China & World Economy / 4–30, Vol. 30, No. 4, 2022

Education Universalization, Rural School Participation, and Population Density

*Xi Zhang, Scott Rozelle**

Abstract

In many developing countries, low population density may be a major reason for low school participation in rural areas, and the problem is likely to worsen with rapid urbanization. However, few studies have investigated empirically the role of population density in rural education, especially the moderating effect of population density on the outcomes of education policies. This study aims to fill this gap in the literature. From 1999 through the early 2000s, China launched a set of major nationwide policies aimed at universalizing 9-year compulsory education in rural areas. Using difference-in-differences and triple difference strategies, we show that the policies significantly increased the probability of junior high school enrollment of rural children and, more importantly, these policies were more effective in densely populated regions. These findings confirm the importance of population density to rural education.

Keywords: enrollment rate, policy effectiveness, rural population density, universal 9-year compulsory education policy
JEL codes: H41, J24, O18

I. Introduction

Today, many developing countries are facing the challenge of improving rural education while the process of urbanization is taking place. On the one hand, because of relatively low urbanization rates and human capital accumulation, the potential for growth in those countries largely depends on rural education for local children (Krueger and Lindahl, 2001; Fleisher et al., 2010; Che and Zhang, 2018; United Nations, 2019). On the other hand, due to factors such as new technologies and economic globalization, the speed of urbanization in developing countries is much greater than it was for developed countries in the past (Wan and Zhang, 2017). As the dominant suppliers of rural education,

*Xi Zhang (corresponding author), Post-doctoral Researcher, Institute of Population and Labor Economics, Chinese Academy of Social Sciences, China. Email: zhangxicass@126.com; Scott Rozelle, Helen F. Farnsworth Senior Fellow, Freeman Spogli Institute for International Studies, Stanford University, the US. Email: rozelle@stanford.edu.

governments are supposed to act appropriately in a rapidly changing environment. Accordingly, rural education policymaking ought to be forward looking and should develop long-range planning.

Why does rapid urbanization matter? One reason is related to rural population density. Although there are many possible causes of poor rural education, low population density in rural areas may be a contributing factor. As many education costs are fixed or quasi-fixed, and can be diluted by students, education involves strong economies of scale (Tholkes and Sederberg, 1990; Chakraborty et al., 2000). Local education faces a dilemma if the population is dispersed. A large central school can take advantage of economies of scale. However, extra commuting costs or boarding costs arise. Perhaps students can attend small schools nearby, but the consequences are either high average education costs or inferior quality schooling (Andrews et al., 2002; Bard et al., 2006; Ares Abalde, 2014). From the supply and demand sides, low population density makes rural education unwieldy. With rapid urbanization, rural population density decreases considerably, and this problem is gaining in importance.

Despite the potential significance of population density for rural education, there are few relevant empirical studies on this topic. There are several studies about economies of scale in education, impacts of commuting and boarding, and even teachers' willingness to serve in remote schools (Sargent and Hannum, 2005; Luschei, 2012; Li and Liu, 2014; Wei, 2016), but an overall evaluation of the causal effect of low population density on rural children's educational attainment is lacking. However, such evaluations are crucial for making projections and guiding actions. One reason for the research gap may involve endogeneity of population density. It is difficult to specify exogenous sources of variation in rural population densities and identify the causality between rural school participation and local population density. Nevertheless, another approach is to identify the effects of educational policies affecting various regions and examine heterogeneity of policy effectiveness among regions with different rural population density. This study exploits a dramatic change in policy to evaluate the role of population density in rural education in China.

Recognizing that 9-year compulsory education had not been universalized in practice, from 1999 through the early 2000s, China launched a set of major education policies aimed at increasing compulsory education enrollment nationwide. The universal 9-year compulsory education policies (UNCEP) consist of several specific supply-sided initiatives, including enhancing school facilities, promoting teacher quality, and providing student subsidies. As the enrollment gap mostly came from rural children, UNCEP measures were in no small part targeted at rural areas, particularly where the enrollment rate was relatively low.

The overall goal of this study is to investigate how rural population density is related to the outcomes of UNCEP measures. To achieve this goal, we have two specific objectives. The first one is to identify and measure the causal effect of UNCEP on rural junior high school participation. The second one is to analyze the heterogeneity of the effect on students living in rural areas characterized by lower and higher population densities.

To achieve our goals, we conducted a difference-in-differences (DID) strategy based on two waves of nationwide survey data from 2000 and 2005. In carrying out this approach, we established treatment and control groups and examined the policies' causal effect on the probability of rural junior high school enrollment over the 5 year period during which UNCEP was being implemented. By exploiting disparities in rural population density across regions in China, we studied the heterogeneity of the policies' outcomes via a triple difference (DDD) strategy.

Our results show that UNCEP was effective in improving rural junior high school enrollment. On average, from 2000 to 2005, the probability of junior high school enrollment of rural eligible children increased by 0.023 in comparison with that of urban children, and the same probability of rural children living in education-lagging regions increased by 0.064 in comparison with that of rural children from education-leading regions. As the parallel trend test shows, there was no existing convergence trend before 1999, and confounding factor tests indicate that the DID effect was not caused by other simultaneous events. Taken together, these results suggest that the policies were effective.

More important, we found that UNCEP was more effective in promoting junior high school participation in densely populated rural areas. After controlling for the initial rural enrollment rates and the fiscal education expenditure growth, proxies for the intensity of the policy, regions with higher population densities enjoyed higher policy-induced increases in rural enrollment probabilities. For every standard deviation of rural population density at the prefectural level, the DID effect varies by 0.019 and 0.041, respectively, under two treatment settings, namely rural children versus urban children and children living in education backward rural regions versus those in education advanced rural regions. The result is robust when estimation methods and control variables are changed.

This study provides evidence that population density indeed matters in rural education. In the short run, policymakers should deal with immediate problems associated with low population density, such as long commuting. Providing school buses and quality boarding facilities in low population density areas might be a helpful way to improve rural school enrollment. In the longer run, however, to fully exploit economies of scale in education and other public services, urbanization and amalgamation of villages should be embraced in future reforms (Jia and Zhong, 2022).

This paper adds to the literature in several ways. First, from the perspective of policy evaluation, studies on UNCEP are enriched and solidified. Estimates of the policies' effects are rare given the policies' significance. Existing studies, including Chyi and Zhou (2014), Shi (2016), and Wang (2018), all focus on the Two Exemptions One Subsidy policy, a part of UNCEP. Their findings to some extent contradict each other. One reason may be that their samples are narrow. Another reason may be that they all make DID estimations exploiting variations in Two Exemptions One Subsidy implementation time but they ignore the impacts of other measures of UNCEP. This paper complements those studies by using a huge national sample to improve representativeness and by viewing all relevant policies as a whole to avoid misidentification.

Second, a new explanation for the urban–rural education gap is provided. Many studies believe that the gap is due to insufficient public expenditure and unfavorable individual and family characteristics in rural areas (Brown and Park, 2002; Chen et al., 2010; Zhao and Glewwe, 2010). The influence of low population density has seldom been examined. Based on data from Netherland, van Maarseveen (2021) reports that one's educational attainment is positively associated with the population density of the environment where one lived as a child, depending on individual and family characteristics. The mechanisms for this remain unclear. The paper contributes to this strand of literature by studying on a developing country and offering the perspective of education supply effectiveness.

Third, in the sense of development economics, new insights are provided into urbanization and rural public services. In the process of urbanization, the welfare of rural residents may not increase that much although labor may become scarcer, because economies of scale related to public services are reduced. More collective action is needed to make rural public activities converge and rebuild economies of scale.

