations across the different percentiles used to derive the reported estimates.

[35] There is consensus in the nutrition literature that the most critical period for growth and development is from conception to 24 months of age, but there is some debate on whether stunting can be fully reversed beyond early stages of life (see, e.g., Prentice et al., 2013; Leroy et al., 2013).

4.2 Robustness checks

We perform two standard placebo tests to evaluate the robustness of our estimations. First, we estimate impacts on groups of population that should not have been affected by the program, referred to as two pseudo treatment groups. As NIP targets students in rural compulsory education, students in urban areas across NIP counties and rural children who do not attend school in these counties should not be affected by the program. The first two columns of Table A.3 in the Appendix report the pseudo treatment effects on HAZ of estimating the DID model defined in equation (1) for enrolled students aged 6-16 in urban areas (first column) and non-school children aged 6-16 in rural areas (second column). The DID estimator is not statistically significant at conventional levels in both cases.

Second, we implement permutation tests where we use our base sample of enrolled rural students aged 6-16 but assume that NIP started earlier. We should expect no impacts prior to 2011. Columns (3) through (5) of Table A.3 in the Appendix report the resulting DID estimators assuming that NIP started after 2004, 2006, and 2009, respectively. In all three cases the pseudo treatment effects on HAZ are not statistically significant. All these robustness checks provide additional support to our empirical approach.

4.3. Additional estimations

4.3.1 Alternative comparison groups

We assess the sensitivity of our results to additional comparison groups. We consider three alternative controls to approximate the counterfactual state of the treatment group. First, instead of pre-matching at the village level, we pair the 8 NIP counties with non-NIP counties using average student, household, and location characteristics; in this case, Mahalanobis distance matching without and with replacement selects 8 non-treatment counties such that we limit the comparison to compulsory rural students in these counties. Second, given that counties within the same province are geographically closer and are expected to share more common characteristics or be affected by similar factors (including potential time-varying unobservables), we narrow down the comparison to rural students of 8 non-NIP counties that belong to the same provinces as the treatment counties. Third, we limit the comparison to rural students of 6 counties that were included in local NIP pilots after 2015 as the counties where the program continued to be expanded could share some specific (unobservable) characteristics with the counties included in the initial national pilot.

Table 3 presents the ATT of the program on HAZ when estimating the DID model using these alternative comparison groups. Similar to our base results, we find statistically significant effects in all three cases. In column (1), when comparing students in pre-matched counties, the estimated effect of NIP on HAZ is 0.45 SDs. In column (2), when comparing students within same provinces, the estimated effect is 0.36 SDs, while in column (3) that compares students in treatment counties with those in subsequent local pilot counties, the effect is 0.35 SDs.

4.3.2 Alternative panel approach

Given that a fraction of students in our sample is observed in more than one instance, it is still worth evaluating the sensitivity of our estimations when alternatively controlling for within-individual correlation and when allowing for the presence of individual random effects. Table 4 reports the results of these two additional estimation exercises. The upper panel of the table shows that the statistical significance of the DID estimators is not affected by clustering the standard errors at the individual level. The lower panel of the table reports, in turn, the DID estimators that result from implementing a generalized least squares (GLS) individual random effects model. The estimated effects of NIP on student HAZ are rather close to the base estimates and range from 0.2 SDs using the full sample to 0.37-0.39 SDs using the matched village samples.[36]

4.3.3 Pseudo panel at the village level

We also construct a pseudo panel at the village level to evaluate the impact of NIP on HAZ using a panel fixed effects model. We work at the village level considering that most villages in our sample are observed across all survey rounds and sampled students within each village are expected to be comparable over time, which is a key assumption for the estimation of this type of methods. We thus calculate within-village averages for the variables described in equation (1). This results in a panel of 448 village-year observations using the full sample and 238 and 213 village-year observations when restricting the analysis to pre-matched areas.

Table 5 shows the results of this alternative estimation approach that indicate a positive effect of the program on student health. When using the full sample of villages in column (1), the estimated effect of NIP on HAZ is 0.34 SDs, while when limiting the comparison to matched villages in columns (2) and (3) the estimated effects are 0.35-0.38 SDs. The precision of the estimates is naturally affected by the reduced number of observations used in the estimations, especially in the last column. Overall, these additional estimations support our main findings.

[36] Note that the random effects model does a much better job accounting for between than within variance across students, which could be explained by the reduced HAZ variation among individuals observed in two waves, as opposed to the HAZ variation across individuals. The overall model fit is worse off than the repeated cross-section base model with village fixed effects.

4.4 Heterogeneous effects

We examine whether the impact of the program on student HAZ varies by individual and household characteristics, including gender, age, if single child, parents' education, and mother's age at childbirth. We accordingly augment the DID model depicted in equation (1) to,

$$\begin{aligned} z_{ijt} &= \beta_0 + \beta_{11}NIP_j + \beta_{12}NIP_j * Hetero_{ijt} + \beta_{21}t_{NIP} + \beta_{22}t_{NIP} * Hetero_{ijt} \\ &+ \beta_{31}NIP_j * t_{NIP} + \beta_{32}NIP_j * t_{NIP} * Hetero_{ijt} + \gamma_1 X_{ijt} + \gamma_2 V_{jt} + u_{ijt} \end{aligned} \tag{3}$$

where $Hetero_{ijt}$ is the indicator variable that identifies students with a specific individual or household characteristic, which is also part of the vector of covariates X_{ijt}. We estimate separate models identifying: female students; students over 10 years old (median age); single child students; students whose father completed over 9 years of schooling (median years of education for father); students whose mother completed over 8 years of schooling (median years of education for mother); and students whose mother was over 24 years old when they were born (median age of mother at childbirth).

Table 6 reports the estimated parameter of interest β_{32} (known as the triple differences or DDD estimator) that captures the potential differential effects of NIP across each the six characteristics evaluated in columns (1) through (6). Panel A presents the results for the full sample and Panels B and C for the matched village samples. We do not find conclusive evidence of varying program effects by different individual or household characteristics. While NIP appears to have a larger impact on girls than boys (0.42-0.47 SDs higher effect), which is line with some studies in other countries showing a greater impact of feeding programs on girls (e.g., Gelli et al., 2019), the differentiated effect is only marginal significant in Panel A. Similarly, the program seems to have a lower impact on students with older mothers (0.34-0.46 SDs lower effect), which could imply more difficulties to change food consumption behaviors among older parents to adopt healthier food choices at home, but the differentiated effect is merely significant in Panel B. For all remaining characteristics, the estimated parameters are not statistically significant.[37]

4.5 Effects on other anthropometric measures

We finally examine the effect of the program on other outcomes of interest. These include an indicator variable identifying if the student is stunted (HAZ < -2), BMI-for-age z-score (BMIz), and weight-for-age z-score (WAZ). As opposed to HAZ, BMIz and WAZ are not necessarily

[37] We also do no find heterogeneous effects between students in early childhood (6-8 years old), middle childhood (8-11 years old), and early adolescence (12-16 years old).

indicative of a prolonged condition but are also standard anthropometric measures used to approximate the health and nutritional status of children in school age (up to 10 years old in the case of WAZ). We estimate a similar DID model as in equation (1) where we replace the dependent HAZ variable with the measures above.[38]

Table 7 presents the estimated DID parameters that approximate the average effect of NIP on stunting (Panel A), BMIz (Panel B), and WAZ (Panel C) using the full sample in column (1) and the matched village samples in columns (2) and (3). We do not find statistically significant effects on the three outcomes suggesting that NIP participation is not associated with a lower probability of being stunted neither with a higher BMI or weight.

