The Long-term and Distributional Effects of School Consolidations: Evidence from China[1]

Yuheng Zhao[2]

Most Recent Version

Abstract: I study the impacts of the largest primary school consolidation ever implemented, using household surveys in China between 1989 and 2015. Employing a generalized Difference-in-Differences framework to evaluate long-term effects, I find that school consolidations increased educational attainment of females but not males, while increasing earnings of both genders. I provide suggestive evidence consistent with mechanisms of enhanced labor market proximity and reduced scope for gender inequality. Individuals in the middle of the income distribution experienced a larger increase in earnings, and increases in inequality were small relative to the increase in average earnings.

Key words: School consolidation, labor market proximity, gender inequality, distributional effects.
JEL Code: I24, I25, O15, O18

---

[1] Funding was provided by The Institute for Economic Development (IED), Boston University. I am indebted to Dilip Mookherjee for his invaluable advice throughout this project. I would also like to thank Sam Bazzi, Peter Blair, Jing Cai, Bilge Erten, Martin Fiszbein, Siddharth George, Joshua Goodman, Mark Hoekstra, Kevin Lang, Mushfiq Mobarak, Daniele Paserman, and Linh T. Tô for helpful conversations and comments, as well as the participants at NEUDC 2020, the Chinese Economists Society North America Conference 2020, and the Boston University Lecznar Memorial Lecture Series for their excellent feedback. All errors are my own.

[2] Department of Economics, Boston University, 270 Bay State Rd, Boston, MA 02215, Email: zhaoyh@bu.edu.

1. INTRODUCTION

Rural out-migration and fertility decline occurring during the process of development motivates policymakers in developing countries to undertake school consolidations in response to declining student populations in rural areas. Such policies are associated with a trade-off between increasing quality and cost-efficiency of education on one hand and widening access on the other (Patrinos and Bustillo, 2018; Secondo, 2020). The impacts of school consolidations on educational investments are theoretically ambiguous, as they can simultaneously raise the cost of schooling and the return to education. Most of the existing literature focuses on high-income economies (Howley, Johnson, and Petrie, 2011; Engberg et al., 2012; Haan, Leuven, and Oasterbeek, 2016; Lee and Lubienski, 2017; Beuchert et al., 2018; Taghizadeh, 2020).[3] Since developing countries are often characterized by greater gender inequality and lower labor market proximity in rural areas, the relevance of developed country experience may be limited.

This paper studies the long-term impacts of the Chinese school consolidation program, which is the largest in the world. Between 1987 and 2013, China's school consolidation program closed around 600,000 rural primary schools, accounting for more than 80% of its primary schools. Through school consolidation, the program aimed to exploit economies of scale and improve the quality of education in remote areas. School construction was initially targeted toward villages with smaller populations, which also tend to be remote. Students were assigned to a new school if their local schools closed. At the same time, consolidated schools received additional resources. While many rural primary schools were closing, the original village-based education system gradually transformed into a town-based one.

I build a simple model which extends Baland and Robinson (2000) and Bau et al, (2021). Assume each household has one child, and there are two periods. In period 1, parents decide how much to invest in their child's education. Education raises skill and thereby wages earned by the child in period 2, which are traded off against education cost plus foregone child labor wage. The skilled wage rate is jointly determined by proximity to an urban labor market and education quality, while the unskilled wage rate only depends on the former. School consolidation moves the school closer to the town and improved education quality. It raises education costs for rural parents owing to longer distances and additional boarding fees. On the other hand, closer proximity to the urban labor market and higher education quality allows students to earn higher wages. If the changes in the skilled wage dominate changes in the unskilled wage and education costs, households will invest more in education. Otherwise, they will invest less. The latter outcome is more likely for low ability

[3] One exception is Hannum, Liu, and Wang (2021). My paper complements this study by looking at the long-term impacts and studying the distributional effects along different dimensions. I will explain its connection to my paper in more detail later in this paper.

children. This simple model also explains why school consolidations can improve labor market outcomes even if they do not increase educational attainment. Additionally, owing to gender wage differentials, school consolidation can result in heterogeneous treatment effects across genders. Treated females would benefit more from school consolidation than males in education if the increase in skill premium resulting from closer proximity to the urban labor market is larger for females. The model does not distinguish the mechanism of enhanced labor market proximity from that of improved education quality, but I later test for presence of these two potential mechanisms.

When estimating the causal effects of school consolidation, the main challenge is the endogeneity of decisions regarding where and when to close schools. Most schools targeted for closure served poor, remote areas. Thus, an analysis that compares villages that did and did not experience school consolidation will produce biased results. To overcome this bias, I use a generalized Difference-in-Differences (DID) strategy by estimating individual-level regression with cohort and village fixed effects. An individual’s date of birth and region of birth determines its exposure to school consolidation. The treatment group consists of individuals who resided in villages that experienced school closure, and had not finished primary schooling when consolidation occurred. I compare the education and labor market outcomes of treated and control cohorts, flexibly controlling for provincial-level time-variant shocks and static differences between villages that experienced school consolidation in different years. In view of the burgeoning literature on the robustness of two-way fixed effects when there are heterogeneous treatment effects across groups and times, I use an alternative estimate proposed by Cengiz et al. (2019) and de Chaisemartin and D’ Haultfoeuille (2022), which is robust to possible problems caused by heterogeneous treatment effects in a staggered DID design.

I find that after the local primary school closure, the average distance to the nearest rural primary school increased by 1.57 km. As the distance increased, fewer students chose transportation by foot. Students exposed to school consolidation spent 344.1 minutes more on homework per week and played more ball sports than the control group. The average treatment effect on educational attainment is small and insignificant. However, this insignificant result masks remarkable heterogeneities. Educational attainment of males remained unchanged while that of females increased—the latter being larger than a catch-up effect. In addition, probably due to higher education costs after consolidation, students were more likely to drop out before the 3rd grade. Analyses of labor market outcomes show that both genders were more active in labor market. On the extensive margin, treated individuals had a 6.77% higher labor force participation rate. On the intensive margin, treated individuals worked 16.4% longer hours. In terms of earning, school consolidations increased the average wage of employed individuals by 20.2% and income by 22.2%.

Next, I discuss the mechanisms that drive these effects. Better labor market outcomes could be due either to enhanced proximities to the urban labor market or better education quality. In addition, the heterogeneous treatment effects across genders may favor females if there are reductions in gender inequalities because there is less discrimination in the high-educated labor market in towns. To explore the heterogeneous treatment effects across villages, I use a triple difference estimation method which interacts school consolidation exposure with different village-level market access and education characteristics. I use distance to economic zone as a measurement of urban labor market proximity and use teacher wage as an imperfect proxy of teacher's capability, which is an important determinant of education quality.[4] Villages further away from the economic zone experienced a larger increase in income. But the interaction between treatment and teachers' wages is not significant. These findings are more consistent with the prediction of enhanced urban labor market proximity rather than improved education quality. I also find evidence for a reduction in gender inequality. Treated females were less likely to experience child labor. They gained weight and height, were more likely to care about being liked by friends and getting good grades. They were also more likely to migrate, spent less time doing housework, and have more control over TV programs, which might be interpreted as increased bargaining power at home.

Because some treated individuals were more likely to drop out as a result of the policy, there were changes in earning inequality. To assess these distributional effects, I follow Carrell, Hoekstra, and Kuka (2018) and ran a generalized DID in which the dependent variable is an indicator of whether the individual's education level or earning level exceeds a given amount. The results suggest that while exposure to school consolidation did not increase education for most low-educated females, those from the middle to the upper end of the education distribution gained more education. School consolidations had two principal effects: females moved from the middle of the education distribution to the higher part of the education distribution, while males were unaffected. In addition, pooling males and females together, individuals moved from the middle part of the income distribution to the upper end. Overall, school consolidations narrowed inequalities between the middle and top end of the income distribution, but also enlarged it between the bottom and middle. I follow Maitra et al. (2022) to evaluate efficiency-equity trade-offs in the utilitarian tradition of Atkinson (1970). The impact of school consolidations is an increasing, concave function of the incomes with an inequality-aversion parameter θ. A higher value of θ indicates a greater weight on the well-being of worse-off individuals in the social welfare function. The results with different values of θ indicate that aggregate welfare increased, implying that changes in inequality were dominated by the increase in average earnings.

My findings suggest that school consolidations in developing countries can lead to a higher labor force participation rate and higher earnings without sacrificing average

[4] I will explain the institution background of wage determination and how it partly captures the teacher's capability in more detail later in this paper.

educational attainment. The heterogeneous treatment effects also had important implications for inequalities: they narrowed inequalities across villages and genders by enhancing labor market proximity and promoting women's empowerment. Those two mechanisms are likely to be prominent in developing countries, but they have been largely overlooked in the previous literature. However, we should bear in mind that those impacts are conditional on students successfully joining the new schools. School consolidations also enlarged inequalities across households because there were more early dropouts, and those low-educated individuals were likely to come from disadvantaged households.

To the best of my knowledge, I provide the first rigorous and comprehensive empirical evidence of the long-term impacts of school consolidations in a developing economy, with a focus on inequalities and the mechanisms that underlie them. A large empirical literature examines how school consolidations influence students' academic achievements and labor market outcomes. However, the distributional effects are understudied, which are central to understanding the allocative consequence of school consolidations. I contribute to this literature by studying the distributional effects in three different ways: the heterogeneous treatment effects across genders, across households, and across villages. In addition, most of the existing literature has focused exclusively on developed countries. Although school consolidations have been implemented at large scales across many developing countries, there is little information about the long-term impacts and mechanisms of school consolidation in the developing world. Hannum, Liu, and Wang (2021), who provide a good start in this direction, focus on the short-term impacts of Chinese school consolidation on years of schooling for those who have not yet finished their education. They find that in China, school consolidation decreased children's years of schooling, and girls were more vulnerable in this process. My paper differs from theirs by looking at lifetime educational attainment rather than years of schooling in the short run. I find that although school consolidations decreased individuals’ years of schooling by delaying enrollment, they increased females’ lifetime education and had insignificant impacts on males’ lifetime education. Ma (2017) examines the trade-offs faced by decision-makers of school consolidations, and found that schools that were closed were less productive, and distance cost was under-weighted in local governments’ decisions. My paper complements these studies by looking at the long-term impacts that could differ from the short-term ones, and I provide new evidence of the underlying mechanisms.