The remainder of this paper is organized as follows. Section II introduces UNCEP in China. Section III describes the data and outlines empirical strategy. Section IV reports and discusses results. Finally, Section V presents the conclusions.

II. Universal 9-year compulsory education policies in China

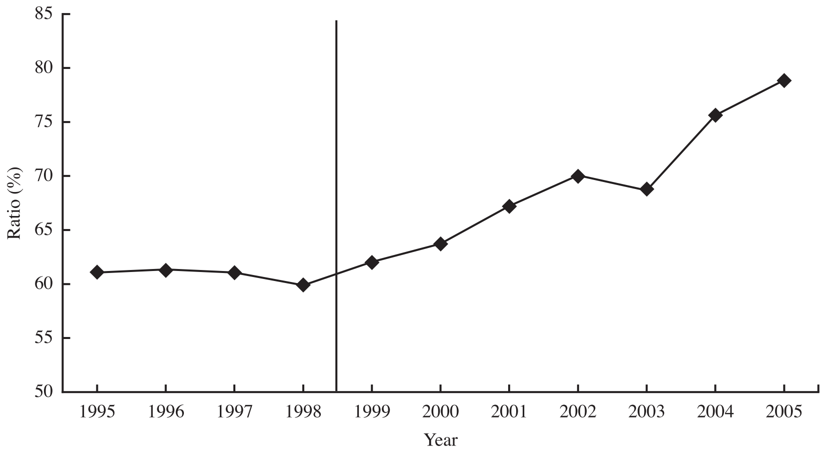
Experiences in many developing countries indicate that compulsory education is not necessarily universalized. In China compulsory education was implemented in 1986 when the *Compulsory Education Law* was enacted by the National People's Congress. From then on, children were required to receive a 9-year education, including 6 years of primary schooling and 3 years of junior high schooling. As a result of the implementation of compulsory education, average educational attainment in China improved significantly (Huang, 2015). However, in many cases, rural children of eligible ages did not attend school – junior high school in particular – because of the poor quality and high cost of compulsory education. According to the Ministry of Education of China (1999a), by 1998, 12 years after the *Compulsory Education Law* was passed, the national gross enrollment rate for junior high school was only 87.3 percent. It was mainly rural children who were absent from schools, as the estimated rural junior high school gross enrollment rate was less than 78 percent. This situation did not appear to be consistent with the intention that the nation should be vitalized by education, science, and technology.

Faced with the gap in compulsory education enrollment, the central government responded forcefully. A set of major policies aimed at elevating the enrollment rate was launched continuously from 1999. We call these policies, collectively, UNCEP. In January 1999, the *Action Plan of Vitalizing Education Towards the 21st Century* (Action Plan for short, Ministry of Education of China, 1999b) was released. This can be regarded as guiding principles to China's education development for the next decade. The priority of universalizing compulsory education was highlighted in the *Action Plan*. Hence, rural compulsory education was brought into sharp focus at all levels of government.

Following the *Action Plan*, the State Council issued the *Decision on Basic Education Reform and Development* in 2001. The decision's essence is supply sided. Generous governmental funding was demanded to ensure payments to rural teachers, to improve rural school buildings, to control school fees, and to subsidize poor students. The well-known Two Exemptions One Subsidy policy, namely tuition exemptions, free textbooks, and living expense subsidies, originated from this decision (Chyi and Zhou, 2014; Shi, 2016; Wang, 2018). The next document is the *Decision on Further Strengthening Rural Education* in 2003 (the State Council of China, 2003). The content is similar to that of the former decision but it placed more emphasis on vertical fiscal transfer payments to rural compulsory education. There was also a series of lower-level governmental documents issued in the early 2000s. These documents provided specific instructions to help expedite the national plan and decisions outlined above. In summary, from 1999 to the early 2000s, the central government mobilized local governments to universalize compulsory education in rural areas by assigning political tasks and offering transfer payments.

Owing to the strong will of the central government, the initiative of universalizing compulsory education in rural areas did not remain on paper. We calculated annual fiscal expenditure per student in the lower secondary education phase in rural and urban areas and then obtained the rural–urban ratio. Figure 1 shows that, before 1999, the ratio roughly held constant but surged thereafter. By 2005, fiscal expenditure per junior high school student in rural areas reached about 80 percent of that in urban areas. This conspicuous transition reflects the strength of the UNCEP.

Figure 1. Per-student budget expenditure on lower secondary education phase of rural student (% of urban student)



Sources: The figure is drawn by the authors based on the data from *Educational Statistics Yearbook of China* 1995−2005 by the Ministry of Education of China; *China Educational Finance Statistical Yearbook* 1996−2006 by the Ministry of Education of China.

III. Data and methodology

1. Data

To explore UNCEP's effect on rural junior high school participation, this study uses microdata from Chinese 2000 Census and 2005 1 Percent National Population Sample Survey. A national census takes place every decade in China. Household by household, every Chinese citizen's information is registered during the census, including the individual's age, gender, ethnicity, educational level, category and location of the household registration (*hukou*),[1] permanent residence, and other household information. The 1 Percent National Population Sample Survey serves as a "mini-Census" as it only covers a representative household sample account for about 1 percent of the total

[1]The *hukou* system is a Chinese household registration system – one of the legacies of China's centrally planned economy. *Hukou* represents a location at which one is supposed to reside. This was a complicated system; however, there are three key points that are relevant to this paper. First, there are two dimensions of *hukou*: it classifies households as "agricultural" and "non-agricultural," and it also classifies them by region. Second, *hukou* is especially important in education, as Chinese students can only attend several assigned public primary and junior high schools based on their *hukou* locations, take the senior high school entrance examinations (*zhongkao*) and have upper secondary education in the prefecture where their *hukou* are, and compete with others through national college entrance examinations (*gaokao*) within the province where their *hukou* are. Third, one's *hukou* is originally inherited from one's parents and can be moved from rural to urban areas or among regions in several ways including acquiring the appropriate qualifications.

population and the questionnaire is almost the same as that of a census. A 1 Percent National Population Sample Survey is conducted five years after a census. In this study we used two random subsamples[2] of the raw data: data from the 2000 Census with approximately 1.2 million observations and data from the 2005 1 Percent National Population Sample Survey with approximately 2.6 million observations. These two subsamples were released from the National Bureau of Statistics of China and are widely used in the existing labor economics literature about China.

We adopted three criteria to define the eligible sample so that we can evaluate UNCEP's impact on junior high school enrollment rates. First, we limited the observations to individuals aged 10–16 years. In China, children start primary school if they reach the age of 6 years at the beginning of the autumn semester (September 1). Primary schooling is typically completed by age 12, as it usually takes 6 years. Considering that some children may start primary schooling earlier or later, and may skip or repeat grades in primary school, we considered children between 10 and 16 years old to be junior-high aged. Second, we excluded students who were not eligible for junior high school – that is, the students who dropped out from primary school, who were still at primary school, and those who had graduated from junior high school. Third, we excluded samples from China's four municipalities, namely Beijing, Tianjin, Shanghai, and Chongqing. These four municipalities are substantially different from prefecture-level cities in terms of scale, administration, and development level. After applying these criteria, all the individuals remaining are representative of Chinese children who were supposed to attend junior high schools.

To explore how UNCEP's effectiveness changed among regions with different population densities, we also utilized some regional characteristics to generate regional variables. Location information in the microdata is at prefectural level, so regional variables must be at or above that level. The most important variable, prefectural rural population density, equals prefectural rural population divided by prefectural area. The prefectural rural population data are obtained from county-level data from the 2000 Census. The areas of prefectures are sourced from the *China City Statistical Yearbook* 2001 (NBS, 2002). We must also control for the growth of fiscal education expenditure per student to reveal only the influence of population density, because UNCEP's intensity may vary among regions. The fiscal education expenditure and the number of students in each prefecture come from the *National Financial Statistical Data of Cities and Counties* (Ministry of Finance of China, 2001, 2007) and *China City Statistical*

[2]The household is the minimum unit in random sampling. This rule means that we can match the parents' and siblings' information to an individual.