The results for stunting are correlated with the CIC estimates depicted in Figure 2 that reveal no impacts among students on the lower end of the HAZ distribution. As discussed earlier, school aged children may be too old to recover from a chronic condition, at least through school meals and nutrition education activities. Hence, the positive average effects of NIP-related interventions on student HAZ do not translate into a lower stunting rate.

Likewise, we find positive impacts on BMIz and WAZ, but the effects are not statistically different from zero. These results are similar to Wang et al. (2019) that do not find NIP impacts on WAZ among seventh grade students, as well as to Nkhoma et al. (2013) that do not find effects of the Malawian SFP on weight (and height) of children aged 6-8 and Gelli et al. (2019) that do not find effects of SFPs in Ghana on BMIz of children aged 5-15. Buttenheim et al. (2011) find positive effects of take-home rations in Lao PDR on WAZ of children aged 3-10 (0.22 SDs) and marginal effects of combining on-site feeding and take-home rations (0.11 SDs), but no effects when only providing on-site feeding. In contrast, Singh et al. (2014) find large positive effects of the SFP in India on WAZ of children aged around 5.5 (0.6 SDs), mainly driven by drought-affected children.

Figure A.1 in the Appendix further reveals that there are no impacts across the entire BMIz (Panel A) and WAZ (Panel B) distributions, based on the depicted percentile treatments effects from equivalent CIC model estimations on these two z-scores. As opposed to height, the lack of NIP effects on weight are generalized. A possible explanation for these differences could be that meals (and recommended foods) in the national pilot favored or paid special attention to promoting the continuum of child growth given larger concerns in rural areas on low height as opposed to obesity or underweight.[39] We do not have though more detailed data

[38] For all three outcomes we follow a linear least squares approach. For stunting, the results of the linear probability model are not sensitive to alternatively implementing a discrete choice model.

[39] Childhood obesity and overweight, for instance, have been more of a concern for SFPs in more developed settings and high-income countries (WFP, 2013). See, e.g., Bhattacharya et al. (2006), Schanzenbach (2009), Millimet et al. (2010) and Gundersen et al. (2012) that examine the association between participation in SFPs and student obesity in the US and find mixed results.

to unravel the exact mechanisms by which the program could be differentially affecting students' height and weight and formally test this hypothesis; for example, the specific meals received and quality of foods and ingredients used.

5. Conclusion

The Chinese government has implemented several nutrition assistance programs over the past decades oriented for poor children that are mainly located in remote rural areas. As part of these efforts, the Nutrition Improvement Program (NIP) was launched in late 2011 targeting compulsory education students in many deprived rural areas. In this paper, we examine whether participation in the national pilot of this large-scale school meal program increased the health condition and nutrition of enrolled students.

The estimation results indicate that NIP participation is, on average, associated with a positive increase in the height-for-age z-score (HAZ) of about 0.22-0.42 standard deviations. The effects are larger among students in a better health condition but small or not significant among most disadvantaged students, which do not translate into a lower stunting rate. The estimated effects do not appear to vary by different individual and household characteristics such as gender, age, if single child, parents' education, and age of mother. In addition, we do not find meaningful effects on other relevant anthropometric measures, including the BMI-for-age and weight-for-age z-scores (BMIz and WAZ). The results are robust to alternative samples and estimation approaches.

Our findings suggest that NIP has played a non-negligible role in improving rural students' health, at least on HAZ over the first years of implementation, but additional support is needed to achieve broader impacts that effectively reach all vulnerable students. The lack of effects on the most deprived and likely stunted students points out that attending malnutrition at childhood and early adolescence is complex. Providing school meals combined with school-based nutrition education seems to improve the health and nutrition condition of children that are relatively better off at school age but is not enough to treat students that suffer from malnutrition. Child malnutrition in rural areas requires special attention with more tailored and comprehensive interventions at very early stages of life (prior to school enrollment) that could then be complemented with school-feeding and nutrition programs. In this line, NIP could achieve larger impacts by building on more targeted (and intensive) maternal, infant, and preschool interventions, including the One Yuan Nutrition Package program for children aged 6 to 24 months in poor areas and subsequent food supplement and nutrition programs for ages 3-5.

Since we rely on a quasi-experimental approach, we acknowledge that we cannot fully discard potential unobservable differences between treatment and comparison areas not accounted for that could still be influencing our results, although we are not aware of other major interventions that could have differentially affected the anthropometric outcomes of students during the period of analysis. Similarly, we focus on the effects of the initial national pilot, but future work should assess extended effects as NIP continues to operate and expands and more data become available, including disentangling the specific channels through which the program is contributing to the observed outcomes. Exploring potential impacts on intra-household food allocation and whether school meals supplant one or more meals at home is also an avenue of future research. Lastly, besides direct impacts on students' health and nutritional status, NIP could be contributing to educational goals and enhancing local social safety nets that should eventually be examined to assess more comprehensive program impacts.

References

Abadie, A., 2005. Semiparametric difference-in-differences estimators. *Review of Economic Studies* 72(1), 1-19.

Adelman, S., Gilligan, D.O., Konde-Lule, J., Alderman, H., 2019. School feeding reduces anemia prevalence in adolescent girls and other vulnerable household members in a cluster randomized controlled trial in Uganda. *The Journal of Nutrition* 149(4), 659-666.

Alderman, H., Behrman, J.R., Glewwe, P., Fernald, L., Walker, S., 2017. Evidence of impact of interventions on growth and development during early and middle childhood. In: Bundy, D.A.P, Silva, N.D., Horton, S., Jamison, D.T., Patton, G.C. (Editors). Child and Adolescent Health and Development. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; Chapter 7.

Alderman, H., Bundy, D., 2012. School feeding programs and development: Are we framing the question correctly? *World Bank Research Observer* 27(2), 204-221.