My paper also contributes to the long-standing discussion about infrastructure investment and development by providing new evidence regarding an increasingly popular but understudied infrastructure policy: school consolidations. The extant literature demonstrates that school expansion and increasing access to school increases educational attainments (Duflo, 2001; Andrabi, Das, and Khwaja, 2013; Burde and Linden, 2013; Kazianga et al., 2013; Muralidharan and Prakash, 2018), which facilitate growth in the developing world. However, whether education infrastructure consolidation precipitated by the demographic transition would contribute to or impede the growth in developing countries remains an open question. I show that school consolidation forced students to attend a new school, which in most cases was located in more populated and prosperous regions. Students who attended new schools had more labor market opportunities, which they would not have had access to if they had remained in their villages.

This study also contributes to a growing literature on gender inequalities and women's empowerment by showing that treatment at an early age can generate persistent effects. Compared with men, women in developing countries have many fewer opportunities in education, the labor market, and political representation. Addressing this issue requires a continuous policy commitment to equality (Duflo, 2012; Ashraf, Field, and Lee, 2014; Baranov et al., 2020). The population studied in this paper, individuals who were between age 0 and 13 when the treatment took place, are understudied in the literature of women's empowerment, which has hitherto focused on adolescent girls or adult women (Ashraf et al., 2020; Dhar, Jain, and Jayachandran, 2022; Bandiera et al., 2020). If school consolidation at an early age has long-term impacts, it would provide a larger window for future interventions that promote women's empowerment and reduce gender inequality.

2. INSTITUTIONAL BACKGROUND

China's primary education in the 1980s featured wide coverage and equalization of education across regions and for both genders. The Chinese government re-established its primary education system after the Cultural Revolution (1966-1976) through "the Decision on Educational System Reform" in 1985 and |"the Compulsory Education Law of the People's Republic of China" in 1986. The central government identified compulsory education as an essential element of basic national policy, and it achieved universal primary school completion by constructing large numbers of primary schools in rural areas.[5] By 2000, 2541 counties, containing 85% of China's total population, had achieved universal primary education. In contrast, in 1949, 80% of the Chinese population was illiterate and less than 20% of school-aged children attended school (Ministry of Education, 2000; Xinhua News Agency, 2009). Primary school gender equity was achieved at the same time: by early 2000, the male enrollment rate of primary school was 99.14% and the female enrollment rate was 99.07%. But a large gender gap in education persisted in higher education: in 2002, females comprised only 44% of tertiary students.

Starting in the mid-1980s, the sharp decline in the number of rural school-aged children began to threaten the sustainability of the broad-coverage model. This transition was induced by burgeoning rural-urban migration and a fertility decline. The birth control policies began to take effect when the size of Chinese birth cohorts

[5] Local governments typically determined that primary schools should be located no more than 2.5 kilometers from villages (Mei et al., 2015).

decreased from 27.1 million in 1970 to about 22.0 million in 1985, reducing the number of rural school-aged children. In 1989, China has 8.9 million internal migrants, and in 1994, that number increased to 23.0 million (Sicular and Zhao, 2004). The absolute number of enrolled students has decreased since 1995.

As cohorts of rural school-aged children declined, local governments viewed school consolidation as an important means of reducing government education expenditures and increasing education efficiency. The school consolidation program was first implemented by some local governments in the late 1980s and early 1990s, and it was officially implemented nationwide after China’s State Council issued the "Decisions on the Reform and Development of Compulsory Education" in 2001.

The program aimed to exploit economies of scale and improve the quality of rural basic education through school consolidation. The initial targets of consolidation were villages with small populations, which also tended to be remote. Students were assigned to a new school when their local school closed. Simultaneously, schools that were the locus of consolidation received additional resources. Treated students were encouraged to go to a designated school, usually located in a nearby town that had a larger population and better labor market proximity. But students whose families could pay extra money were able to attend other schools. As rural primary schools closed, the village-based education system was gradually transformed into a town-based one.

As Figure 1 shows, between 1987 and 2013, the number of primary schools continuously and significantly declined. During this period, more than 80% of rural primary schools disappeared. This decline occurred primarily in primary schools located in remote, rural communities and had only a limited impact on township and urban communities. The numbers of urban and township primary schools fluctuated only moderately, and they even increased slightly beyond their 1987 total. Note that the decrease in rural primary schools started many years before 2001, when school consolidations were more likely to be incentivized by the education efficiency needs of local communities rather than the later promotion of the central government.

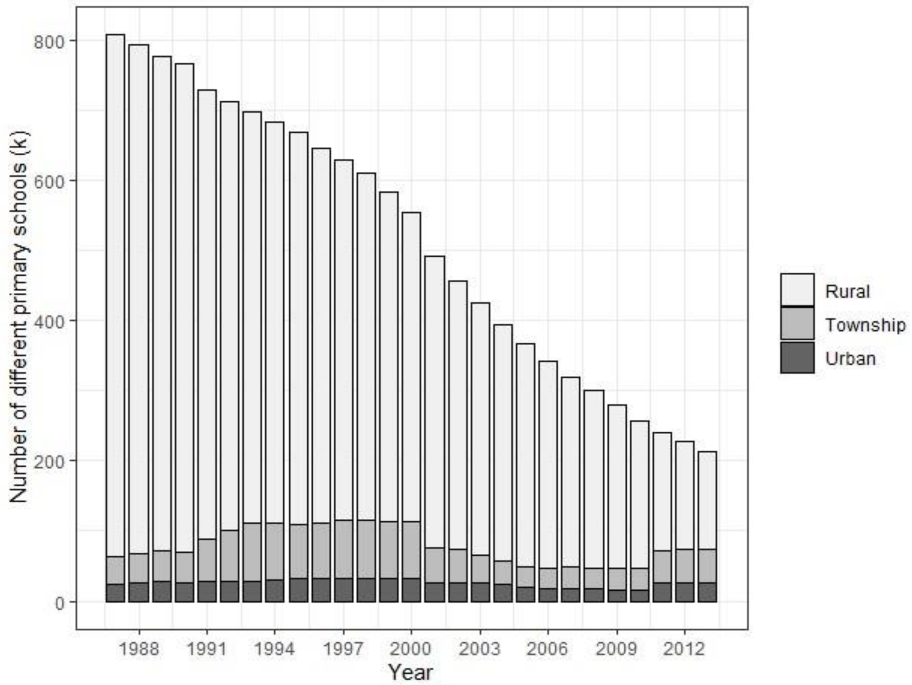

Figure 1: Composition of Primary Schools in China 1987-2013

Sources: Ministry of Education.

The decrease in rural primary schools was accompanied by a reduction in rural students of a smaller magnitude. As shown in Table 1, between 1987 and 2013, the number of rural primary schools decreased by 81.1% (603.7 thousand), while the number of rural primary students decreased by 69.3% (72.4 million). In contrast, by 2013, the number of the township and urban primary students had more than doubled. In 1987, rural primary schools accommodated more than 81.5% of all primary school students in China. By 2013, after the program was formally terminated, there was almost an equal number of primary schools in rural, township, and urban communities. All of these schools now had more students and larger scales, but the locus of increase was predominantly concentrated among township and urban primary schools.

Table 1: Changes in the Number of Primary School and Students

	N of students (m)			N of school (k)			N of students per school		
	1987	**2013**	**Change**	**1987**	**2013**	**Change**	**1987**	**2013**	**Change**
Rural	104.6	32.2	-69.3%	744.0	140.3	-81.1%	140.6	229.3	63.0%
Township	12.2	33.7	175.3%	38.7	47.2	21.8%	316.3	714.8	126.0%
Urban	11.5	27.7	141.6%	24.7	26.1	5.4%	464.5	1064.5	129.2%

Sources: Ministry of Education.

An initial goal of the school consolidation program — providing a better education while not sacrificing access — was not always attained in practice. For instance, Lei and Xu (2011) found that the average travel distance from students' homes to schools increased from 3.2 to 8.1 kilometers, while the average time it took students to go to school increased from 26 to 44 minutes. In addition to the cost of commuting, boarding fees, and safety issues limited the opportunities for the rural students to attend better-resourced consolidated schools. That is, school consolidation did not in all cases reduce education inequalities. Yang (2012) finds that despite the better facilities and teachers in consolidated schools, the dropout rate of rural primary students increased from 0.60% in 2008 to 0.88% in 2011.

3. CONCEPTUAL FRAMEWORK

I build a simple model in the spirit of Baland and Robinson (2000) and Bau et al, (2021) to illustrate the impacts of school consolidation on households' educational investment and children's income when they grow up. The optimal level of schooling acquisition is determined by the return to education and opportunity cost. Even if the impacts on education are different between boys and girls, both genders may perform better in the labor market. The heterogeneous treatment effects across villages with different initial market conditions and school qualities can help us identify potential mechanisms.

3.1 Model setup

Assume each household has one child, and there are two periods. Parents only consume in the first period. In period 1, parents have exogenous endowments y_1, and decide the amount of education investment e and child labor $1-e$. Education leads to higher income in the future with the opportunity costs determined by forgone child labor wage w_1 and education cost. The education cost is determined by both the direct cost c_e (i.e., tuition fee, transportation fee), and children's ability α. A higher α means a higher ability which lowers learning cost.

Parents' utility also depends on children's consumption in period 2 with an altruism index δ and a discount factor ρ. The consumption of children in period 2 depends on education e, and the gender g specific adult wage rate for unskilled (low-educated) and skilled (high-educated) workers w_L^g and w_H^g. The skilled wage rate w_H^g is jointly determined by education quality and labor market proximity, while the unskilled wage rate w_L^g is only determined by the labor market proximity. An individual with education e earns $(1-e)w_L^g + ew_H^g$. The utility functions u have the following properties: $u' > 0$ and $u'' \leq 0$. Hence the maximization problem of parents' utility can be written as

$$\max_e U(e) = u\left(C_1^p(y_1, e)\right) + \rho\delta u\left(C_2^c(e)\right) \tag{1}$$

s.t

$$C_1^p = y_1 + (1-e)w_1 - \frac{c_e}{\alpha}e$$
$$C_2^c = (1-e)w_L^g + ew_H^g$$

I make the following assumptions on the relative size of parameters:

Assumption 1: The household's initial endowment is larger than the full education cost, $y_1 > c_e/\alpha$.

Assumption 2: The opportunity cost of education dominates the effective direct cost of schooling, $w_1 > c_e/\alpha$.