Yearbook (NBS, 2002, 2006), respectively. There are also some other regional control variables, which are derived from the data mentioned above or other open sources.

2. Empirical strategies

Our empirical efforts are divided into three parts. First, we performed a baseline DID regression to measure the relative change between the junior high school enrollment probability of the treatment group and that of the control group. Second, we conducted an indirect test of the parallel trend assumption to assert that the DID effect is causal. Third, after attributing the relative increase in the treatment group's enrollment probability to UNCEP, we analyzed UNCEP's differential effectiveness among regions with different population densities.

(1) Baseline model

Based on the pooled cross-sectional data described above, a DID strategy is applied to identify the causal effect of UNCEP on rural junior high school enrollment rates. The empirical specification is outlined as the following linear probability model:

$$enroll_{ipt} = \alpha treat_{ipt} + \beta year05_t + \delta treat_{ipt} \times year05_t + \pi_t \boldsymbol{X_{ipt}} + \gamma_p + \varepsilon_{ipt}, \quad (1)$$

where $enroll_{ipt}$ is a dummy variable indicating whether child i in prefecture p was enrolled in a junior high school at survey time t. The independent variable $treat_{ipt}$ is a dummy variable indicating whether the child belongs to the treatment group. There are two ways to define the treatment and control groups. In one the treatment group consists of eligible rural children and the control group consists of their urban counterparts. This setting is based on UNCEP's targeting of rural areas. The second only considers rural children and defines the treatment group as those living in regions with an initial below-median rural enrollment rate, and the control group as all others. The second setting works because UNCEP's egalitarian nature necessarily led to stronger policy measures, like closer government supervision and greater fiscal expenditure, in regions where compulsory education enrollment was lagged.[3] The variable *year*05 is a time dummy variable indicating whether the observation was from the 2005 1 Percent National Population Sample Survey or not. The coefficient δ of the interaction item $treat \times year05$ measures the relative change in enrollment probabilities of the two groups between 2000 and 2005. A positive estimation of the coefficient δ suggests that the junior high school enrollment rate increased faster in the treatment group than in the control group.

[3]Under this setting, *treat* only varies with prefectures p, so it should be written as $treat_p$.

$\boldsymbol{X}_{ipt}$ is a series of individual characteristics, including *female*, indicating whether the respondent was female; *minority*, indicating whether the respondent belonged to an ethnic minority; *age* and *age_square*, representing one's age and its quadratic term; *father_edu* and *mother_edu,* referring to the years of education of one's father and mother,[4] respectively; *siblings*, denoting the number of siblings; and ln*area_pc*, the logarithmic housing area per capita, acting as a proxy variable for family income and wealth. In view of systematic sampling differences between the 2000 Census and the 2005 survey and the time-varying impact of individuals' features on enrollment, the model also includes all the interaction items of $\boldsymbol{X}_{ipt}$ and *year*05. Finally, γ_p represents the fixed effects of observations' resident prefectures,[5] added to control time-invariant regional characteristics.

The validity of regressions using pooled data with different sampling rules requires clarification. Our sample of 2000 comes from the 2000 Census via simple random sampling, which is ideal from the regression perspective. The 2005 1 Percent National Population Sample Survey adopted probability proportionate to size sampling. However, the subsample used here experienced non-probability resampling. Hence, there is a hidden danger that unequal sampling weights may affect the causal effects estimation. This issue can be divided into three distinct problems: heteroskedasticity, endogenous sampling, and possible bias of average partial effects (Solon et al., 2015). To manage heteroskedasticity, in every single regression we used robust standard error clustered at the prefectural level. Endogenous sampling refers to cases in which the regression error term is related to the sampling criteria. In our pooled data, weight variation originated from time and resident locations. We controlled for the time dummy and all regional fixed effects, and thus the probability of selection should no longer be related to the error term. With regard to the problem of biased average partial effects caused by heterogeneous effects, the paper aimed to study the heterogeneous effect rather than simply averaging it out. In summary, the issue of unequal weights cannot confound the following results.

---

[4]Parents' years of education are calculated according to the education levels that are reported based on normal schooling length, which is 6 years of primary school, 3 years of junior high school, 3 years of high school, 4 years of undergraduate education, and 3 years of graduate education. For example, "graduated from primary school" is assigned 6 years, and "graduated from junior high school" is assigned 9 years (6 + 3). The years of an education phase are reduced by half if people dropout or fail to graduate. For example, "dropout from junior high school" is assigned 7.5 years (6 + 3/2).

[5]In our sample, over 97 percent observed that their *hukou* locations coincided with current residential prefectures, so migration across regions was immaterial.

(2) The parallel trend test

The validity of the DID strategy relies on a underlying parallel trend between the treatment group's dependent variable and that of the control group. Our data only has two waves rather than multiple periods, so we conducted an indirect parallel trend test by analyzing the upgrading rate from primary school to junior high school among different age cohorts with only the 2005 data, following the method described by Lu and Zhang (2019). Two facts must be explained to interpret the procedure fully. First, the junior high school enrollment rate is closely related to the upgrading rate from primary school to junior high school. In the eligible sample for Equation (1), 83.1 percent of those not in junior high schools never enrolled in one at all. The enrollment rate and the upgrading rate are supposed to change simultaneously. Second, students in China usually attend junior high school when they reach the age of 12. We can therefore estimate the time series of the upgrading rate based on a cross-sectional dataset by calculating the share of individuals who had ever entered junior high school in each corresponding age cohort. For example, in the 2005 1 Percent National Population Sample Survey data, people who were registered as 22 years old were 12 in 1995. In this group, the share of people who had junior high school experiences can serve as an estimator of the upgrading rate from primary school to junior high school in 1995.

To claim valid DID identification requires that the gap between upgrading rates of the control group and treatment group did not change prior to UNCEP's implementation but did so after. To examine the trend of the gap quantitatively, we estimate a regression model similar to Equation (1). The specific form is:

$$JH_{ipc} = \alpha treat_{ipc} + \beta_c \boldsymbol{COHORT}_c + \delta_c \boldsymbol{COHORT}_c \times treat_{ipc} + \pi_c \boldsymbol{X_{ipc}} + \gamma_p + \varepsilon_{ipc}, \quad (2)$$

where i, p, and c represent individual i, prefecture p, and cohort c, respectively. The dependent variable, *JH*, indicates whether an individual had entered junior high school. ***COHORT*** is a set of cohort dummies indicating the year in which the individual was supposed to start junior high schooling, from *cohort*1995 to *cohort*2005. For example, people born between September 1982 and August 1983 typically ought to graduate from primary school and attend junior high school in September 1995 and thus have value 1 for *cohort*1995 and 0 for other cohort dummies. The cohort corresponding to 1994 is set as an omitted benchmark cohort.

The treatment variable is slightly different from that in Equation (1). Here, we divide individuals into the treatment and control groups based on their *hukou* registration instead of permanent residence. People with agriculture *hukou* are considered to be rural residents at age 12, and *hukou* locations are regarded as where they were supposed to go to junior high school. Hence, the *hukou* category and location tell us whether an individual was treated heavily or slightly. We do not continue to assign treatment variables according to resident information because young people probably leave their home towns to find work when they reach the legal working age. Biases emerge if we identify people's origins based on their current places of residence. However, *hukou* are difficult to change unless the individual joins the army or has a college degree and is regarded as a skilled laborer, so *hukou* can assist in tracing the past.