Afridi, F., 2010. Child welfare programs and child nutrition: Evidence from a mandated school meal program in India. *Journal of Development Economics* 92, 152-165.

Ao, D., Wu, F., Yun, C.-F., Zheng, X.-Y., 2019. Trends in physical fitness among 12-year-old children in urban and rural areas during the social transformation period in China. Journal of Adolescent Health 64, 250-257.

Ashenfelter, O., Card, D., 1985. Using the longitudinal structure of earnings to estimate the effects of training programs. *Review of Economics and Statistics* 67(4), 648-660.

Athey, S., Imbens, G., 2006. Identification and inference in nonlinear difference-in-differences model. *Econometrica* 74(2), 431-497.

Berry, J., Mehta, S., Mukherjee, P., Ruebeck, H., Shastry, G.K., 2020. Implementation and effects of India's national school-based iron supplementation program. *Journal of Development Economics* 144, 102428.

Best, C., Neufingerl, N., van Geel, L., van den Briel, T., Osendarp, S., 2010. The nutritional status of school-aged children: Why should we care? *Food and Nutrition Bulletin* 31(3), 400-417.

Bhattacharya, J., Currie, J., and Haider, S., 2006. Breakfast of Champions? The Nutritional Effects of the School Breakfast Program. *Journal of Human Resources* 41(3), 445-466.

Blundell, R., Costa Dias, M., 2000. Evaluation methods for non-experimental data. *Fiscal Studies* 21(4): 427-468.

Blundell R., Costa Dias, M., Meghir, C., Van Reenen, J., 2004. Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association* 2(4): 569-606.

Brown, P.H., Park, A., 2002. Education and poverty in rural China. *Economics of Education Review* 21(6), 523-541.

Bundy, D.A.P, Silva, N.D., Horton, S., Patton, G.C., Schultz, L., Jamison, D.T., 2017. Child and adolescent health and development: Realizing neglected potential. In: Bundy, D.A.P, Silva, N.D., Horton, S., Jamison, D.T., Patton, G.C. (Editors). Child and Adolescent Health and Development. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; Chapter 1.

Bundy, D.A.P., Burbano, C., Grosh, M., Gelli, A., Jukes, M., Drake, L., 2009. Rethinking School Feeding. Social Safety Nets, Child Development and the Educational Sector. Washington, DC: World Bank.

Buttenheim, A., Alderman, H., Friedman, J., 2011. Impact evaluation of school feeding programs in Lao PDR. Policiy Research Working Paper 5518, The World Bank.

Chen, S., Liu, W., Song, H., 2020. Broadband internet, firm performance, and worker welfare: Evidence and mechanism. *Economic Inquiry* 58(3), 1146-1166.

China Development Research Foundation (CDRF), 2017. Progress of nutrition improvement plan for students in poverty-stricken rural areas. Accessed January 2021. http://tsf.cdrf.org.cn/Content/Detail/31/39/926

Chyi, H., Zhou, B., 2014. The effects of tuition reforms on school enrollment in rural China. *Economics of Education Review* 38, 104-123.

Connelly, R., Zheng, Z., 2003. Determinants of school enrollment and completion of 10 to 18 year olds in China? *Economics of Education Review* 22(4), 379-388.

de Onis, M., Branca, F., 2016. Childhood stunting: a global perspective. *Maternal & Child Nutrition* 12, 12-26.

de Onis, M., Onyango, A.W., Borghi, E., Siyam, A., Nishida, C., Siekmann, J., 2007. Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization* 85(9): 660-667.

Deng, Z.J., Mao, G.X., Wang, Y.J., Liu, L., Chen, Y., 2016. Evaluation of nutritional status of school-age children after implementation of "Nutrition Improvement Program" in rural area in Hunan, China. *Journal of Contemporary Pediatrics* 18, 851-856. [In Chinese]

Drake, L., Fernandes, M., Aurino, E., Kiamba, J., Giyose, B., Burbano, C., Alderman, H., Mai, L., Mitchell, A., Gelli, A., 2017. School feeding programs in middle childhood and adolescence. In: Bundy, D.A.P, Silva, N.D., Horton, S., Jamison, D.T., Patton, G.C. (Editors). Child and Adolescent Health and Development. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; Chapter 12.

Fraumeni, B.M., He, J., Li, H., Liu, Q., 2019. Regional distribution and dynamics of human capital in China 1985-2014. *Journal of Comparative Economcis* 47, 853-866.

Gelli, A., Aurino, E., Folson, G., Arhinful, D., Adamba, C., Osei-Akoto, I., Masset, E., Watkins, K., Fernandes, M., Drake, L., Alderman, H., 2019. A school meals program implemented at scale in Ghana increases height-for-age during midchildhood in girls and in children from poor households: A cluster randomized trial. *The Journal of Nutrition* 149, 1434-1442.

Gelli, A., Daryanani, R., 2013. Are school feeding programs in low-income settings sustainable? Insights on the costs of school feeding compared with investments in primary education. Food nutrition bulletin 34, 310-317.

Glewwe, P., Miguel, E.A., 2008. The impact of child health and nutrition on education in less developed countries. *Handbook of Development Economics*, Volume 4. Chapter 56. Elsevier.

Grillenberger, M., Neumann, C.G., Murphy, S.P., Bwibo, N.O., Van't Veer, P., Hautvast, J.G., West, C.E., 2003. Food supplements have a positive impact on weight gain and the addition of animal source foods increases lean body mass of Kenyan schoolchildren. *The Journal of Nutrition* 133(11 Suppl 2), 3957S-3964S.

Gundersen, C., Kreider, B., Pepper, J., 2012. The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics* 166(1), 79-91.

Ho, D.E., Imai, K., King, G., Stuart, E.A., 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15(3): 199-236

International Food Policy Research Institute (IFPRI), 2016. Global Nutrition Report 2016: From Promise to Impact: Ending Malnutrition by 2030. IFPRI books, International Food Policy Research Institute, Number 978-0-89629-584-1, September. Washington DC.

Jomaa, L.H., McDonnell, E., Probart, C., 2011. School feeding programs in developing countries: Impacts on children's health and educational outcomes. *Nutrition Reviews* 69(2), 83-98.

Kazianga, H., de Walque, D., Alderman, H., 2014. School feeding programs, intrahousehold allocation and the nutrition of siblings: Evidence from a randomized trial in rural Burkina Faso. *Journal of Development Economics* 106, 15-34.

Kerr, A., 2019. Household investment in durable appliances and outcomes for children: Evidence from China. *Labour Economics* 58, 110-127.

Kong, N., Osberg, L., Zhou, W., 2019. The shattered “Iron Rice Bowl”: Intergenerational effects of Chinese State-Owned Enterprise reform. *Journal of Health Economics* 67, 10220.