The first-order condition with respect to e is:

$$u'(C_1^p)\left(-w_1 - \frac{c_e}{\alpha}\right) + \rho\delta u'(C_2^c)(w_H - w_L) = 0$$

Consider the case when the utility function takes the log form, $u(x) = \log x$, then the optimal education investment for a household is

$$e^* = \frac{\rho\delta}{1+\rho\delta} \cdot \frac{\alpha(y_1 + w_1)}{\alpha w_1 + c_e} - \frac{1}{1+\rho\delta} \cdot \frac{1}{\frac{w_H^g}{w_L^g} - 1} \tag{2}$$

The subsequent earnings of the child when it becomes an adult in period 2 is

$$y_c{}^* = (1 - e^*)w_L + e^* w_H = \frac{\rho\delta}{1+\rho\delta}\left[w_L + \frac{\alpha(y_1 + w_1)}{\alpha w_1 + c_e}(w_H - w_L)\right] \tag{3}$$

Equation (2) can be interpreted as follows. The education investment increases in household initial endowment y_1, children's ability α, discount factor ρ, altruism index δ, and the high-educated wage w_H. Education decreases in education costs c_e and low-educated wage w_L. The impact of the child labor wage w_1 depends on the relative size of y_1 and c_e/α. Since I assume $y_1 > c_e/\alpha$, an increase in w_1 decreases education investment, as the opportunity cost of education goes up.

Equation (2) also gives an explanation for the gender difference in education investment in rural China. If girls are believed to be more suitable for child labor than boys, either because of social norms or because girls physically mature faster than boys, then they will receive less education and do more child labor. In addition, if parents care less about girls because they will leave the household after marriage, δ is smaller, and girls will have less education.

3.2 The impacts of school consolidations

From equation (2), children's education will decrease in c_e and increase in $\frac{w_H^g}{w_L^g}$. School consolidation changes these two parameters at the same time: after school consolidation, households pay higher education costs due to longer distances and additional boarding fees, hence c_e increases; the new school has better education quality, hence w_H increases; the new school locates in a more populated place and more prosperous labor market, hence both w_H^g and w_L^gincrease, the change in $\frac{w_H^g}{w_L^g}$ can go either direction.[6] If the change in w_H^g dominates changes in c_e, and w_L^g, households invest more in education. Otherwise, they will invest less. In addition, school consolidations may increase labor market outcomes even if the educational attainment does not increase. Reorganize the earnings in period 2 as $w_L^g + e\left(w_H^g - w_L^g\right)$. If e remains unchanged, earnings still increase if w_L^g increases. And it is undetermined from the model whether males or females will benefit more in the labor market. Although the overall impacts of school consolidation depend on the relative importance of the changes in c_e, and $\frac{w_H^g}{w_L^g}$, this simple model generates useful implications.

***Proposition 1*: An increase in c_e has a larger negative impact on children with lower ability α. This result comes from the following expression:**

$$\frac{\partial^2 e^*}{\partial c_e \partial \alpha} = \frac{\rho\delta}{1+\rho\delta} \cdot \frac{(y_1 + w_1)(\alpha w_1 - c_e)}{(\alpha w_1 + c_e)^3} \tag{4}$$

Equation (4) is positive if $w_1 > c_e/\alpha$. Since low ability children have lower education investment before the school consolidation, it implies that there could be more low grades dropout rate after school consolidation.

3.3 Discussion of potential mechanisms

3.3.1 Labor market proximity and education quality

This conceptual framework suggests that a higher $\frac{w_H^g}{w_L^g}$ has a positive impact on education and children's adult earnings. The model allows both better labor market proximity and better education quality to contribute to the increase of $\frac{w_H^g}{w_L^g}$. However, the heterogeneous treatment effects across villages may help me to pin down which mechanism might be more plausible.

The school consolidation targets remote and less populated villages. Those targeted villages usually have limited labor market proximity and poor education quality, but

[6] Since the overall dropout rate is low, which I would provide empirical evidence later, I mainly consider the case that students attended the new schools when discussing the change in parameter to simplify illustration. For households decided not to send their children to the new schools, w_H and w_L remain unchanged because they stayed in the village, and the educational attainment would decrease due to a higher education cost c_e.

the two dimensions are not perfectly correlated. School consolidations level up both factors, as students attend the new schools enjoy the same resources. If the improvement of labor market proximity is dominant to a higher wage rate, individuals from villages with worse ex-ante labor market proximity would earn more after school consolidation. If the improvement of education quality dominates, individuals from villages with worse ex-ante education quality would earn more. Hence, I can test the two potential mechanisms by the heterogeneous treatment effects across villages.

3.3.2 Reduction in gender inequalities

Whether females may benefit from school consolidation or are vulnerable to the increased education cost is an open question. From the above model, it depends on whether females could enjoy the benefit of higher $\frac{w_H^g}{w_L^g}$ after the consolidation. I argue that if the labor market in town in more gender-equalized for high-educated worker, females would have more education compared with males after school consolidation, which can be referred to as the mechanism of reduction in gender inequalities.

***Proposition 2*: Females would have larger increase in education than man, if skill premium change is larger for female after consolidation, i.e., $\frac{\Delta w_H^f}{\Delta w_L^f} > \frac{\Delta w_H^m}{\Delta w_L^m}$.**

$$\frac{\partial e^*}{\partial \frac{w_H^g}{w_L^g}} = \frac{1}{1+\rho\delta} \cdot \frac{1}{\left(\frac{w_H^g}{w_L^g} - 1\right)^2} \tag{5}$$

Equation (5) is positive, hence whether females would have larger increase in education depends on the relationship between $\frac{\Delta w_H^f}{\Delta w_L^f}$ and $\frac{\Delta w_H^m}{\Delta w_L^m}$. The new labor market in town provides more opportunities for both unskilled workers and skilled workers. But the unskilled labor market is less favorable for females due to physical requirements. For example, construction and some manufacturing industries were labor-intensive and prefer male workers, while the office staff and technical worker are less discriminated against females. Hence, w_H and w_L both increase for males, while w_H increases more than w_L for females. I will test this relative change in gender specific wage rate later. Ceteris paribus, *Proposition 2* implies that the second component on the right-hand side of equation (2), $-\left(\frac{1}{1+\rho\delta} \cdot \frac{w_L^g}{w_H^g - w_L^g}\right)$, will be less negative for females. If the change in $\frac{w_H^g}{w_L^g}$ dominates the change in c_e, females would gain more education because there is less gender discrimination in the skilled labor market. In that case, there would be a reduction in gender inequalities because girls would receive more human capital investment and do less child work, which has long-term impacts on females' earnings and women's empowerment when they grow up.

4. DATA

The data set of this paper comes from the China Health and Nutrition Survey (CHNS) and dates from 1989 to 2015. CHNS draws a sample of about 7,200 households composed of more than 30,000 individuals living in 15 provinces and municipal cities that vary substantially in geography, economic development, public resources, and health indicators. CHNS is an unbalanced panel, but for this study I treat it as repeated cross section data.

Figure 2 shows a map of survey provinces with the sample size and time of entry for each province. Counties in the provinces were stratified by income, and a weighted sampling scheme was used to randomly select four counties in each province. In addition, the provincial capital and a lower-income city were selected when feasible. Villages and townships within the counties and urban and suburban neighborhoods within the cities were selected randomly. The average time gap between the two waves of the survey is 2.6 years.

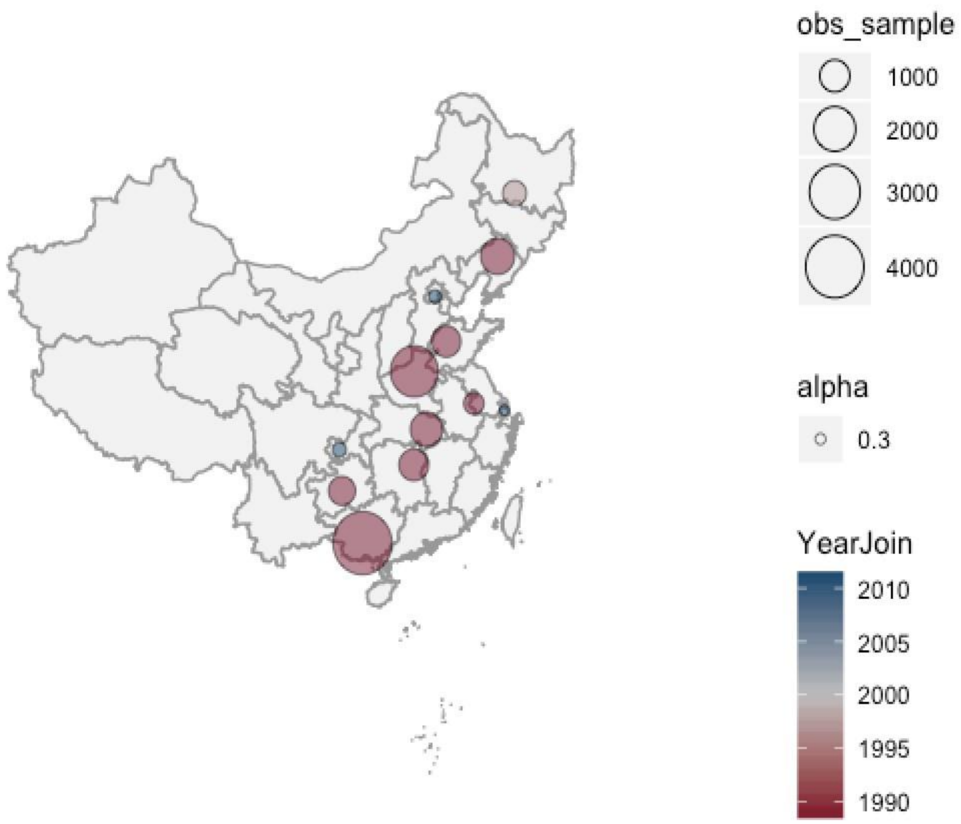

Figure 2: CHNS Survey Provinces and Sample Size

Sources: China Health and Nutrition Survey.

Although CHNS is not based on a nationally representative sample, it provides wide coverage of the population, which is essential for nationwide policy analysis. The sampled provinces in CHNS constituted about 57.3% of the total population in China in the 2010 census. The sample area covers both highly developed coastal areas and less developed middle-west areas, as well as minority autonomous regions, which vary substantially in geography, economic development, and public resources. The survey also tracked individuals who migrated to work but were still registered in their villages. Before 2011, the average attrition rate in each wave of the survey was about 11.2% (Liang, 2011).