The coefficient of the variable *treat* measures the size of the upgrading rate gap between the treatment and control groups in 1994. The coefficients of the interaction terms of cohort dummies and *treat* measure changes in the gap over time. If the parallel trend assumption holds, we should observe that the coefficients of *cohort*1995 × *treat* to *cohort*1998 × *treat* are insignificant, and those of *cohort*1999 × *treat* to *cohort*2005 × *treat* are significantly positive.

Individual characteristics ***X*** are the same as in Equation (1), except that the variable ln*area_pc* is excluded because of the reverse causality problem. Again, we allow the impacts of ***X*** on the upgrading rate to vary over time, by adding the interaction terms of ***X*** and cohort dummies. As in Equation (1), γ_p represents the prefecture fixed effects.

(3) Triple difference heterogeneity analysis: Confounding policies and rural population density

Due to other simultaneous policies, a parallel pre-trend may not be sufficient to assert the causality between UNCEP and the increase in rural junior high school enrollment rate. Lu and Zhang (2019) found that China's Higher Education Expansion (HEE) from 1999 strongly stimulated senior high school participation among rural students, because HEE greatly increased their chance of colleges admissions. It is likely that HEE also affected junior high school eligible children and elevated the enrollment rate. Another policy is China's Rural School Consolidation (RSC) through the early 2000s, which shut down dispersed small primary schools and junior high schools and enhanced central and large ones. Evidence shows that the RSC may have positive effects on school participation by improving education quality (Liang and Wang, 2020).

Both HEE and RSC are characterized by major policy intensity variations. If the improvement in junior high school participation is caused by these two policies, we should observe a positive association between policy intensity and DID effect heterogeneity. To test the alternative hypotheses, we estimated a DDD model derived from Equation (1). It takes the following form:

$$enroll_{ipt} = \alpha treat_{ipt} + \beta year05_t + \delta treat_{ipt} \times year05_t + \varphi_1 treat_{ipt} \times intensity_p + \varphi_2 year05_t \times intensity_p + \theta treat_{ipt} \times intensity_p \times year05_t + \pi_t \boldsymbol{X_{ipt}} + \gamma_p + \varepsilon_{ipt}, \quad (3)$$

where *intensity* indicates HEE and RSC policy intensity, taking values of the expansion multiple of local college admissions from 1999 to 2005 and the share of schools closed from 2000 to 2005, respectively. If an insignificant θ is obtained, we can claim that the DID effect is not due to these two policies. Hence, UNCEP's causal effect on junior high school participation will be further verified.

The DDD method also applies to the heterogeneity analysis of rural population density's moderating effect, the key component of this study. We estimated the following model:

$$\begin{aligned} enroll_{ipt} &= \alpha treat_{ipt} + \beta year05_t + \delta treat_{ipt} \times year05_t + \varphi_1 treat_{ipt} \times density_p + \\ &\varphi_2 year05_t \times density_p + \theta treat_{ipt} \times density_p \times year05_t + \phi_1 treat_{ipt} \times Z_p + \\ &\phi_2 year05 \times Z_p + \lambda treat_{ipt} \times Z_p \times year05_t + \pi_t \boldsymbol{X_{ipt}} + \gamma_p + \varepsilon_{ipt}, \end{aligned} \tag{4}$$

where *density* indicates the prefectural rural population density. All the interaction terms, *density*, *treat*, and *year*05 (including the triple interaction term) are added to capture the differential effect of UNCEP along the rural population density dimension. The interaction terms of some other prefectural-level variable *Z*, *treat*, and *year*05 are also introduced into the model to control for UNCEP's uneven intensity. Such prefectural variables include local initial rural junior high school enrollment rate in 2000 and growth rate of local fiscal education expenditure per student from 2000 to 2005. If the UNCEP is more effective in densely populated regions, θ, the coefficient of *rural* × *density* × *year*05, should be significantly positive.

3. Summary statistics

Based on the data and the baseline model setting, we have summary statistics of individual-level variables sorted by year, shown in Table 1. The junior high school enrollment rate of eligible children increased by about 4 percent from 2000 to 2005. Rural children had an initial enrollment rate below average, but the 5 percent progress is above average, so an uncontrolled DID effect exists under the rural/urban treatment setting (treatment_1). Within the rural sample, a similar pattern also holds with the backward regions and advanced regions treatment setting (treatment_2), as the enrollment rates of the treatment group and the control group were 0.80 and 0.94 in 2000, and 0.90 and 0.96 in 2005, respectively. The 2005 data place more weight on observations in western regions, which were less developed. As a result, the real average of father's education level and mother's education level should be higher than the means shown. The means of other variables, treatment_2, for example, may also be different from real national averages in 2005.

Table 1. Summary statistics at the individual level

Variables	2000			2005		
	Mean	Standard deviation	Observations	Mean	Standard deviation	Observations
	(1)	(2)	(3)	(4)	(5)	(6)
Full sample						
Studying in a junior high school (1 = yes)	0.91	0.29	42,627	0.95	0.23	50,787
Treatment_1 (1 = rural area)	0.71	0.45	42,627	0.62	0.48	50,787
Gender (1 = female)	0.48	0.50	42,627	0.49	0.50	50,787
Minority (1 = yes)	0.09	0.28	42,627	0.11	0.31	50,787
Age (year)	14.2	1.14	42,627	14.4	1.14	50,787
Father's education level (year)	8.79	2.71	42,627	8.63	2.84	50,787
Mother's education level (year)	7.38	3.07	42,627	7.34	3.27	50,787
Siblings (person)	1.35	1.02	42,627	1.31	1.06	50,787
Average housing space (m^2)	22.8	14.5	42,627	23.7	15.4	50,787
Subsample in rural area						
Studying in a junior high school (1 = yes)	0.88	0.32	30,175	0.93	0.26	31,499
Treatment_2 (1 = backward region)	0.41	0.49	30,175	0.53	0.50	31,499
Gender (1 = female)	0.48	0.50	30,175	0.49	0.5	31,499
Minority (1 = yes)	0.10	0.29	30,175	0.13	0.34	31,499
Age (year)	14.2	1.14	30,175	14.5	1.13	31,499
Father's education level (year)	8.36	2.50	30,175	7.98	2.51	31,499
Mother's education level (year)	6.78	2.83	30,175	6.51	2.98	31,499
Siblings (person)	1.53	1.00	30,175	1.49	1.05	31,499
Average housing space (m^2)	22.4	13.6	30,175	22.5	13.8	31,499

Sources: Authors' calculation based on 2000 Census and 2005 National 1 Percent Population Survey.

Table 2 outlines the summary statistics of the prefecture-level variables. It can be seen that rural population densities are quite dispersed among different prefectures. The intensities of UNCEP and the simultaneous HEE and RSC also vary greatly. Thus, the DDD estimations are feasible.

Table 2. Summary statistics at the prefectural level

Variables	Mean	Standard deviation	Min	Max	Observations
	(1)	(2)	(3)	(4)	(5)
Rural population density (person/km²)	262.3	172.2	3.1	828.8	243
Per capita education expenditure in 2000 (RMB)	548.1	240.8	185.2	1,559.0	243
Per capita education expenditure in 2005 (RMB)	1408.9	699.9	540.7	5,718.6	243
Growth rate of per capita education expenditure (intensity of UNCEP, %)	162.5	66.0	–13.7	409.0	243

(Continued on the next page)

(Table 2 continued)

Variables	Mean	Standard deviation	Min	Max	Observations
	(1)	(2)	(3)	(4)	(5)
Rural junior high enrollment rate in 2000 (intensity of UNCEP, %)	88.5	8.4	46.2	100.0	243
Expansion rate of admissions of local colleges (intensity of HEE, %)	404.5	89.6	234.1	623.4	243
Share of primary schools closed (intensity of RSC, %)	34.6	30.0	−215.4	97.7	243

Sources: Authors' calculation based on *China Statistical Yearbook* 2001 and 2006 by the National Bureau of Statistics of China; *China City Statistical Yearbook* 2001 and 2006 by the National Bureau of Statistics of China; *National Financial Statistical Data of Cities and Counties* 2000 and 2005 by the Ministry of Finance of China; 2000 Census.