Krishnaratne, S., White, H., Carpenter, E., 2013. Quality education for all children? What works in education in developing countries?” 3ie Working Paper 20, International Initiative for Impact Evaluation, New Delhi.

Kristjansson, E., Gelli, A., Welch, V., Greenhalgh, T., Liberato, S., Francis, D., Espejo, F., 2016. Costs, and cost-outcome of school feeding programmes and feeding programmes for young children. Evidence and recommendations. *International Journal of Educational Development* 48, 79-83.

Kristjansson, E.A., Robinson, V., Petticrew, M., MacDonald, B., Krasevec, J., Janzen, L., Greenhalgh, T., Wells, G., MacGowan, J., Farmer, A., Shea, B.J., Mayhew, A., Tugwell, P., 2007. School feeding for improving the physical and psychosocial health of disadvantaged elementary school children. *Cochrane Database of Systemic Reviews* 7(1): CD004676.

Leroy, J.L., Ruel, M., Habicht, J.P., 2013. Critical windows for nutritional interventions against stunting. *American Journal of Clinical Nutritrion* 98(3), 854-855.

Li, Y.P., Hu, X.Q., Zhao, J., Yang, Z.G., Ma, G.S., 2009. Application of the WHO growth reference (2007) to assess the nutritional status of children in China. *Biomedical and Environmental Sciences* 22,130-135.

Liang, C., Wang, S., 2020. The influence of public school allocation on human capital: Evidence from the school consolidation in China. *Economic Research Journal* 55(09), 138-154. [In Chinese]

Millimet, D.L., Tchernis, R., Husain, R., 2010. School nutrition programs and the incidence of childhood obesity. *Journal of Human Resources* 45(3), 640-654.

Millman, A., Tang, D., Perera, F.P., 2008. Air pollution threatens the health of children in China. *Pediatrics* 122(3), 620-628.

Neumann, C.G., Murphy, S.P., Gewa, C., Grillenberger, M., Bwibo, N.O., 2007. Meat supplementation improves growth, cognitive, and behavioral outcomes in Kenyan children. *The Journal of Nutrition* 137(4), 1119-1123.

Nkhoma, O.W., Duffy, M.E., Cory-Slechta, D.A., Davidson, P.W., McSorley, E.M., Strain, J., O'Brien, G.M., 2013. Early-stage primary school children attending a school in the Malawian School Feeding Program (SFP) have better reversal learning and lean muscle mass growth than those attending a non-SFP school. *The Journal of Nutrition* 143(8), 1324-1330.

Popkin, B.M., Du, S., Zhai, F., Zhang, B., 2009. Cohort profile: The China health and nutrition survey - monitoring and understanding socio-economic and health change in China, 1989-2011. *International Journal of Epidemiology* 39(6), 1435-1440.

Prentice, A.M., Ward, K.A., Goldberg, G.R., Jarjou, L.M., Moore, S.E., Fulford, A.J., Prentice, A., 2013. Critical windows for nutritional interventions against stunting. *American Journal of Clinical Nutritrion* 97(5), 911-918.

Qi, L., Zhao, J., 2012. Nutrition intervention and human capital development of boarding students in poor areas. *Management World* (02), 52-61+72. [In Chinese]

Ren, W., Rammohan, A., Wu, Y., 2014. Is there a gender gap in child nutritional outcomes in rural China? *China Economic Review* 31, 145-155.

Schanzenbach, D.W., 2009. Do school lunches contribute to childhood obesity? *Journal of Human Resources* 44(3), 684-709.

Singh, A., Park, A., Dercon, S., 2014. School meals as a safety net: An evaluation of the Midday Meal Scheme in India. *Economic Development and Cultural Change* 62, 275-306.

SUN Movement (2010) Scaling up nutrition: a framework for action. *Food and Nutrition Bulletin* 31(1), 178-186.

The Lancet (2008). Maternal and Child Undernutrition, Special Series. January, 2008. Accessed January 2021. http://www.thelancet.com/series/maternal-and-child-undernutrition

The World Bank, 2018. The State of Social Safety Nets 2018. Washington D.C., World Bank. Accessed January 2021. https://openknowledge.worldbank.org/handle/10986/29115

The World Bank, 2009. Poor Regions to Poor People: The Evolution of China's Poverty Alleviation Agenda. Poverty and Inequality Assessment in China. [In Chinese]

van Stuijvenberg, M.E., 2005. Using the school feeding system as a vehicle for micronutrient fortification: Experience from South Africa. *Food Nutrition Bulletin* 26 (2 Suppl 2), S213-S219.

Victora, C., Adair, L., Fall, C., Hallal, P.C., Martorell, R., Richter, L., Sachdev, H.S., 2008. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet* 371(9609), 340-357.

Wang, D., Fawzi, W.W., 2020. Impacts of school feeding on educational and health outcomes of school-age children and adolescents in low- and middle-income countries: protocol for a systematic review and meta-analysis. *Systematic Reviews* 9:55.

Wang, J., Zhou, L., Yao, S., 2019. The impact of the Nutrition Improvement Program on children's health in rural areas: Evidence from China. *Emerging Markets Finance Trade*, DOI: 10.1080/1540496X.2019.1706047.

Watkins, K., Gelli, A., Hamdani, S., Masset, E., Mersch, C., Nadazdin, N., Vanhees, J., 2015. Sensitive to nutrition? A literature review of school feeding effects in the child development lifecycle. HGSF Working Paper Series No. 16. Home Grown School Feeding.

World Food Programme, 2019. School Feeding in 2018. Beyond the Annual Performance Report 2018 Series. October. Rome: WFP. Accessed January 2021. https://docs.wfp.org/api/documents/WFP-0000110344/download/

World Food Programme, 2017. China Country Strategic Plan (2017-2021). Accessed January 2021. https://www.wfp.org/operations/cn01-china-country-strategic-plan-2017-2021

World Food Programme, 2013. State of School Feeding Worldwide. Rome: WFP.

World Food Programme, 2009. The World Food Programme in China 1979-2009. Accessed January 2021. [In Chinese] https://zh.wfp.org/publications/lianheguoshijieliangshijihuashuzaizhongguo1979-2009huzhuhezuosanshinian

Xu, Y., Hang, L., 2017. Height inequalities and their change trends in China during 1985-2010: results from 6 cross-sectional surveys on children and adolescents aged 7-18 years. *BMC Public Health* 17(1): 473.

Yang, G., Bansak, C., 2020. Does wealth matter? An assessment of China's rural-urban migration on the education of left-behind children. *China Economic Review* 59, 101365.