This paper utilizes data from the rural sample directly influenced by the school consolidation program. I denote the year of school closure as *t*, if there is a local primary school in the previous survey year, but it disappears in the survey year *t*. Figure 3 shows the distribution of school closure years in my sample. On the basis of these data, I defined communities that once had a local primary school that was later closed as treated villages, and I defined communities that had at least one local primary school throughout the sample period as untreated villages.[7]

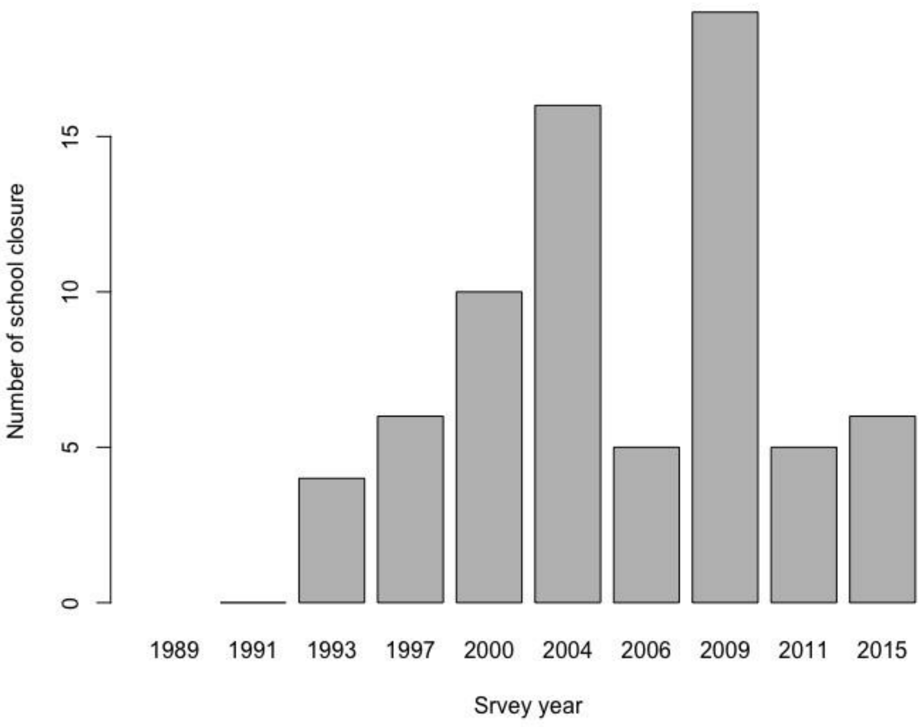

Figure 3: Number of school closures observed each year

Sources: China Health and Nutrition Survey.

As shown in Figure 3, school consolidation took place in the CHNS survey sample no later than 1993, roughly when the Chinese local school consolidation program began. The CHNS covers 181 rural communities (villages), and around 40% of them closed

[7] Communities that had no local primary school throughout the sample period are dropped to alleviate the concerns of bias due to heterogeneous treatment effects in a two-way fixed effect estimation, which I will discuss later.

local primary schools during the sample period. The largest number of school consolidations was documented in 2009, when local primary schools in 19 villages closed. Because the CHNS surveys were not carried out every year during the study period, I also calculated the average number of school closures during multiple survey years. Between 1989-2015, on average 2.73 villages closed their local primary school each year. Figure 3 suggests that the implementation of the school consolidation program accelerated until 2009. After 2012, when the central government ended the consolidation program, the annual number of closed rural primary schools went down.

The survey also collected rich demographic and social-economic information at the individual levels. Table 2 shows the summary statistics in my sample. This paper focuses on individuals resided in the rural area who were born after 1966. Those individuals were less likely than older cohorts to be influenced by the Cultural Revolution. They would have been of primary school age, and rural primary schools were less interrupted than other schools by the political movement. My sample consists of 17,835 individual observations. The average age was 17.09, 41% of them were still in school by the time of the survey. Among those who had already finished their education, the average education was 8.53 years, and the labor force participation rate was 85%. And 3% of individuals between age 6 and 15 in my sample were child laborers. Compared with their male counterparts, females had less education and a lower labor force participation rate. They worked fewer hours than males, earned less, and were also more likely to be child laborers in my sample.

Table 2: Summary statistics for individual characteristics

Variables	N	(1) Full sample	(2) Females	(3) Males
Age	17,835	17.09	17.82	16.45
		(11.34)	(11.48)	(11.17)
Whether the individual is still in school	14,515	0.41	0.39	0.43
		(0.492)	(0.488)	(0.494)
Education for individuals finished schooling	7,831	8.53	8.41	8.64
		(3.209)	(3.274)	(3.144)
Labor force participation	8,863	0.85	0.79	0.90
		(0.357)	(0.404)	(0.296)
Log of working hours	3,546	7.26	7.22	7.30
		(0.917)	(0.949)	(0.884)
Log of income	6,252	8.58	8.48	8.67
		(1.630)	(1.579)	(1.671)
Log of wage	3,255	8.90	8.82	8.97
		(1.428)	(1.391)	(1.455)
Child labor dummy	5,758	0.03	0.04	0.02
		(0.169)	(0.187)	(0.152)

Note: Sample includes individuals resided in rural areas who were born after 1966.

Figure 4 shows the age distribution of the treatment group and control group 1 (all individuals in villages that did not undergo consolidation) and control group 2 (individuals in treated villages who were older than 13 years old when local schools closed). The red bar represents the treated individuals, the green bar control group 1, and the blue bar control group 2. By construction, the average age of the treated group is the youngest of the three, while the average age of control group 2 is the oldest. Putting the red and blue bar together, the overall age distribution in the treated villages looks similar to the control villages.

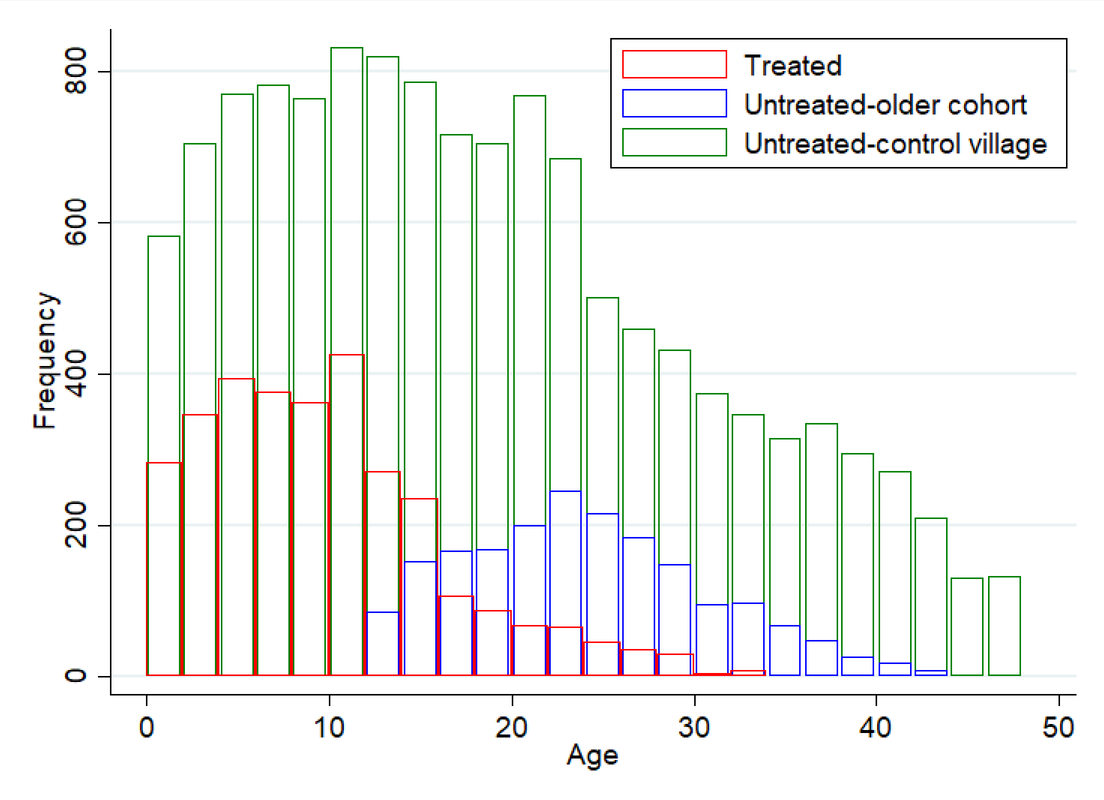

Figure 4: Age distribution for treatment group and control groups

Sources: China Health and Nutrition Survey.

5. IDENTIFICATION STRATEGY

5.1 Generalized DID and the event study

The identification strategy employed in this paper follows Duflo (2001). The basic idea behind the identification strategy is a classic DID practice. The date of birth and the region of birth jointly determine an individual's exposure to the program of school consolidation.

Chinese students normally attend primary school starting at age 7 and graduate at age 13. Individuals are divided into those who live in a village with school consolidation, and those who live in a village without consolidation as defined in section 4. All students who live in a village without school consolidation form the first part of the control group. For those in a village with consolidation, suppose a local primary school was closed in year *t* and that all children born in year (*t-13*) or before were 13 or older by the time of closure and, thus, were not exposed to the school consolidation program. I define children born in year (*t-13*) or before as the old cohort and children born after year (*t-13*) as the young cohort. The young cohort is the treatment group, while the old cohort forms the second part of the control group.

To study the impact of school consolidation programs in rural China, I compare the education and labor market outcomes of individuals from the treatment and control groups. The difference can be interpreted as the causal effect of the program, given two assumptions. First, in the absence of the program, there would have been a parallel education trend between the treatment group and the control group. Second, there was no spillover effect—that is, education resources and students were not disproportionally reallocated after school consolidation.[8]

The identification assumption should not be taken for granted because the pattern of educational attainment could vary systematically across villages. However, I can provide suggestive evidence for an implication of the identification assumption because individuals aged 13 or older at the time of local primary school closure were not exposed to school consolidation. The change in educational attainment between cohorts in this age group should not differ systematically across regions. Thus, I can employ the old cohorts as placebo treatment groups to account for the potential systematic difference between treated and untreated villages. The different treatment year across villages also alleviates concerns about the violation of parallel trends.

Thus, I run the following regression to specify the impact of school consolidation:

$$Y_{ijct} = \alpha + \beta_1 D_{ijc} + \beta_\tau X_i + \theta_j + \mu_c + \omega_{pt} + \varepsilon_{ijct} \tag{6}$$

where Y_{ijct} is educational attainment or labor market outcomes in survey year t of individual i from community j born in cohort c. D_{ijc} is a dummy indicating whether the individual belongs to the treatment group (less than 13 years old at local school closure), X_i is a set of individual and household controls, θ_j is a village fixed effect, μ_c is a birth cohort fixed effect, ω_{pt} is the province-specific survey year fixed effect, and ε_{ijct} is the unobserved error term. The parameter of interest is β_1, which captures the average treatment effect of school consolidation. Supposing the impact of an increase in commuting cost dominates the better qualities or labor market proximities in consolidated schools, β_1 should be negative. Otherwise, β_1 should be positive.