Notes: The expansion rate of admissions of local colleges is a provincial-level variable. Negative values for share of primary schools closed mean that regional primary schools increased during 2000 to 2005. The rural junior high enrollment rate in 2000 is calculated based on microdata, so it can be very low or high (even 100 percent) if the sample in a prefecture is not large enough. UNCEP, universal 9-year compulsory education policies.

IV. Results

1. Rural junior high school participation in 2000: Cross-sectional association between enrollment rate and population density

Before considering UNCEP's effect and how it is moderated by rural population density, it is helpful to demonstrate the differences in initial rural junior high school enrollment rates among regions. As shown in Table 3, rural population density declined from east to west in 2000. Interestingly, the variation in the rural junior high school enrollment rate followed the same pattern.

Table 3. Rural population density and junior high school enrollment rate among different regions in China (2000)

Region	Average of rural population density (person/km^2)	Average of rural junior high school enrollment rate (%)
Eastern prefectures	318.8	91.8
Central prefectures	242.1	86.0
Western prefectures	125.2	77.2

Sources: Authors' calculation based on the data from *China Statistical Yearbook* (National Bureau of Statistics of China 2001); *China City Statistical Yearbook* (National Bureau of Statistics of China 2002); 2000 Census.

Is the similarity between the difference in population density and that of the enrollment rate a meaningful association or just a trivial correlation? After all, there is also a negative development gradient from the eastern coastal regions to the western regions. Prefectures with low rural population density are also those with a weak education supply, and the residents have characteristics working against school participation, such as belonging to ethnic minorities and having low incomes. In response to the question, based on the rural eligible sample of 2000, we ran a simple regression in which the strength of education supply and individual-level characteristics are controlled. From Table 4, we find that rural population density still matters even though impacts of other factors are eliminated. A 1 percent change in population density corresponds to an change in enrollment probability of approximately 0.03 percent. As a proxy variable for public education supply, prefectural fiscal education expenditure per student also has a positive effect on the probability of rural children's junior high school enrollment. However, both the economic importance and statistical significance of the coefficient of *edu_exp_ps* are weaker than that of *density*. Of course, the results in Table 4 are not sufficient to verify the causal effect of population density on rural education because of endogeneity, yet it is enough to highlight the topic's importance.

Table 4. Determinants of rural junior high school enrollment probability: Population density, education supply, and individual characteristics

Variables	Dependent variable: Being at junior high school (yes = 1, no = 0)
density	0.031*** (0.003)
edu_exp_ps	0.019** (0.008)
female	−0.050*** (0.017)
minority	−0.041 (0.005)
age	0.547*** (0.046)
age_sq	−0.020*** (0.002)
father_edu	0.013** (0.001)
mother_edu	0.010*** (0.001)
siblings	−0.010*** (0.001)
hous_area_pc	0.035*** (0.005)
R^2	0.074
Observations	26,194

Notes: *** and ** represent significance at the 1 and 5 percent levels, respectively. *density* and *edu_exp_ps* indicate logarithmic rural population density and per-student fiscal education expenditure, respectively. See Table 1 and section III for individual variable's definitions. Robust standard errors clustered at prefectural level are in parentheses.

The effects of individual and family characteristics are also in line with common sense, as females and minorities were less likely to attend school, and those with better educated parents, fewer siblings, or wealthier families were more likely to be at school. The relationship between enrollment probability and age is non-monotonic. Children too young or too old were likely to be out of school. In Table 4, the coefficients of the individual-level variables are all highly significant, justifying controlling for such variables in DID and DDD regressions.

2. Baseline difference-in-differences effects

Table 5 reports the baseline results of DID estimation for UNCEP's effect on the junior high school enrollment rate in rural areas. Column (1) shows the result of DID estimation based on the full sample and rural/urban treatment setting. We found that after controlling for other observables, rural children's enrollment probability was 0.034 less than that of urban children in 2000, but the former increased by 0.023 over the latter from 2000 to 2005. The estimated coefficient of *year*05 is omitted because the interaction terms of *year*05 and the individual characteristics ***X*** are controlled and thus coefficient of *year*05 per se does not have a clear meaning. In column (2), we present the result of regression using only the rural sample and applying the second treatment setting, grouping individuals according to their regions of residence: those living in prefectures where the initial rural junior high school enrollment rate is below the median of regional rural enrollment rates were placed in the treatment group, and the others were placed in the control group. It can be seen that the treatment group's probability of enrollment increased 0.069 over that of the control group. Baseline DID estimation is therefore in line with our expectation.

Table 5. Difference-in-differences estimations of rural junior high school enrollment rates (2000−2005)

Variables	Dependent variable: Being at junior high school (yes = 1, no = 0)	
	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.034*** (0.003)	
treat × *year*05	0.023*** (0.003)	0.069*** (0.007)
*year*05	Yes	Yes
Individual characteristics	Yes	Yes
Individual characteristics × *year*05	Yes	Yes
Prefectural FE	Yes	Yes
R^2	0.123	0.136
Observations	93,414	61,674

Notes: *** represents significance at the 1 percent level. Robust standard errors clustered at prefectural level are in parentheses. The variable *treat* in column (2) is omitted because of its perfect collinearity with regional fixed effects. FE, fixed effects.

3. Causality tests

(1) Parallel trend test

We have found that the junior high school enrollment rate of the treatment group grew faster than that of the control group from 2000 to 2005, a promising DID result, so the next step is to attribute it to UNCEP. To check causality, we conducted the indirect parallel trend test described in Section III. The results of two regression settings are presented in Tables 6 and 7, respectively.

Table 6. The trend of the upgrading probability gap between the treatment and control groups (rural and urban sample; *treat* = treatment_1)

Dependent variable: Had been at junior high school (yes = 1, no = 0)			
treat	−0.031*** (0.006)	*cohort*2000 × *treat*	0.014** (0.007)
*cohort*1995 × *treat*	0.006 (0.009)	*cohort*2001 × *treat*	0.016** (0.008)
*cohort*1996 × *treat*	−0.001 (0.009)	*cohort*2002 × *treat*	0.032*** (0.007)
*cohort*1997 × *treat*	−0.001 (0.008)	*cohort*2003 × *treat*	0.032*** (0.007)
*cohort*1998 × *treat*	0.003 (0.009)	*cohort*2004 × *treat*	0.036*** (0.007)
*cohort*1999 × *treat*	0.010 (0.007)	*cohort*2005 × *treat*	0.028*** (0.008)
Cohort dummies		Yes	
Individual characteristics		Yes	
Individual characteristics × Cohort dummies		Yes	
Prefectural FE		Yes	
R^2		0.146	
Observations		138,902	

Notes: *** and ** represent significance at the 1 and 5 percent levels, respectively. Robust standard errors clustered at prefectural level are in parentheses. FE, fixed effects.

In Table 6, the coefficient of *treat* is −0.031 and significant, which means that in comparison with the urban primary school graduates, rural primary school graduates' upgrading probability is less by 0.031 in 1994. The coefficients of *cohort*1995 × *agri* to *cohort*1998 × *agri* are all close to zero and insignificant – evidence suggesting that there was no relative change between rural and urban upgrading rates from 1994 to 1998. However, the coefficients start increasing from *cohort*1999 × *agri* and become significant from *cohort*2000 × *agri*, indicating that the rural upgrading rate grew faster from 1999.[6]

[6]The reason why the coefficient of *cohort*2005 × *agri* shrinks may be that some primary school graduates are not enrolled into junior high school immediately but months later. In the 2005 data, these people are included as "junior high school or above educated" for early cohorts, but not for the 2005 cohort.