Zhan, J., Yang, H., Jiang, X., Lu, H., Guo, X., Hong, Y., 2019. Analysis of students' nutritional status of the "Nutrition Improvement Program for Rural Compulsory Education Students" Area. *Advances in Social Science, Education and Humanities Research* 371, 362-367, 2nd International Workshop on Education Reform and Social Sciences (ERSS 2019), Atlantis Press.

Zhang, F., Hu, X., Tian, Z., Zhang, Q., Ma, G., 2014. Literature research of the Nutrition Improvement Programme for Rural Compulsory Education Students in China. *Public Health Nutrition* 18(5), 936-943.

Zhang, S., Ji, C., Wang, H., 2020. Determinants of the actual inpatient reimbursement rate under the New Rural Cooperative Medical Scheme. *Management Review* 32(10), 22-33. [In Chinese]

Zhaoyu, T. 2011. Free Lunch: Let your child have a hot lunch! China Social Organization 2011 (2): 10. [In Chinese]

Zong, X.N., Li, H., 2014. Physical growth of children and adolescents in China over the past 35 years. *Bulleting of the World Health Organization* 92(8), 555-564.

Table 1. Summary statistics

Variable	2004	2006	2009	2011	2015	Total
Height-for-age z-score (HAZ)	-0.438	-0.330	-0.137	-0.075	0.265	-0.165
	(1.109)	(1.282)	(1.168)	(1.209)	(1.203)	(1.215)
If stunted (HAZ < -2)	0.082	0.089	0.049	0.057	0.017	0.061
	(0.275)	(0.285)	(0.216)	(0.233)	(0.128)	(0.239)
Body Mass Index BMI-for-age z-score (BMIz)	-0.330	-0.391	-0.319	-0.246	-0.061	-0.278
	(1.108)	(1.196)	(1.333)	(1.455)	(1.713)	(1.358)
Weight-for-age z-score (WAZ)	-0.381	-0.377	-0.139	-0.035	0.203	-0.147
	(1.163)	(1.246)	(1.271)	(1.332)	(1.607)	(1.350)
If female	0.479	0.466	0.453	0.500	0.455	0.471
	(0.500)	(0.499)	(0.498)	(0.500)	(0.498)	(0.499)
Age	11.134	10.396	10.484	10.331	9.683	10.454
	(2.946)	(2.808)	(2.705)	(2.861)	(2.691)	(2.850)
If single child	0.777	0.834	0.864	0.885	0.835	0.835
	(0.417)	(0.372)	(0.343)	(0.319)	(0.372)	(0.371)
Years of schooling father	8.474	8.294	8.127	8.183	8.778	8.375
	(2.297)	(2.403)	(2.311)	(2.217)	(2.254)	(2.309)
Years of schooling mother	6.604	6.911	7.131	7.416	7.984	7.163
	(2.983)	(2.773)	(2.619)	(2.615)	(2.372)	(2.741)
Age of mother at birth of child	24.409	24.499	24.500	24.014	23.467	24.203
	(6.722)	(5.557)	(5.624)	(5.381)	(5.771)	(5.897)
Household size (members)	4.519	4.834	5.003	4.891	5.494	4.920
	(1.296)	(1.505)	(1.596)	(1.461)	(1.928)	(1.586)
Per capita household income (2015 Thousand Yuan)	4.821	5.566	8.022	9.304	12.628	7.810
	(3.937)	(9.351)	(9.598)	(9.262)	(19.615)	(11.501)
Economic environment score (scale 0-10)	4.141	5.051	5.046	5.648	5.437	5.001
	(2.416)	(2.409)	(2.553)	(2.615)	(2.822)	(2.610)
Health quality score (scale 0-10)	3.823	3.952	5.017	5.360	4.694	4.512
	(1.920)	(2.017)	(2.286)	(2.284)	(2.378)	(2.247)
Social services score (scale 0-10)	1.603	2.231	2.391	2.665	3.590	2.432
	(1.281)	(1.937)	(2.132)	(2.045)	(2.896)	(2.176)
If NIP beneficiary	0.336	0.384	0.343	0.351	0.312	0.345
	(0.473)	(0.487)	(0.475)	(0.478)	(0.464)	(0.475)
# observations	730	586	572	522	539	2,949

Note: The first row for each variable reports the mean and the second row the standard deviation in parentheses. BMIz is based on a total of 2,659 students for which their z-score is available while WAZ is based on 1,392 students between 6 and 10 years old for which their z-score is available. The scores (on a scale 0-10) for economic environment, quality of health services, and social services are calculated at the village level by the China Health and Nutrition Survey (CHNS) team and reported in the dataset.

Table 2. Differences-in-differences estimations on height-for-age z-score (HAZ)

Coefficient	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample		Matched villages			
			No replacement		With replacement	
			Dependent variable: HAZ			
If NIP beneficiary x If treatment period	0.222**	0.247**	0.347***	0.401***	0.364***	0.423***
	(0.105)	(0.104)	(0.116)	(0.111)	(0.128)	(0.123)
If female		-0.191***		-0.157***		-0.172***
		(0.041)		(0.047)		(0.050)
Age		-0.054***		-0.062***		-0.059***
		(0.007)		(0.009)		(0.009)
If single child		0.144***		0.123**		0.120**
		(0.052)		(0.056)		(0.057)
Years of schooling father		0.022**		0.034***		0.029**
		(0.010)		(0.013)		(0.013)
Years of schooling mother		0.028***		0.024**		0.022*
		(0.010)		(0.011)		(0.012)
Age of mother at birth of child		-0.002		0.003		0.003
		(0.003)		(0.004)		(0.004)
Household size (members)		0.004		0.007		0.011
		(0.017)		(0.017)		(0.019)
Per capita household income (Thousand Yuan)		0.001		0.002		0.001
		(0.002)		(0.003)		(0.004)
Economic environment score (scale 0-10)		-0.003		-0.024**		-0.024**
		(0.009)		(0.010)		(0.010)
Health quality score (scale 0-10)		0.010		0.011		0.014
		(0.010)		(0.012)		(0.013)
Social services score (scale 0-10)		0.004		-0.005		-0.006
		(0.010)		(0.011)		(0.014)
# observations	2,949	2,949	2,016	2,016	1,818	1,818
R-squared	0.231	0.261	0.214	0.254	0.224	0.260
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village and year. The results reported in columns (1)-(2) are based on the full sample, and the results in columns (3)-(4) and (5)-(6) are based on the matched samples from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement.