I can further break down my sample and estimate equation (6) in every age-at-closure group. Consider the new equation:

$$Y_{ijct} = \alpha + \sum_{l=0}^{21} \beta_{1l}\left(D_{ijc} \cdot D_{il}\right) + \beta_\tau X_i + \theta_j + \mu_c + \omega_{pt} + \varepsilon_{ijct} \tag{7}$$

where D_{il} is a dummy that indicates whether individual i is age l at primary school

[8] My identification would not be threatened even if the level of education and other village characteristics were different in villages with and without school closure.

closure. β_{1l} can be interpreted as an estimate of the impact of school consolidation on a given age-at-closure group. Regression (7) can be viewed as a generalization of the classic test for parallel trends. As established by Duflo (2001), there should be a testable restriction on the pattern of β_{1l}. Since children aged 13 and older at school closure were not impacted by the program, β_{1l} should be zero for $l \geq 13$, and it should be significantly different from zero for $l < 13$.

In addition to the average treatment effect of school consolidation, this paper examines heterogeneous treatment effects. To test for the heterogeneous treatment effect of school consolidation across villages, I run the following regression:

$$Y_{ijct} = \alpha + \beta_1 D_{ijc} + \beta_2 C_i + \beta_3 \left(D_{ijc} \cdot C_i\right) + \beta_\tau X_i + \theta_j + \mu_c + \omega_{pt} + \varepsilon_{ijct} \qquad (8)$$

where C_i is village-level characteristics. For example, C_i could be a dummy that takes 1 if the village is within two kilometers of an economic zone. Or it can be a continuous measurement of labor market proximities and teacher's capability. $D_{ijc} \cdot C_i$ is the intersection term. β_3 is the coefficient of triple difference. If there are heterogeneous treatment effects of school consolidation across villages, β_3 would be significantly different from zero.

5.2 Robustness to the heterogeneous treatment effects

It is worth noting that the generalized DID in equation (1), which is also called the two-way fixed effect model, can produce misleading estimates (de Chaisemartin and D'Haultfoeuille, 2022; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Roth and Sant'Anna, 2021; Sun and Abraham, 2021; Cengiz et al., 2019). The two-way fixed effect is essentially a weighted average of treatment effects across groups and over time. If there are heterogeneous treatment effects, and if some of the weights are negative, the estimated average treatment effect for the treated (ATT) would be biased.

I consider three approaches to alleviate the concerns of bias due to heterogeneous treatment effects. First, de Chaisemartin and D' Haultfoeuille (2022) show that with a binary treatment, the weights on ATTs are likely to be positive when there is no always-treated group and no time periods where most groups are treated. Hence, I dropped individuals from villages that never had local primary schools throughout the sample period, to mitigate the negative weights. By doing so, the estimator satisfies the "no-sign-reversal" property, with which the estimand attached to it can only be positive if the treatment is not Pareto dominated by the absence of treatment.

Second, Cengiz et al. (2019) propose an alternative estimate that is not subject to this problem, which could be applied to cross-section data in which cohorts of birth play the role of time. I first create datasets for each specific local school closure, then stack all of the event-specific data to calculate an average effect across all events using a single set of treatment effects. By aligning events by event time rather than calendar time, it is equivalent to a setting where the events happen all at once. Compared with my baseline specification, it uses more stringent criteria for admissible control groups and prevents the negative weighting of some events. Hence, it is more robust to possible problems caused by heterogeneous treatment effects in a staggered DID design.

Third, I calculate the weights attached to the two-way fixed effects regressions using the method proposed by de Chaisemartin, D'Haultfoeuille, and Deeb (2020). If many weights are negative, and if the ratio of $|\beta_3|$ divided by the standard deviation of the weights is not very large, there is a concern that the two-way fixed effect estimator is biased. On the other hand, if most weights are positive and the sum of negative weights is low, then the heterogeneous treatment effects are unlikely to be problematic in this context.

6. EMPIRICAL RESULTS

This section introduces the impacts of school consolidations on education and earnings. I show that following school consolidation, new schools were located further from students' homes but had better quality. The average treatment effects of school consolidation on education are insignificant. However, the effects mask large heterogeneities across genders: the education of females increased significantly, while the impacts on males were insignificant and negative. Pooling females and males together, treated individuals of both genders subsequently performed better in the labor market. They had higher labor force participation rates and more working hours, and they enjoyed higher wages and incomes. Those findings are consistent with the implications of my conceptual framework.

6.1 Longer distances to schools and better school quality

School consolidations increased the commuting costs of rural primary school students. The CHNS survey asked for the distance to the nearest primary school if there was no local school in the village. If there was a local primary school in the village, I denoted the distance as zero. CHNS also asked students for information about their transportation to school. Table 3 presents the results of regressions that take the same form as equation (1), while replacing the Y_{ijct} by distance, transportation method, time spent walking to school and on homework, and sports in school. Only individuals who had not finished their primary school education at the time of the survey are considered. Columns 1-3 represent the increase in education costs after school consolidation. After the local primary school closure, the average distance to the nearest rural primary school increased by 1.57 km. As the distance increased, treated students were 11.3 percentage points less likely to walk to school, and those who kept walking spent an additional 22 minutes every day on their way to school.

At the same time, school consolidations brought better school quality. CHNS did not directly ask for information about school quality. Columns 4 and 5 use two measurements as imperfect proxies of school qualities. First is the time students spent on homework; I assume that schools that had better education qualities taught more, and, thus, their students needed to spend more time on homework.[9] The second measure is the number of ball sports, such as soccer, basketball, volleyball, tennis, and badminton that students do in school. Schools that offered more ball sports are assumed to have had better facilities than those that offered fewer sports. After school consolidation, treated students spent 344.1 minutes more on homework per week. In addition, they played more ball sports in school. Table 3 is consistent with the story that after school consolidation, households faced a trade-off between higher commuting costs and better school quality.

Table 3: Higher Cost and Better School Quality after Consolidations

	(1)	(2)	(3)	(4)	(5)
	Education cost			School quality	
	Distance to school	None walk travel	Time on walking	Time on homework	Number of ball sports
School consolidation	1.572***	0.113*	22.25*	344.1**	0.370**
	(0.208)	(0.0605)	(12.72)	(136.7)	(0.153)
Cohort FE + Village FE	√	√	√	√	√
Province*Year FE	√	√	√	√	√
Dependent Variable Mean	0.368	0.262	41.07	475.4	0.431
Observations	2,528	2,095	1,546	1,565	1,721
R-squared	0.889	0.478	0.157	0.243	0.288

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level. The dependent variables in columns 1 to 5 are the distance to the nearest primary school, whether students chose to walk to school, the time spent on walking (daily) and homework (weekly), and the number of equipped sports students did in school (soccer, basketball, volleyball, tennis, and badminton). The regression sample includes individuals born after 1966 who were older than 6 and were still in school at the time of the survey. School consolidation is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village; otherwise it is 0.

6.2 The impacts on education

6.2.1 Treatment effects on educational attainment

Table 4 shows the impacts of school consolidation on education. Columns 1-3 show the result of the two-way fixed effects specification that controls cohort fixed effects and village fixed effects. Females on average had significantly less education than males, and the average treatment effect of school consolidation on education was small and insignificant. However, there were large heterogeneous effects across genders. The coefficient of school consolidation was significantly positive for females

[9] It is common in China for students who attend better schools to spend more time on homework, especially in rural areas, where students face fierce pressure to enter better schools.

but insignificant for males. Columns 4-6 add a minority dummy and add province times year fixed effects to flexibly control provincial-level time-varying shocks. The results are consistent with columns 1-3. Specifically, females gained 0.709 more years of education after school consolidation, but the treatment effect on males was negative and insignificant. In addition, the difference in school consolidation coefficients was 0.961 years, and the difference in average education between males and females was 0.394 in column 4, suggesting that the treatment effects reflect more than just a catch-up effect across genders. I take columns 4 to 6 as the baseline specification of this paper.

Table 4: Females Get More Education after Consolidations

	(1) All	(2) Female	(3) Male	(4) All	(5) Female	(6) Male
School Consolidation	0.195	0.622**	-0.179	0.192	0.709***	-0.252
	(0.251)	(0.263)	(0.324)	(0.234)	(0.268)	(0.298)
Female	-0.350***			-0.394***		
	(0.0899)			(0.0887)		
Minority				0.225	0.527**	0.0551
				(0.199)	(0.252)	(0.245)
Cohort FE + Village FE	√	√	√	√	√	√
Province*Year FE	-	-	-	√	√	√
Dependent Variable Mean	8.928	8.777	9.066	8.928	8.777	9.066
Observations	7,247	3,456	3,785	7,218	3,443	3,766
R-squared	0.358	0.403	0.364	0.378	0.424	0.391

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level . The dependent variable is educational attainment measured by years of education completed. The regression sample includes individuals born after 1966 who were older than 15 and finished schooling at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village, and otherwise is 0. Minority is a dummy that takes 1 if the individual does not belong to the Han ethnic group.

Figure 5 looks at the treatment effects in smaller sub-groups by showing the event study of school consolidation's impacts on education across genders. Panel A and panel B plot the β_{il} and 95% confidence intervals for sub-samples of females and males, respectively. I divide individuals in villages that have local primary school closures into roughly 3-year birth cohort windows. For example, the *13-15* group consists of individuals who were 13-15 years old when the school closed. Taking individuals aged 22 and 23 at school closure as the reference group, the coefficients are small and fluctuate around zero for individuals older than 13, become positive for females younger than 13 at school closure, and are negative for males. I only use these results as suggestive evidence because most of the coefficients are marginally insignificant. But the sign of these coefficients is consistent with our findings in Table 5. As expected, consolidation did not affect the education of cohorts not exposed to it, and after school consolidation, the treatment effects were positive for females only. However, we should bear in mind that although parallel trends in the period before treatment suggest counterfactual parallel trends, parallel pre-trends were neither necessary nor sufficient for the parallel counterfactual trends condition to hold (Kahn-Lang and Lang, 2019).