Table 7. The trend of the upgrading probability gap between the treatment and control groups (rural sample only; *treat* = treatment_2)

Dependent variable: Had been at junior high school (yes = 1, no = 0)			
*cohort*1995 × *treat*	−0.013 (0.014)	*cohort*2001 × *treat*	0.033** (0.012)
*cohort*1996 × *treat*	−0.016 (0.012)	*cohort*2002 × *treat*	0.039*** (0.011)
*cohort*1997 × *treat*	0.007 (0.013)	*cohort*2003 × *treat*	0.042*** (0.010)
*cohort*1998 × *treat*	0.000 (0.012)	*cohort*2004 × *treat*	0.051*** (0.011)
*cohort*1999 × *treat*	0.006 (0.012)	*cohort*2005 × *treat*	0.053*** (0.011)
*cohort*2000 × *treat*	0.017 (0.011)		
Cohort dummies		Yes	
Individual characteristics		Yes	
Individual characteristics × cohort dummies		Yes	
Prefectural FE		Yes	
R^2		0.146	
Observations		108,543	

Notes: *** and ** represent significance at the 1 and 5 percent levels, respectively. Robust standard errors clustered at prefectural level are in parentheses. Variable *treat* is omitted in the regression due to perfect collinearity between *treat* and fixed effects of prefectures. FE, fixed effects.

Table 7 presents the results of the regression based on the rural sample and treatment_2, which are quite similar to the results in Table 5. Before measures of UNCEP being taken, there was no convergence tendency between the upgrading probability of the treatment group and that of control group, but after 2000. In conclusion, there is a parallel pre-trend in the DID identification. Moreover, the results in Tables 6 and 7 can serve as robustness checks of the baseline results. The DID in upgrading probabilities of the treatment and control groups from 2000 echoes the DID in enrollment probabilities shown in Table 5. The relative improvement of rural school participation is double checked.

(2) Impacts of confounding policies

Regarding confounding policies, HEE and RSC, Table 8 provides the results of the DDD tests. As shown in Table 8, the coefficient of the triple interaction term is close to zero and statistically insignificant. The DID effect does not vary with the intensity of the HEE, which is entirely different from the findings of Xing (2014) or Lu and Zhang (2019). The HEE is therefore not the cause of the convergence of enrollment probabilities of the treatment and control groups. However, there is a question about why the HHE can promote enrollment in senior high school rather than junior high schools. We believe individuals who can be affected by the change in college admission chances are not at the same margin as those who still hesitate to enter junior high school. The HHE only affected relatively competitive students, and they would almost certainly enter junior high school. Those reluctant to finish compulsory education could only respond to supply-side policies.

Table 8. Triple difference estimation regarding higher education expansion policy

	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.043*** (0.004)	
treat × *year*05	0.027*** (0.004)	0.071*** (0.007)
treat × *HEEintensity*	0.024*** (0.005)	
HEEintensity × *year*05	−0.002 (0.003)	0.005 (0.004)
treat × *HEEintensity* × *year*05	−0.005 (0.005)	−0.002 (0.007)
*year*05	Yes	Yes
Individual characteristics	Yes	Yes
Individual characteristics × *year*05	Yes	Yes
Prefectural FE	Yes	Yes
R^2	0.124	0.137
Observations	93,414	61,674

Notes: *** represents significance at the 1 percent level. Robust standard errors clustered at prefectural level are in parentheses. Because of perfect collinearity, *HEEintensity* is omitted in column (1) and *treat*, *HEEintensity*, and *treat* × *HEEintensity* are omitted in column (2). *HEEintensity* implies expansion rate of admissions of local colleges. *HEEintensity* has been standardized at prefectural level. FE, fixed effects.

As Table 9 shows, the coefficients of the DDD terms regarding the Rural School Consolidation policy are also economically and statistically insignificant. This result suggests that the baseline DID effect cannot be an outcome of the RSC either. According to the literature, the RSC has advantages and disadvantages, and the net effect is ambiguous (Cai et al., 2017; Liang and Wang, 2020). Our results are, to some extent, in line with those studies. Despite great regional variations in the RSC policy intensity (described in part III), the rural-urban junior high school enrollment gap is not correlated with it. The RSC seems to have a close to zero net effect on compulsory education enrollment. Real influences may merely exist in the upper secondary education phase or at intensive margins, such as academic achievements (Liu et al., 2010; Liang and Wang, 2020).

Table 9. Triple difference estimation regarding rural school consolidation policy

	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.029*** (0.003)	
treat × *year*05	0.022*** (0.004)	0.064*** (0.007)
treat × *RSCintensity*	0.009** (0.004)	
RSCintensity × *year*05	−0.000 (0.003)	0.001 (0.005)
treat ×*RSCintensity* × *year*05	−0.003 (0.004)	0.001 (0.006)
*year*05	Yes	Yes
Individual characteristics	Yes	Yes
Individual characteristics × *year*05	Yes	Yes
Prefectural FE	Yes	Yes
R^2	0.090	0.103
Observations	77,107	50,512

Notes: *** and ** represent significance at the 1 and 5 percent levels, respectively. Robust standard errors clustered at prefectural level are in parentheses. Because of perfect collinearity, *RSCintensity* is omitted in column (1) and *treat*, *RSCintensity*, and *treat* × *RSCintensity* are omitted in column (2). *RSCintensity* implies share of primary schools closed. *RSCintensity* has been standardized at prefectural level. FE, fixed effects.

Some other major events happened in the early 2000s and they might have affected rural education, like China’s entry into WTO in 2001 and Rural Tax for Fee Reform from 2001. The literature indicates that China’s entry into WTO probably reduced the enrollment rate of rural children because the demand for low-skilled labor increased and the opportunity cost of schooling increased (Zhang, 2015). With regard to the Rural Tax for Fee Reform, Zhou and Chen (2015) found that, because of the reform, county-level fiscal conditions worsened and thus rural public services decreased. Alm and Liu (2013) also reported that the reform appeared to have damaged the villages’ financing capacity, and hence to have lowered their overall expenditure. So rural education was very likely to deteriorate. We can therefore only underestimate the effects of UNCEP if China’s entry into WTO or Rural Tax for Fee Reform plays a part. Moreover, trend analysis indicated that the turning point was exactly 1999 and only UNCEP fits this.

4. Heterogeneity: Role of rural population densities

As described in Section III, DDD regressions can also be used to show how population density affects the effectiveness of UNCEP. Notably, UNCEP is egalitarian and focuses on urging children in backward rural areas to attend school, so the strength or intensity of UNCEP necessarily varied across regions depending on the initial level of school participation. For regions where initial enrollment rates are low, central and local governments inevitably put more effort into promoting school participation, by closer supervision and greater fiscal expenditure for instance, until the enrollment gap was filled. As Tables 3 and 4 show, there is a positive correlation between rural population density and the initial rural junior high school enrollment rate; hence it is very likely that regions with lower rural population densities have larger DID effects. However, that is not because the policies were more effective in such regions, but the policy intensity is stronger. To isolate policy effectiveness from the gross policy effect, policy intensity must be carefully controlled. Obviously the initial enrollment rate can act as a good proxy for policy intensity. Nevertheless, initial enrollment rate might not fully capture the variation of the increase in expenditure on public education, so we also control the latter directly.