Table 3. Differences-in-differences estimations on height-for-age z-score (HAZ), alternative comparison groups

Coefficient	(1) Comparison group: counties matched with treatment counties	(2) Comparison group: counties in the same province	(3) Comparison group: subsequent local pilot counties
		Dependent variable: HAZ	
If NIP beneficiary x If treatment period	0.445***	0.362***	0.350**
	(0.121)	(0.108)	(0.144)
# observations	1,831	1,851	1,530
R-squared	0.272	0.241	0.258
Control variables	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village and year. The results reported in column (1) use as the comparison group rural students of non-NIP counties that were previously matched with the treatment counties using Mahalanobis distance matching, the results in column (2) use as the comparison group rural students of non-NIP counties in the same province as the NIP counties, and the results in column (3) use as the comparison group rural students of counties that were part of NIP local pilots after 2015. The control variables are the same as the ones described in equation (1).

Table 4. Differences-in-differences estimations on height-for-age z-score (HAZ), alternative panel approach

Coefficient	(1)	(2)	(3)
	Full sample	Matched villages	
		No replacement	With replacement
	Dependent variable: HAZ		
	Panel A: Clustering by individual		
If NIP beneficiary x If treatment period	0.247**	0.401***	0.423***
	(0.122)	(0.130)	(0.139)
# observations	2,949	2,016	1,818
R-squared	0.261	0.254	0.260
	Panel B: GLS individual random effects model		
If NIP beneficiary x If treatment period	0.197**	0.370***	0.393***
	(0.097)	(0.109)	(0.123)
# observations	2,949	2,016	1,818
R-squared within	0.002	0.009	0.020
R-squared between	0.148	0.146	0.142
R-squared overall	0.140	0.133	0.128
Control variables	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses. In Panel A the model is estimated by pooled ordinary least squares where the standard errors are clustered by individual. In Panel B the model is estimated by generalized least squares (GLS) individual random effects. The results reported in column (1) are based on the full sample, and the results in columns (2) and (3) are based on the matched samples from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement. The control variables are the same as the ones described in equation (1). The model in Panel A includes village fixed effects.

Table 5. Differences-in-differences estimations on height-for-age z-score (HAZ), pseudo panel at the village level

Coefficient	(1)	(2)	(3)
	All villages	Matched villages	
		No replacement	With replacement
		Dependent variable: HAZ	
If NIP beneficiary x If treatment period	0.338**	0.381**	0.353*
	(0.170)	(0.180)	(0.209)
# observations	448	238	213
R-squared within	0.278	0.351	0.375
R-squared between	0.007	0.198	0.166
R-squared overall	0.155	0.277	0.282
Control variables	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village. The pseudo panel is constructed at the village level across the five survey waves. The results reported in column (1) are based on the full sample of villages, and the results in columns (2) and (3) are based on the matched samples of villages from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement. The control variables are the same as the ones described in equation (1) averaged at the village-year level.

Table 6. Differences-in-differences estimations on height-for-age z-score (HAZ), heterogenous effects

Coefficient	(1) Female student	(2) Student > 10 years old	(3) Single child student	(4) Father > 9 years of education	(5) Mother > 8 years of education	(6) Mother > 24 years old at childbirth
	Dependent variable: HAZ					
	Panel A: Full sample					
If NIP beneficiary x If treatment period	0.467*	-0.045	-0.352	0.225	0.180	-0.341
x Heterogeneous indicator	(0.240)	(0.241)	(0.338)	(0.280)	(0.234)	(0.232)
# observations	2,949	2,949	2,949	2,949	2,949	2,949
R-squared	0.262	0.255	0.262	0.261	0.260	0.264
	Panel B: Matched villages (no replacement)					
If NIP beneficiary x If treatment period	0.424	-0.113	-0.114	0.374	0.216	-0.459*
x Heterogeneous indicator	(0.260)	(0.257)	(0.348)	(0.322)	(0.251)	(0.264)
# observations	2,016	2,016	2,016	2,016	2,016	2,016
R-squared	0.256	0.246	0.255	0.255	0.255	0.257
	Panel C: Matched villages (with replacement)					
If NIP beneficiary x If treatment period	0.442	-0.082	-0.185	0.203	0.228	-0.381
x Heterogeneous indicator	(0.277)	(0.270)	(0.361)	(0.396)	(0.269)	(0.282)
# observations	1,818	1,818	1,818	1,818	1,818	1,818
R-squared	0.262	0.253	0.261	0.261	0.261	0.262
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

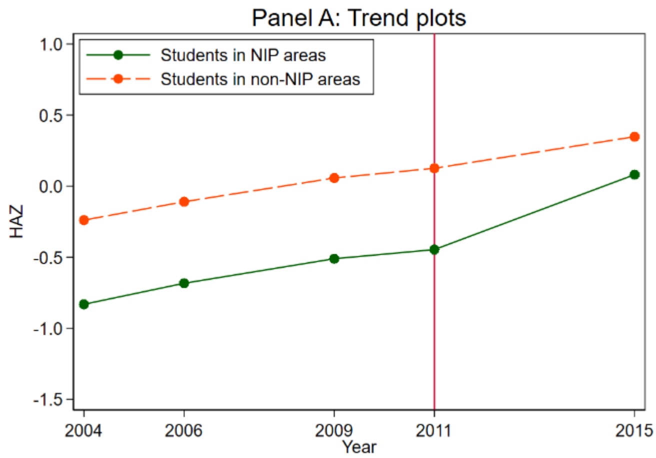
Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village and year. The results reported in Panel A are based on the full sample, and the results in Panels B and C are based on the matched samples from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement. The heterogeneous indicator is a dummy variable that varies across columns to identify: female students in column (1); students over 10 years old in column (2); single child students in column (3); students whose father completed over 9 years of schooling in column (4); students whose mother completed over 8 years of schooling in column (5); and students whose mother was over 24 years old when they were born in column (6). All regressions also include interaction terms of the heterogeneous indicator with the treatment dummy variable and the treatment period indicator. The control variables in each column are the corresponding heterogeneous indicator plus all remaining control variables described in equation (1).