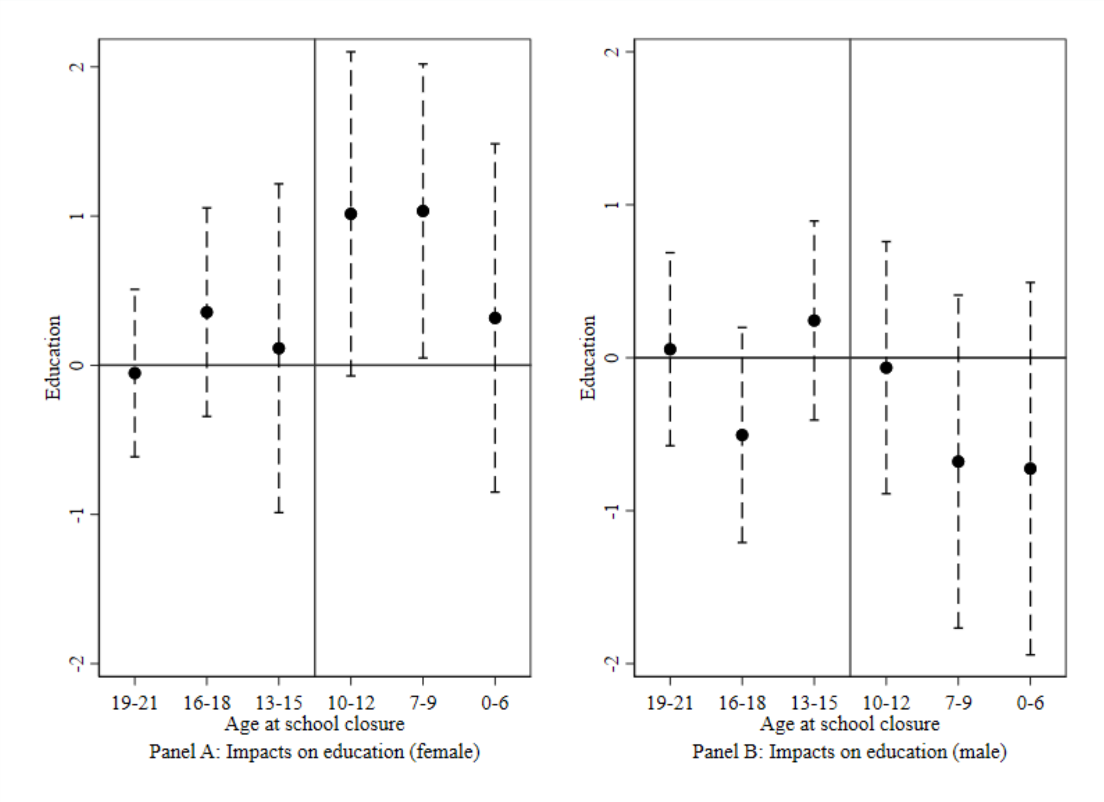

Figure 5: Heterogeneous Treatment Effects on Education across Genders

Notes: The figure shows the effect of school consolidation on educational attainment in my baseline DID specification by six age-at-closure groups. Panels A and B are female and male samples, respectively, born after 1966 who were older than 15 and finished schooling at the time of the survey. Each dot represents the impact on lifetime educational attainment for each age group defined at the time of local school closure. The confidence interval is at the 95% level. The reference group are individuals aged 22 and 23 by the time of local school closures.

6.2.2 Dropouts and delaying enrollment

Although the school consolidations did not harm average lifetime education, it might have had two negative impacts on schooling. First, children might have delayed enrollment if parents had safety concerns or if parents were unsure about the benefits of school consolidation. Second, there might have been more low-grade dropouts if the increase in education costs had larger impacts on children with low abilities.

Table 5 column 1 shows that individuals exposed to school consolidations before age 9 were 1.5 percentage points more likely to drop out before the third grade. This increase is salient because the average share of early dropouts is only 1.8%. Columns 2 and 3 conduct the same specification except that the dependent variables are the probability of dropping out before the 6th and 9th grades. The significant coefficient in column 1 and the insignificant coefficients in columns 2 and 3 suggest that school consolidations did not bring additional dropouts beyond the 6th and 9th grades. It is likely that students who would drop out without school consolidations now dropped out at younger ages. Therefore, although school consolidations did not harm the average educational attainment, it might enlarge inequalities, because individuals with the least education were more likely to be worse off. This result is consistent with the implication of my model that an increase in education cost has a larger negative impact on children with lower ability.

To test for the short-run effect of delaying enrollment I need information on the age when individuals entered primary school. For individuals still in school, I can use their year of educational attainment and current age to back out the age when they first entered primary school. If a student was age A at the time of the survey and had n years of schooling, we denote its entering age as (A-n). The underlying assumption is that the grade repetition rate is low.[10] Local school closures would only impact the enrollment age of individuals who had not yet started primary school; thus, in the treated village, at the time of school closure, the treated group in this specification should be those younger than age 6.[11] In column 4 of Table 5, I examine the impacts on (A-n) for individuals who are in school. This allows me to look at delays in enrollment age while excluding the effects of dropping out before they ever go to school. Column 4 shows that after school consolidation, treated individuals delayed their enrollment by 0.202 years.

Table 5: Students are More Likely to Drop out Early and Delay Enrollments

	(1)	(2)	(3)	(4)
	< 3rd grade dropout	< 6th grade dropout	< 9th grade dropout	Age at enrolment
School consolidation before age 9	0.0150**	-0.0141	0.0313	
	(0.00703)	(0.0223)	(0.0569)	
School consolidation before age 6				0.202**
				(0.0942)
Cohort FE + Village FE	√	√	√	√
Province*Year FE	√	√	√	√
Dependent Variable Mean	0.018	0.177	0.717	7.021
Observations	9,754	8,095	6,779	4,277
R-squared	0.092	0.194	0.331	0.337

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level . The dependent variables in columns 1 to 3 are the dummy of whether individuals drop out before the 3rd, 6th, 9th grade, and the regression samples are individuals born after 1966 who at the time of the survey were older than 12, 15, and 18, respectively. The dependent variable in column 4 is their estimated age at primary school enrollment, and the regression sample includes individuals who born after 1966 who were still in school by the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a

[10] Governments placed a restriction on repetition rate. For example, the Gansu province required that the repetition rate in primary school be no more than 2%.

[11] One way to test this delaying enrollment effect is to test the probability of being enrolled by given ages (e.g., the probability of being enrolled by age 6, age 7, age 8, etc.). But in the case where the probability of being enrolled by 7 decreases after school consolidation, we cannot determine whether this reflects a delay in enrollment or students dropping out before they enroll in school.

primary school closure in the village, and otherwise is 0.

6.2.3 Robustness checks

I did four sets of robustness checks.

(1) In Table 6 columns 1 and 2, I used a new sample of individuals older than 20 at the time of the survey. The new sample uses a natural cutoff rather than an endogenous threshold (whether individuals were still in school or not by the time of survey), and it is less concerned with sample selection issues. However, there are two limitations: first, the sample size is smaller because individuals out of school but younger than 20 are not included; second, some individuals had not finished schooling yet at 20, and they are less comparable with full-time workers. The coefficient of school consolidations in the female sample was 0.995, which was larger and more significant than the baseline specification.

(2) In columns 3 and 4, I excluded the communities in towns to alleviate concerns of potential spillover effects. After school consolidations, more students and resources were allocated to the new schools. If the increase in students outweighed the increase in teachers, the control group might be negatively affected due to an increase in the student-teacher ratio. Such spillover effects can cause a positive bias in estimates. On the other hand, if the increase in teachers outweighed the increase in students, it could cause a negative bias in estimates. To address this issue, I excluded the communities in towns, which are more likely to be affected by the reallocation of students and teachers. The magnitudes of school consolidations' effect are similar to those in columns 1 and 2.

(3) In columns 5 and 6, I used an alternative estimator employed by Cengiz et al. (2019) that is robust to the heterogeneous treatment effects. I first created event-specific datasets and then stacked all of the event-specific data to calculate an average effect across all events using a single set of treatment effects. Due to this construction process, the sample size is much larger than my baseline specification. The magnitudes of school consolidations are close to the baseline specification, suggesting that issues about negative weighting using staggered treatments were unlikely to be driving my results.

(4) Following de Chaisemartin, D'Haultfoeuille, and Deeb (2020), I calculated the share of negative weights for two-way fixed effects estimations in Table 4 columns 2 and 3. Denote the coefficient of school consolidation for females and males as β_f and β_m, respectively. Under the common trends assumption, β_f estimates a weighted sum of 107 average treatment effects (ATTs), of which 104 (97.2%) ATTs receive a positive weight, and 3 ATTs receive a negative weight. β_m estimates a weighted sum of 105 ATTs, of which 104 (99.0%) receive a positive weight. Because there are only small shares of ATTs that have negative weights, it is likely that my results are robust

to the heterogeneous treatment effects.

Overall, the above practice suggests that the baseline results on education are robust to the heterogeneous treatment effects and the sample choices.

Table 6: Robustness Checks of Impacts on Education

	(1)	(2)	(3)	(4)	(5)	(6)
	Age>20		Age>20 + exclude communities in town		Cengiz et al. (2019) method	
	Female	Male	Female	Male	Female	Male
School consolidation	0.995***	-0.223	0.907**	-0.382	0.655**	-0.291
	(0.342)	(0.395)	(0.435)	(0.535)	(0.264)	(0.299)
Cohort FE + Village FE	√	√	√	√	√	√
Province*Year FE	√	√	√	√	√	√
Dependent Variable Mean	8.974	9.353	8.418	8.835	9.007	9.321
Observations/Switcher	2,888	3,014	2,235	2,373	25,939	28,197
R-squared	0.430	0.416	0.347	0.301	0.426	0.384

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level . The dependent variable is educational attainment. The regression sample in columns 1 and 2 includes females and males born after 1966 who were older than 20 at the time of the survey. Samples in columns 3 and 4 further exclude individuals from communities (villages) in towns. Samples in columns 5 and 6 are females and males born after 1966 who were older than 15 and finished schooling at the time of the survey. The sample sizes in columns 5 and 6 are much larger due to the construction process in Cengiz et al. (2019). Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and that otherwise is 0.

6.3 Impacts on labor market outcomes

6.3.1 Better labor market outcomes for both genders

Table 7 shows the impacts of school consolidations on a series of labor market outcomes. Panel A shows the results of individuals older than 15 years of age at the time of the survey. From both intensive and extensive perspectives, individuals exposed to school consolidation were more likely to be active in the labor market. From the extensive perspective, column 1 shows that treated individuals were 7.0 percentage points more likely to work or actively seek a job in the labor market, which is an 7.9% increase from an 89.0% average labor force participation rate. From the intensive perspective, column 2 shows that treated individuals had 15.8% more annual working hours. Consistent with the increased labor force participation rate and working hours, treated individuals on average earned 29.0% more income and 19.1% higher wages.