In the DDD regressions, we add all the interaction terms of treatment variables, time dummy, and policy intensity variables (including triple interaction terms). For the treatment_2 setting, there are two policy intensity variables, regional rural junior high school enrollment rate in 2000 and growth rate of local fiscal education expenditure per student from 2000 to 2005, respectively. In the treatment_1 setting, the local urban junior high enrollment rate acts as the reference for the rural enrollment rate. We further control the regional initial gap for urban and rural enrollment rates in this setting to exclude the possibility that the larger enrollment gap drew more policy attention.

Table 10 presents the results. Regardless of treatment settings, the estimated coefficient of $treat \times density \times year05$ is significantly positive. That is, in terms of encouraging rural children to enter junior high school, UNCEP is more effective in densely populated regions. The influence of rural population density is also very strong. For every standard deviation of rural population density at the prefectural level, the DID effect varies by 0.019 and 0.041, respectively, under two treatment settings, a striking heterogeneity in contrast with the average DID effect of 0.023 and 0.069. Hence, we can claim that low population density in rural areas is indeed a serious impediment to local education development.

There is a difference between our heterogeneity results and the results of Duflo (2001). In that study, a large-scale school construction plan in Indonesia was evaluated and the results suggest that the program had no effect in densely populated regions, and a large effect in sparsely populated regions. This finding does not necessarily contradict ours because UNCEP's nature is distinct from that of school construction. Measures of UNCEP are aimed at strengthening existing schools rather than building new schools. By generous government funding, the quality of teachers and facilities were improved and costs borne by students were cut. While new school construction significantly reduces the cost of going to school in sparsely populated regions, the effect of school strengthening is limited for children living in remote villages. These two facts verify the importance of population density from different perspectives. However, faced with rapid urbanization in developing countries, radical school construction plans are neither economical nor practical. The UNCEP measures will probably become appealing options for many countries to improve rural education. We therefore believe that our results provide useful information for future policymaking.

Table 10. The differential effectiveness of UNCEP on rural junior high school enrollment probabilities among regions with different population densities

	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.065*** (0.003)	
treat × *year*05	0.039*** (0.005)	0.010 (0.012)
treat × *density*	−0.006 (0.005)	
density × *year*05	0.003 (0.006)	0.009 (0.006)
treat × *density* × *year*05	0.019*** (0.007)	0.041*** (0.011)
*year*05	Yes	Yes
Individual characteristics	Yes	Yes
Individual characteristics × *year*05	Yes	Yes
Prefectural FE	Yes	Yes
All the interaction terms of *Z* (intensity of UNCEP), *treat*, and *year*05	Yes	Yes
R^2	0.099	0.104
Observations	76,578	50,398

Notes: *** represents significance at the 1 percent level. Robust standard errors clustered at prefectural level are in parentheses. Because of perfect collinearity, *density* is omitted in column (1) and *treat*, *density*, and *treat* × *density* are omitted in column (2). FE, fixed effects. UNCEP, universal 9-year compulsory education policy.

5. Robustness checks

Some of the previous regressions have played the role of robustness checks. For example, the two treatment settings are robustness checks to each other; the parallel trend tests can be regarded as robustness checks to the baseline results as the tests show a significant DID from 2000. Tables 11 and 12 further demonstrate the robustness of our key results. In Table 11, individual variables are not controlled. The coefficients of DDD terms are very close to that in Table 10. It is proven that the DDD effect is strong and insensitive to individual level controls, so the effect is unlikely to be driven by changes in rural households' characteristics such as the income.

Table 11. The different effectiveness of UNCEP on rural junior high enrollment probabilities among regions with different population densities

	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.112*** (0.002)	
treat × *year*05	0.063*** (0.005)	0.004 (0.011)
treat × *density*	0.004 (0.004)	
density × *year*05	0.006 (0.005)	0.006 (0.006)
treat × *density* × *year*05	0.017** (0.007)	0.045*** (0.011)
*year*05	Yes	Yes
Individual characteristics	No	No
Individual characteristics × *year*05	No	No
Prefectural FE	Yes	Yes
All the interaction terms of *Z* (intensity of UNCEP), *treat*, and *year*05	Yes	Yes
R^2	0.057	0.055
Observations	76,578	50,398

Notes: *** and ** represent significance at the 1 and 5 percent levels, respectively. Robust standard errors clustered at prefectural level are in parentheses. Because of perfect collinearity, *density* is omitted in column (1) and *treat*, *density*, and *treat* × *density* are omitted in column (2). *density* and *Z* have been standardized at prefectural level. Variables at individual level are not controlled. FE, fixed effects. UNCEP, universal 9-year compulsory education policy.

Table 12 reports results (average marginal effects) generated by the probit model. For the treatment_1 setting, the probit model reduces the coefficient of *treat* × *density* × *year*05 and inflates the variance, so the coefficient is no longer significant. However, under the treatment_2 setting, the coefficient of *treat* × *density* × *year*05 is still highly significant. Average marginal effects of the probit model indicate probabilistic marginal effects of independent variables when all independent variables take values of sample averages. An unequally weighted sample is used and many regional dummies are controlled, so the average marginal effects are hard to interpret. Besides, probit or logit models make strong assumptions about the distribution of random terms. By comparison, the results generated from linear probability model have immediate interpretations and require fewer econometric restrictions, so we think they are more relevant.

Table 12. The differential effectiveness of universal 9-year compulsory education policy (UNCEP) on rural junior high enrollment probabilities among regions with different population densities (probit model)

	(1) Rural and urban sample *treat* = treatment_1	(2) Rural sample only *treat* = treatment_2
treat	−0.048*** (0.003)	
treat × *year*05	0.023*** (0.005)	−0.013 (0.018)
treat × *density*	−0.000 (0.006)	
density × *year*05	0.006 (0.008)	0.006 (0.009)
treat × *density* × *year*05	0.013 (0.009)	0.044*** (0.011)
*year*05	Yes	Yes
Individual characteristics	Yes	Yes
Individual characteristics × *year*05	Yes	Yes
Prefectural FE	Yes	Yes
All the interaction terms of *Z* (intensity of UNCEP), *treat* and *year*05	Yes	Yes
Pseudo R^2	0.181	0.162
Observations	76,024	49,898

Notes: *** represents significance at the 1 percent level. Robust standard errors clustered at prefectural level are in parentheses. Coefficients are average marginal effects. Because of perfect collinearity, *density* is omitted in column (1) and *treat*, *density*, and *treat* × *density* are omitted in column (2). The variables *density* and *Z* have been standardized at prefectural level. FE, fixed effects. UNCEP, universal 9-year compulsory education policy.

V. Conclusions

This paper examined how rural population density is related to the outcome of supply-sided education policies, taking advantage of a set of nationwide expansionary education policies in China, the universal 9-year education policies. We found that UNCEP was effective in increasing the rural junior high school enrollment rate, as the enrollment probability of eligible rural children was increased by 0.023 on average. More important, after controlling for other confounding factors, the policies were more productive in regions with higher rural population densities.

This study provides evidence that population density is a significant factor in rural education. Obviously, the low-density dilemma of rural education cannot be solved spontaneously without collective action, so the government's role is of prime importance. In the short run, policymakers should give more attention to the immediate problems associated with low population density, such as long-distance commuting.

Providing school buses and lodgings in low population-density areas might be an effective way to elevate or maintain the rural education enrollment rate. In the long run, however, to improve the efficiency of the whole education system, urbanization and village amalgamation should be embraced to fully exploit economies of scale.

References

Alm, J. and Y. Liu, 2013, "Did China's Tax-for-Fee Reform improve farmers' welfare in rural areas?" *Journal of Development Studies*, Vol. 49, No. 4, pp. 516–32.