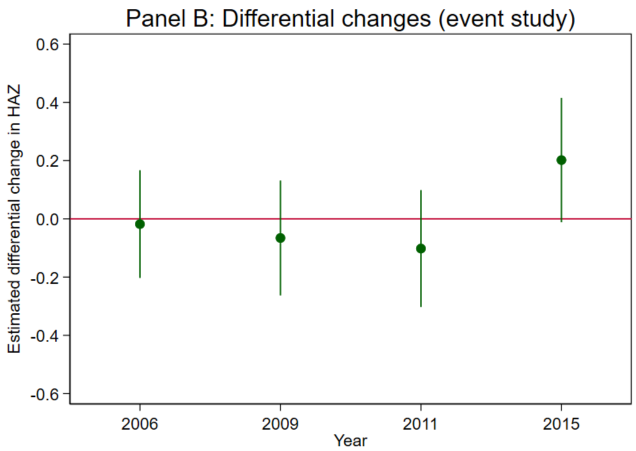
Table 7. Differences-in-differences estimations on stunting, Body Mass Index BMI-for-age z-score (BMIz), and weight-for-age z-score (WAZ)

Coefficient	(1)	(2)	(3)
	Full sample	Matched villages	
		No replacement	With replacement
	Panel A: Dependent variable: If stunted		
If NIP beneficiary x If treatment period	0.001	-0.008	-0.014
	(0.022)	(0.024)	(0.026)
# observations	2,949	2,016	1,818
R-squared	0.124	0.127	0.131
	Panel B: Dependent variable: BMIz		
If NIP beneficiary x If treatment period	0.108	0.201	0.161
	(0.158)	(0.173)	(0.181)
# observations	2,659	1,817	1,642
R-squared	0.187	0.139	0.148
	Panel C: Dependent variable: WAZ		
If NIP beneficiary x If treatment period	0.048	0.235	0.281
	(0.217)	(0.213)	(0.218)
# observations	1,392	959	866
R-squared	0.295	0.235	0.245
Control variables	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village and year. Panels A and B consider students aged 6-16, while Panel C considers students aged 6-10. The results reported in column (1) are based on the full sample, and the results in columns (2) and (3) are based on the matched samples from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement. The control variables are the same as the ones described in equation (1).

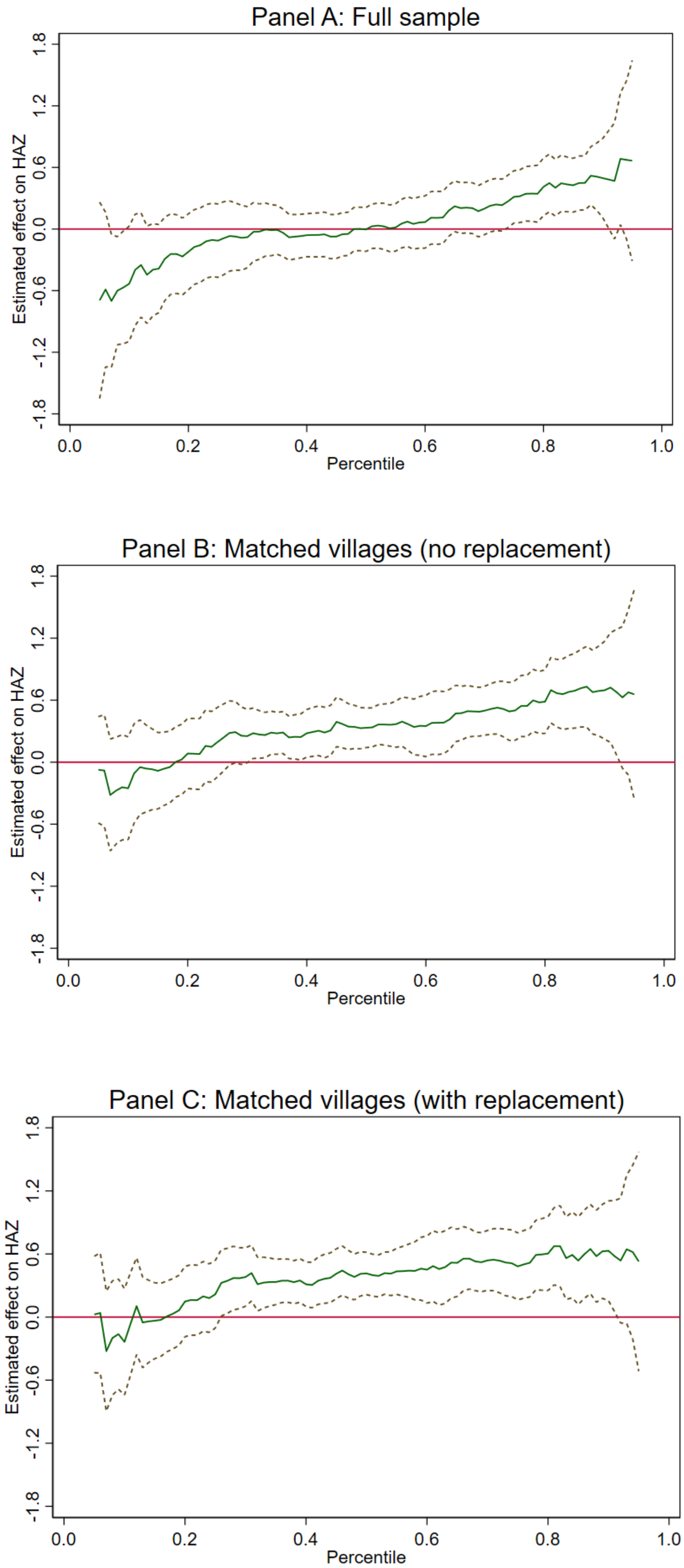
Figure 1. Trend plots and differential changes over time in height-for-age z-score (HAZ) between treatment and control cohorts





Note: Panel A reports the sample average HAZ for each cohort in each survey year. Panel B reports the estimated differential changes over time in HAZ between cohorts obtained by regressing HAZ on year indicators (base category is 2004), interactions of the year indicators with the treatment dummy variable (i.e., if the student is in a village where NIP was implemented), and the set of control variables described in equation (1), including village fixed effects. The point estimates are the corresponding coefficients of the interaction terms between the year indicators and treatment dummy variable, and the vertical lines are the 95% confidence intervals. Sample size = 2,949 observations.

Figure 2. Percentile treatment effects on height-for-age z-score (HAZ) based on changes-in-changes estimation



Note: Panel A is based on the full sample (2,949 observations), and Panels B and C are based on the matched sample from previously pairing treatment and control villages using Mahalanobis distance matching without and with replacement (2,016 and 1,818 observations, respectively). The dashed lines are 95% confidence intervals obtained from 2,000 bootstrap replications. The percentile in the horizontal axis is expressed in decimals.

Appendix. Supplementary Tables and Figures

Table A.1. Average differences between treatment and control villages, full and matched samples

Variable	All villages	Matched villages: No Replacement	Matched villages: With replacement
If female	0.073	0.058	0.051
	(0.028)	(0.028)	(0.030)
Age	0.171	0.158	0.207
	(0.183)	(0.182)	(0.195)
If single child	-0.077	-0.018	-0.022
	(0.015)	(0.019)	(0.020)
Years of schooling father	-0.296	0.146	0.164
	(0.171)	(0.179)	(0.192)
Years of schooling mother	-0.715	-0.559	-0.380
	(0.204)	(0.224)	(0.240)
Age of mother at birth of child	-0.957	-0.365	-0.258
	(0.389)	(0.418)	(0.442)
Household size (members)	0.051	-0.363	-0.421
	(0.108)	(0.121)	(0.128)
Per capita household income (2015 Thousand Yuan)	-0.404	0.426	0.434
	(0.915)	(0.977)	(0.932)
Economic environment score (scale 0-10)	0.907	1.038	1.056
	(0.278)	(0.317)	(0.326)
Health quality score (scale 0-10)	0.304	0.336	0.356
	(0.235)	(0.286)	(0.304)
Social services score (scale 0-10)	-0.020	-0.033	0.031
	(0.240)	(0.271)	(0.271)
# observations	448	238	213

Note: The first row for each variable reports the average differences between treatment and control villages, and the second row the standard errors of the differences in parentheses. Villages matched using one-to-one Mahalanobis distance matching without and with replacement.