Panels B and C look at subsamples of females and males, respectively. In contrast to the gender heterogeneity in education, both treated males and females enjoyed better outcomes in the labor market after consolidation, though the specific channels are different for males and females. Females exposed to school consolidation had a 10.8 percentage points higher labor force participation rate and earned 27.5% higher wages. Treated males had 22.1% more working hours and earned 24.9% higher income. While the coefficients between males and females are not significantly different, the difference in magnitudes suggests that the increase in labor force participation is more likely to be driven by females while the increase in working hours is mainly driven by males.

Table 7: Impacts of School Consolidations on Labor Market Outcomes

	(1) LFP	(2) ln(workhour)	(3) ln(income)	(4) ln(wage)
Panel A: full sample				
School consolidation	0.070**	0.158**	0.290***	0.191**
	(0.0274)	(0.0754)	(0.106)	(0.0840)
Dependent Variable Mean	0.890	7.452	8.723	8.989
Observations	7,698	3,318	5,656	3,020
R-squared	0.181	0.277	0.468	0.666
Panel B: females				
School consolidation	0.108**	0.0671	0.193	0.275*
	(0.0532)	(0.117)	(0.130)	(0.140)
Dependent Variable Mean	0.820	7.436	8.637	8.911
Observations	3,766	1,583	2,702	1,363
R-squared	0.234	0.346	0.483	0.676
Panel C: males				
School consolidation	0.0256	0.221**	0.249*	0.0772
	(0.0226)	(0.105)	(0.132)	(0.107)
Dependent Variable Mean	0.957	7.466	8.802	9.055
Observations	3,925	1,712	2,943	1,634
R-squared	0.106	0.300	0.498	0.705
Cohort FE + Village FE	√	√	√	√
Province-Year FE	√	√	√	√

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level. The dependent variables in columns 1 to 4 are the dummy of whether individuals are in the labor force, the log work hours, log incomes, and log wages, respectively. The regression sample includes individuals who were born after1966, were older than 15 at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and that otherwise is 0. Panel A is the full sample, while panels B and C are subsamples of females and males respectively.

Figure 6 Panel A presents the results of the event study of labor force participation with linear probability models. Taking individuals aged 22 and 23 at school closure as the reference group, the coefficients were negligible and fluctuated around zero for individuals older than 13 and became more positive for individuals younger than 13 at school closure. Although there were some fluctuations of coefficients among treated individuals, the effects generally were larger for younger cohorts. This result reinforces my finding, presented in Table 7 Panel A column 1: after graduation, individuals exposed to school consolidation had more labor force participation. Figure 6 Panel B shows the event study of log working hours. Consistent with the pattern shown in panel A, the coefficients were small for individuals older than 13, and they became larger and positive for most age-at-closure groups that were younger than 13 at school closure.

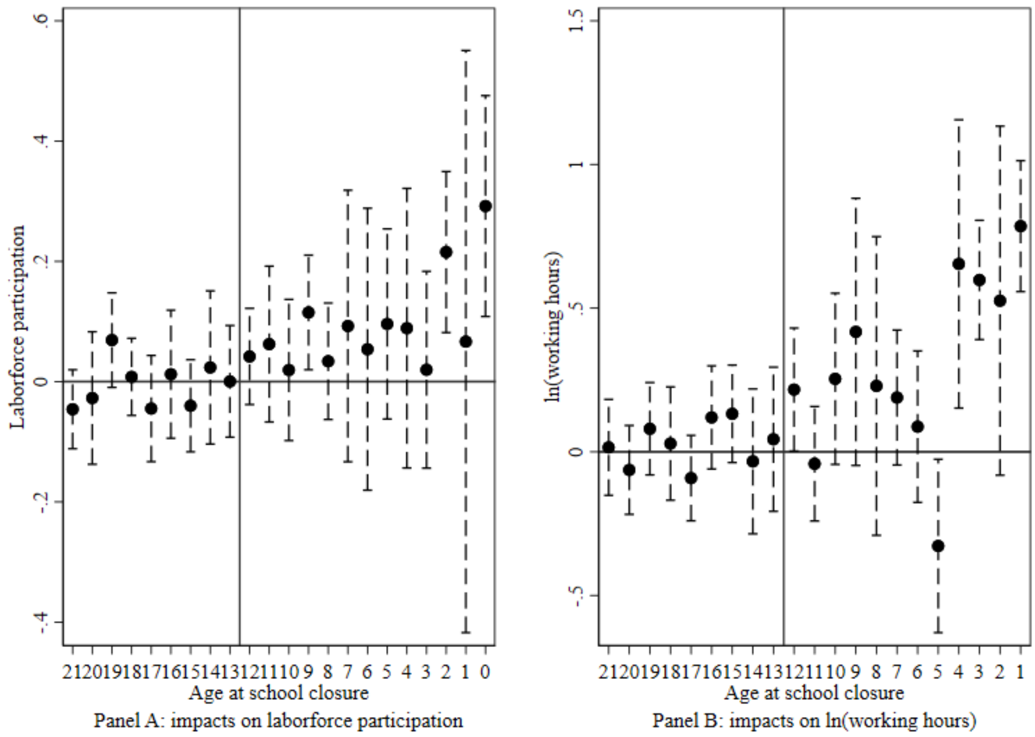

Figure 6: Event Studies of Impacts on Labor Force Participation and Working Hours

Notes: Panels A and B show the effect of school consolidation on the labor force participation rate and working hours by twenty-one age-at-closure groups. The regression sample includes individuals born after 1966 who were older than 15 at the time of the survey. Each dot represents the impact of school consolidation for each age group defined at the time of local school closure. The confidence interval is at the 95% level. The reference group are individuals aged 22 and 23 at the time of local school closures.

Next, I report in Figure 7 the event study results of the effects of school consolidation on log wage earning and log individual income. As shown in panel A and panel B, individuals who were younger than 13 at time of school closure generally had higher wages and incomes (with large fluctuations in some cohorts). The event study results suggest that the results in DID are unlikely to be driven by pre-trends or outliers. The coefficients of individuals equal or younger than age zero (not yet born) at school closure were larger than those of the rest of the treatment group for both log income and log wages. There are two potential explanations for larger and less precisely estimated coefficients of the zero-age group. First, if parents invested more in children's health during pregnancy after consolidation, the zero-age group would have larger treatment effects than those already born at school closure. Second, the zero-age group includes all treated individuals born after the consolidation rather than those born in the specific year. The larger coefficients capture the heterogeneous treatment effects across age groups if the impacts decreased with age at school consolidation and the treated age for the zero-age group is much younger than zero. When I drop observations in the zero-age group, the main results in columns 3 and 4 of Table 7 still hold.

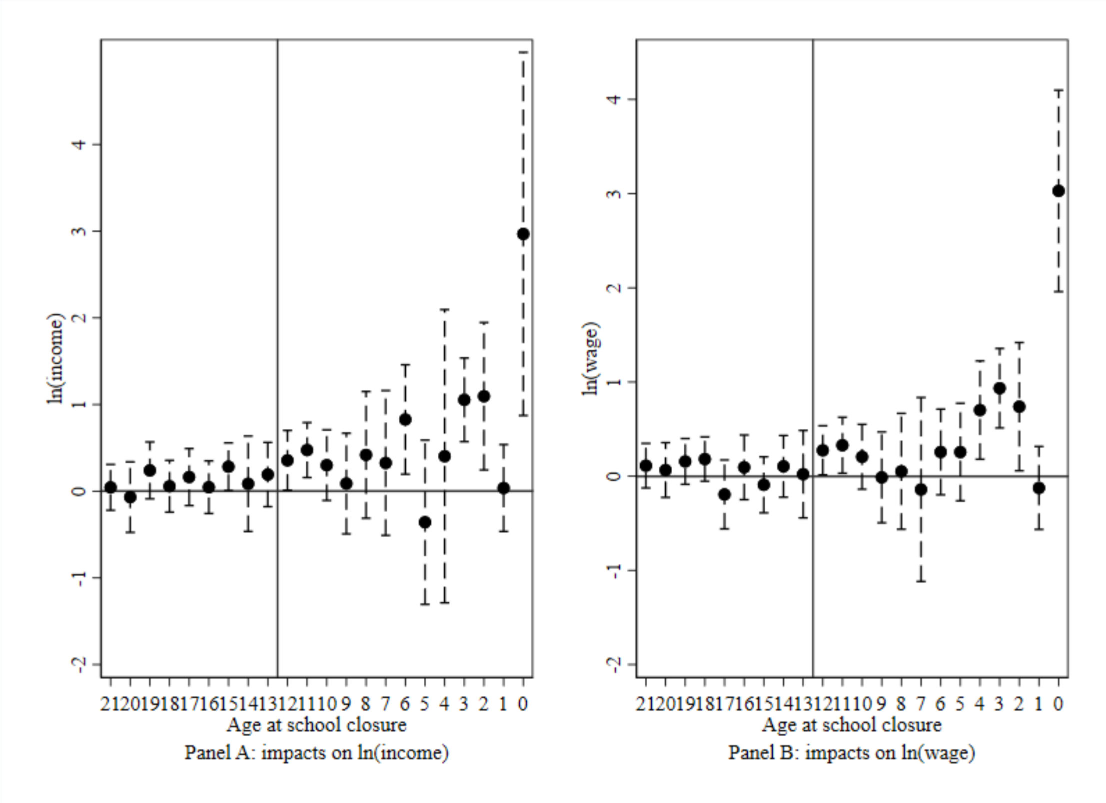

Figure 7: Event Studies of Impacts on Incomes and Wages

Notes: Panels A and B show the effect of school consolidation on log incomes and wages by twenty-one age-at-closure groups. The regression sample includes individuals born after 1966 who were older than 15 at the time of the survey. Each dot represents the impact of school consolidation for each age group defined at the time of local school closure. The confidence interval is at the 95% level. The reference groups are individuals aged 22 and 23 at the time of local school closures.

In Figure 6 and Figure 7, I observe larger positive effects of school consolidation if the treated individuals were younger at the time of school closure. Two factors may explain this pattern. First, the younger-than-7 cohort had not yet begun school when rural schools closed as part of school consolidation, and, thus, they did not experience the disruptions caused by school closure and relocation. Second, the younger-than-7 cohort spent longer periods of time in their new schools, which might have provided better resources than their old schools. The first factor led the younger cohort to enjoy greater educational stability, the second allowed the younger cohort to have longer exposures to superior resources. Both had long-term beneficial effects.

6.3.2 Robustness check

Table 8 shows robustness checks for the impact of school consolidations on labor market outcomes, following the same logic in the robustness checks for education outcomes in Table 6. In panel A, I used a new sample of individuals older than 20 at the time of the survey, and in panel B, I further excluded the communities in the towns. The coefficients in most specifications remained significant and the magnitudes were larger than those in the baseline regression. The only exception was the coefficient on log working hours in panel A, which was smaller and insignificant but remained positive.