Andrews, M., W. Duncombe and J. Yinger, 2002, "Revisiting economies of size in American education: Are we any closer to a consensus?" *Economics of Education Review*, Vol. 21, No. 3, pp. 245–62.

Ares Abalde, M., 2014, "School size policies: A literature review," *OECD Education Working Papers* No. 106, Organization for Economic Co-operation and Development Publishing, Paris.

Bard, J., C. Gardener and R. Wieland, 2006, "Rural school consolidation: History, research summary, conclusions, and recommendations," *Rural Educator*, Vol. 27, No. 2, pp. 40–8.

Brown, P. H. and A. F. Park, 2002, "Education and poverty in rural China," *Economics of Education Review*, Vol. 21, No. 6, pp. 523–41.

Cai, W., G. Chen and F. Zhu, 2017, "Has the compulsory school merger program reduced the welfare of rural residents in China?" *China Economic Review*, Vol. 46, pp. 123–41.

Chakraborty, K., B. Biswas and W. C. Lewis, 2000, "Economies of scale in public education: An econometric analysis," *Contemporary Economic Policy*, Vol. 18, No. 2, pp. 238–47.

Che, Y. and L. Zhang, 2018, "Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s," *The Economic Journal*, Vol. 128, No. 614, pp. 2282–320.

Chen, B., P. Zhang and R. Yang, 2010, "Government educational expenditure, human capital investment and urban-rural inequality in China," *Guanli Shijie* (*Management World*), Vol. 26, No. 1, pp. 36–43.

Chyi, H. and B. Zhou, 2014, "The effects of tuition reforms on school enrollment in rural China," *Economics of Education Review*, Vol. 38, No. 1, pp. 104–23.

Duflo, E., 2001, "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *American Economic Review*, Vol. 91, No. 4, pp. 795–813.

Fleisher, B., H. Li and M. Zhao, 2010, "Human capital, economic growth, and regional inequality in China," *Journal of Development Economics*, Vol. 92, No. 2, pp. 215–31.

Huang, W., 2015, "Understanding the effects of education on health: Evidence from China." *IZA Discussion Paper* No. 9225 [online; cited April 2022]. Available from: https://docs.iza.org/dp9225.pdf.

Jia, N. and H. Zhong, 2022, "The causes and consequences of China's municipal amalgamations: Evidence from population redistribution," *China & World Economy*, Vol. 30, No. 4, pp. 174–200.

Krueger, A. B. and M. Lindahl, 2001, "Education for growth: Why and for whom?" *Journal of Economic Literature*, Vol. 39, No. 4, pp. 1101–36.

Li, L. and H. Liu, 2014, "Primary school availability and middle school education in rural China," *Labour Economics*, Vol. 28, pp. 24–40.

Liang, C. and S. Wang, 2020, "The influence of public school allocation on human capital: Evidence from the school consolidation in China," *Jingji Yanjiu* (*Economic Research Journal*), Vol. 55, No. 9, pp. 138–54.

Liu, C., L. Zhang, R. F. Luo, S. Rozelle and P. Loyalka, 2010, "The effect of primary school mergers on academic performance of students in rural China," *International Journal of Educational Development*, Vol. 30, No. 6, pp. 570–85.

Lu, M. and X. Zhang, 2019, "Towards an intelligent country: China's higher education expansion and rural children's senior high school participation," *Economic Systems*, Vol. 43, No. 2, Article 100694.

Luschei, T. F., 2012, "In search of good teachers: Patterns of teacher quality in two Mexican states," *Comparative Education Review*, Vol. 56, No. 1, pp. 69–97.

Ministry of Education of China, 1999a, Bulletin of National Education Development of 1999 [online; cited May 2022]. Available from: http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/tnull_841.html.

Ministry of Education of China, 1999b, *Action Plan of Vitalizing Education Towards the 21st Century* [online; cited May 2022]. Available from: https://www.gmw.cn/01gmrb/1999-02/25/GB/17978%5EGM3-2505.HTM.

Ministry of Education of China, 1996–2006, *Educational Statistics Yearbook of China*, Beijing: China Statistics Press (in Chinese).

Ministry of Education of China, 1997–2007, *China Educational Finance Statistical Yearbook*, Beijing: China Statistics Press (in Chinese).

Ministry of Finance of China, 2001, National Financial Statistical Data of Cities and Counties (*Quanguo Dishixian Caizheng Tongji Ziliao*), Beijing: China Financial and Economic Publishing House (in Chinese).

Ministry of Finance of China, 2007, National Financial Statistical Data of Cities and Counties (*Quanguo Dishixian Caizheng Tongji Ziliao*), Beijing: China Financial and Economic Publishing House (in Chinese).

NBS (National Bureau of Statistics of China), 2001, 2006, *China Statistical Yearbook*, Beijing: China Statistics Press (in Chinese).

NBS (National Bureau of Statistics of China), 2002, 2007, *China City Statistical Yearbook*, Beijing: China Statistics Press (in Chinese).

Sargent, T. and E. Hannum, 2005, “Keeping teachers happy: Job satisfaction among primary school teachers in rural northwest China,” *Comparative Education Review*, Vol. 49, No. 2, pp. 173–204.

Shi, X., 2016, “The impact of educational fee reduction reform on school enrolment in rural China,” *The Journal of Development Studies*, Vol. 52, No. 12, pp. 1791–1809.

Solon, G., S. J. Haider and J. M. Wooldrige, 2015, “What are we weighting for?” *Journal of Human Resources*, Vol. 50, No. 2, pp. 301–16.

The State Council of China, 2003, *Decision on Further Strengthening Rural Education* [online; cited May 2022]. Available from: http://www.gov.cn/zhengce/content/2008-03/28/content_5747.htm.

Tholkes, R. J. and C. H. Sederberg, 1990, “Economies of scale and rural schools,” *Research in Rural Education*, Vol. 7, No. 1, pp. 9–15.

United Nations, 2019, *World Urbanization Prospects: The 2018 Revision*, New York: United Nations.

Maarseveen, R. V., 2021, “The urban–rural education gap: Do cities indeed make us smarter?” *Journal of Economic Geography*, Vol. 21, No. 5, pp. 683–714.

Wan, G. and Y. Zhang, 2017, “Accelerating urbanization explained: The role of information,” *ADBI Working Paper* No. 674, Asian Development Bank Institute, Tokyo.

Wang, Y., 2018, “Educational and nutritional consequences of education subsidy in rural China,” *China Economic Review*, Vol. 51, pp. 167–80.

Wei, Y., 2016, “Teacher mobility in rural China: Evidence from northwest China,” Doctoral Dissertation, Michigan State University [online; cited April 2022]. Available from: https://d.lib.msu.edu/etd/3896/datastream/OBJ/download/Teacher_mobility_in_rural_China___Evidence_from_Northwest_China.pdf.

Xing, C., 2014, “Education expansion, migration and rural-urban education gap: A case study on the effect of university expansion,” *Jingjixue Jikan* (*China Economics Quarterly*), Vol. 13, No. 1, pp. 207–32.

Zhang, C., 2015, “Middle-education trap? Export expansion, employment growth and the individual decision of education investment,” *Jingji Yanjiu* (*Economic Research Journal*), Vol. 50, No. 12, pp. 115–27.

Zhao, M. and P. Glewwe, 2010, “What determines basic school attainment in developing countries? Evidence from rural China,” *Economics of Education Review*, Vol. 29, No. 3, pp. 451–60.

Zhou, L. and Y. Chen, 2015, “Local fiscal burden and public services provision: Impact of rural Tax-for-Fee Reform,” *Jingjixue Jikan* (*China Economics Quarterly*), Vol. 14, No. 2, pp. 417–34.

(Edited by Shuyu Chang)