Table A.2. Survey attrition and school enrollment between treatment and comparison areas across survey waves

Coefficient	(1) Dependent variable: If individual observed only one time	(2) Dependent variable: If individual enrolled into school
If 2006 wave	-0.223***	0.021
	(0.032)	(0.018)
If 2009 wave	-0.174***	0.007
	(0.035)	(0.020)
If 2011 wave	-0.113***	-0.013
	(0.040)	(0.020)
If 2015 wave	0.116***	-0.016
	(0.034)	(0.022)
If 2006 wave x If located in NIP county	0.032	-0.046
	(0.050)	-0.028
If 2009 wave x If located in NIP county	-0.047	-0.017
	(0.053)	(0.029)
If 2011 wave x If located in NIP county	-0.020	-0.027
	(0.060)	(0.032)
If 2015 wave x If located in NIP county	0.018	-0.058
	(0.047)	(0.036)
# observations	3,192	3,192
R-squared	0.119	0.070
Control variables	Yes	Yes
Village Fixed Effects	Yes	Yes

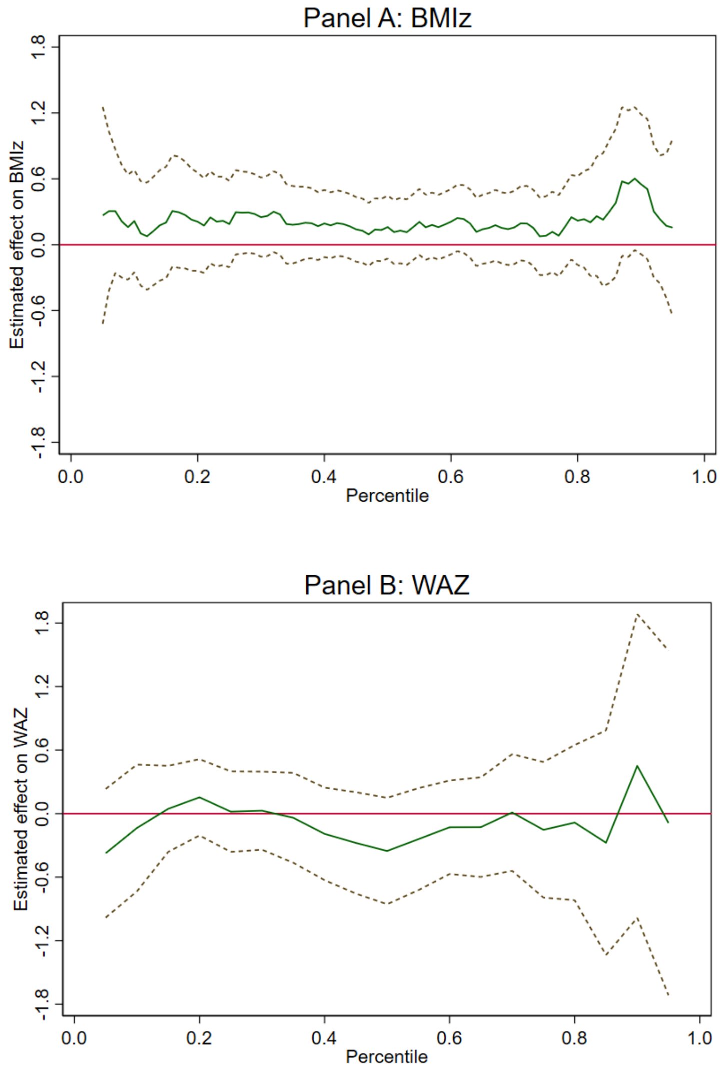
Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by village and year. The results reported in column (1) are based on all 6-16 years old children in rural areas across the 8 NIP counties and 24 non-NIP counties considered for the base analysis, where the dependent variable is an indicator variable equal to one for those children solely observed during one survey round, and zero otherwise. The results reported in column (2) are based on all 6-16 years old children in rural areas across the 8 NIP counties and 24 non-NIP counties, where the dependent variable is an indicator variable equal to one for those children enrolled into school, and zero otherwise. The control variables are the same as the ones described in equation (1).

Table A.3. Placebo tests

Coefficient	(1)	(2)	(3)	(4)	(5)
	Enrolled students in urban areas	Non-enrolled children in rural areas	Permutation tests		
			NIP started after 2004	NIP started after 2006	NIP started after 2009
			Dependent variable: HAZ		
If located in NIP county x If treatment period	-0.034	-0.164	0.001	0.014	0.071
	(0.212)	(0.422)	(0.082)	(0.077)	(0.084)
# observations	829	243	2,949	2,949	2,949
R-squared	0.230	0.498	0.260	0.260	0.260
Control variables	Yes	Yes	Yes	Yes	Yes
Location Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes

Note: ***, **, * Significant at 1%, 5%, and 10% level. Robust standard errors reported in parentheses clustered by location and year. The results reported in column (1) are based on 6-16 years old enrolled students in urban areas across the 8 NIP counties and 24 non-NIP counties considered for the base analysis. The results reported in column (2) are based on 6-16 years old children not enrolled into school, which are in rural areas across the 8 NIP counties and 24 non-NIP counties considered for the base analysis. The results reported in columns (3)-(5) are based on the full working sample assuming that NIP started after 2004, 2006 and 2009, respectively. The control variables are the same as the ones described in equation (1).

Figure A.1. Percentile treatment effects on Body Mass Index BMI-for-age z-score (BMIz) and weight-for-age z-score (WAZ) based on changes-in-changes estimation



Note: Panel A is based on a total of 2,659 observations corresponding to students for which their BMIz is available, and Panel B is based on 1,392 observations corresponding to students between 6 and 10 years old for which their WAZ is available. The dashed lines are 95% confidence intervals obtained from 2,000 bootstrap replications. The percentile in the horizontal axis is expressed in decimals.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW

Washington, DC 20005 USA

Tel.: +1-202-862-5600

Fax: +1-202-862-5606

Email: ifpri@cgiar.org