Table 8: Robustness Checks of Impacts on Labor Market Outcomes

	(1)	(2)	(3)	(4)
	LFP	ln(workhour)	ln(income)	ln(wage)
Panel A: age>20				
School consolidation	0.0711**	0.103	0.268**	0.291**
	(0.0331)	(0.0757)	(0.113)	(0.116)
Dependent Variable Mean	0.706	7.452	8.979	9.185
Observations	5,994	2,912	4,496	2,548
R-squared	0.201	0.301	0.447	0.629
Panel B: age>20 + exclude communities in town				
School consolidation	0.0757*	0.268**	0.404***	0.413***
	(0.0440)	(0.128)	(0.128)	(0.128)
Dependent Variable Mean	0.624	7.382	8.882	9.147
Observations	4,668	2,002	3,532	1,698
R-squared	0.220	0.316	0.420	0.550
Cohort FE + Village FE	√	√	√	√
Province-Year FE	√	√	√	√

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level. The dependent variables in columns 1 to 4 are the dummy of whether individuals are in the labor force, the log work hours, log incomes, and log wages, respectively. The regression sample of panel A includes individuals born after1966 who were older than 20 years old at the time of the survey. Panel B uses the same sample as panel A but excludes the individuals from communities (villages) in towns. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and that otherwise is 0.

7. MECHANISMS

So far, I have found that school consolidations led to better labor market outcomes by increasing the labor force participation rate, working hours, wages, and incomes for both genders. But only females had more education after consolidations. As discussed in the conceptual framework, better labor market outcomes could be due either to enhanced proximity to the urban labor market or better education quality. In addition, the heterogeneous treatment effects across genders may favor females if there is a reduction in gender inequality due to less discrimination in the high-educated labor market in town.

In this section, I adress two questions: (1) Which is more likely to explain better labor market outcomes: labor market proximity or education quality? (2) Is there suggestive evidence for enduring reductions in gender inequality?

7.1 Enhanced labor market proximity or better education quality

I discuss the heterogeneous treatment effects across villages with different initial market conditions and school qualities in Table 9. Panel A columns 1 and 2 interact school consolidation with a dummy of whether there is an open trade area, an open city, or a special economic zone near the village (within two hours by bus), hereafter called economic zones. Columns 3 and 4 interact school consolidation with teacher (absolute) wages in the village. Teacher wages can largely reflect teacher's capability in the Chinese education system, and I use it as an imperfect proxy of education quality.[12] Columns 5 and 6 pool the two interactions together. School consolidations increased both labor market proximity and teacher capability for treated individuals. Hence, if improved labor market proximity was the main channel for better labor market outcomes, I would see larger positive impacts of school consolidations on villages far from economic zones. If improved teacher capability was the main channel, I would see larger impacts on villages with a lower teacher wage.

Panel A columns 1 and 2 provide evidence that in villages that undergo primary school consolidation, treated individuals from villages far from economic zones gained 35.4% higher incomes and 42.2% higher wages than individuals from villages close to economic zones after school consolidation. But individuals from villages with a low teacher's wage did not earn higher incomes or wages. When including both interactions, the pattern of heterogeneous treatment effects across villages with different labor market proximity still holds.

Panel B uses as an alternative measure of teacher capability the relative wage between the teacher's and the low-skilled male workers' wages. This measure captures the incentive of primary school teachers and addresses the potential concern that the teacher's wage in the public schools was set by the government and lacks variation. I did not find in Panel B columns 3 and 4 significant impacts of the teachers' relative wages on log incomes and log wages. A potential concern with this measure is that the relative wages differ if the low-skilled male worker's wages were different, but the teacher wages were rigid. Hence, the coefficient of interaction term might partly reflect differences in village's labor market conditions.

---

[12] I do not have enough information to directly measure education quality in my CHNS sample. The common inputs to schools are class size, teachers' capabilities, etc. Taking into consideration both institutional factors (how wages are determined—in this case, primarily by professional title) and data availability, I use wages to measure the capabilities of teachers. The wage of a primary school teacher in China is largely determined by his or her professional title. There are six levels of primary school teacher titles. The more educated and experienced teachers usually have higher titles and, thus, higher wages. Therefore, ceteris paribus, villages that have higher teacher wages tend to have more capable teachers.

Panel C uses distance to the railway station as an alternative proxy for labor market proximity. Railway stations are usually material distribution centers that provide many job opportunities. The railway was also one of the main method of transportation for rural residents who wanted to migrate for work. Hence, villages close to railway stations usually had better labor market access. Including both the interaction of dummy of far from railway station and teacher's absolute wages in columns 5 and 6 reveals that individuals from villages that were far from the rail station gained 38.9% more in wages and 41.5% more in income after school consolidations.

These results are consistent with the following mechanism: students exposed to school consolidations earned more because those who attended new schools had more labor market opportunities. These improved labor market outcomes might not be the product of better teacher capability. Individuals in town-level labor markets had more choices and faced higher demand than those in village-level labor markets, even if they had the same education quantity and quality.

Table 9: Enhanced Labor Market Proximity and Better Labor Market Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(wage)	ln(income)	ln(wage)	ln(income)	ln(wage)	ln(income)
Panel A: econ zone vs. teacher's absolute wage						
Consolidation*far from econ zone	0.354***	0.422***			0.423***	0.529***
	(0.102)	(0.145)			(0.137)	(0.135)
Consolidation*teacher's absolute wage			0.215	0.196	0.097	0.084
			(0.210)	(0.250)	(0.241)	(0.271)
Observations	3,002	5,621	2,052	3,339	2,044	3,329
R-squared	0.665	0.468	0.464	0.355	0.466	0.356
Panel B: econ zone vs. teacher's relative wage						
Consolidation*far from econ zone	0.354***	0.422***			0.551***	0.592***
	(0.102)	(0.145)			(0.172)	(0.165)
Consolidation*teacher's relative wage			0.104	-0.266	0.132	-0.199
			(0.152)	(0.213)	(0.120)	(0.166)
Observations	3,002	5,621	1,871	2,989	1,871	2,989
R-squared	0.665	0.468	0.497	0.352	0.500	0.354
Panel C: distance to train station vs. teacher's absolute wage						
Consolidation*far from train station	0.304**	0.298			0.389**	0.415*
	(0.130)	(0.195)			(0.178)	(0.231)
Consolidation*teacher's absolute wage			0.215	0.196	0.050	0.107
			(0.210)	(0.250)	(0.264)	(0.275)
Observations	2,504	4,715	2,052	3,339	1,710	2,795
R-squared	0.663	0.472	0.464	0.355	0.484	0.359
Dependent Variable Mean	9.011	8.758	9.011	8.758	9.011	8.758
Cohort FE + Village FE	√	√	√	√	√	√
Province-Year FE	√	√	√	√	√	√

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level. The dependent variable in columns 1, 3, and 5 is individuals' log wages, and the dependent variable in columns 2, 4, and 6 is individuals' log incomes. The regression sample includes individuals born after1966 who were older than 15 at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village, and otherwise is 0. Panels A, B, and C interact school consolidation with the dummy of whether villages are close to an economic zone, the distance to the nearest train station, teachers' absolute wages, and teachers' wage relative to the wage of unskilled male workers

7.2 Reduction in gender inequality

In this subsection, I discuss why in my sample only females gained more lifetime education after school consolidations. Whether females benefitted more or less than their male counterparts is theoretically undetermined. Section 7.1 provides suggestive evidence that individuals who attended a new school after consolidations enjoyed better labor market proximity. However, for two reasons different genders might not equally enjoy the additional opportunities. First, rural households in developing countries usually have strong gender preferences for boys. If parents cannot afford the additional costs after school consolidation, they are more likely to force girls to drop out of school. Second, low-education labor markets in towns discriminate more against females. Thus, if females have access to the new schools after consolidations, they may prefer to have more education.

Town areas had many non-agricultural job opportunities for workers with different education levels. Among them, the low-education labor market (including construction and some manufacturing industries) was labor-intensive and favorable for males, while the high-education labor market discriminated less against females. If females wanted to enjoy the job opportunities brought by school consolidations, it would be desirable for them the obtain more education. In a unitary model, if the education return for females after school consolidations is high enough, girls receive more resources and work less early in life as child laborers. These early-stage human capital investments were critical to gender equality, and could have long-term impacts on women's empowerment when they grow up.

Table 10 tests my *Assumption 3* that the labor market for high-educated workers in town is more gender-equal. I define (relatively) high-education workers as individuals with more than a primary school education, and low-education workers as those with less than a primary school education. Panel A columns 1 and 2 shows that for individuals with more than 6 years of education in the CHNS sample, females in the village earned 25.5% less than males, and this gender income gap fell to about 12.8% among those who worked in the town. Panel A columns 3 and 4, which consider individuals with less than 6 years of education, show that the gender gap is larger among low-educated workers in towns than among workers in villages. Since the number of observations in Panel A column 4 is small, Panel B uses the 2005 mini census sample as a robustness check. Panel B columns 1 and 2 suggest that, on the one hand, the gender gap in town for individuals with more than 6 years of education is about 30.6%, which is smaller than a 36.2% gap in villages; on the other hand, the gender gap in towns for individuals with less than 6 years of education is about 40.0%, which is larger than a 30.9% gap in villages. Table 10 suggests that the gender income gap for high-education (low-education) workers is smaller (larger) in town than in village labor markets. My *Assumption 3* in the conceptual framework holds.

Table 10: Smaller Gender Income Gap in High-educated Labor Market in Town

	(1)	(2)	(3)	(4)
	ln(income)			
	education>6		education<=6	
	village	town	village	town
Panel A: CHNS sample				
Female	-0.255***	-0.128*	-0.133	-0.986***
	(0.0443)	(0.0628)	(0.101)	(0.197)
Dependent Variable Mean	8.735	9.206	8.245	8.547
Observations	3,427	1,037	1,127	22
R-squared	0.452	0.629	0.424	0.934
Cohort FE + Village FE	√	√	√	√
Province-Year FE	√	√	√	√
Panel B: 2005 mini census sample				
Female	-0.362***	-0.306***	-0.309***	-0.400***
	(0.0131)	(0.00939)	(0.0116)	(0.0127)
Dependent Variable Mean	5.824	6.365	5.443	5.811
Observations	378,788	154,864	329,188	50,360
R-squared	0.272	0.212	0.275	0.332
Cohort FE + Prefecture FE	√	√	√	√

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level in Panel A and at the prefecture level in Panel B. The dependent variable is individuals' log incomes. The regression sample in Panel A includes individuals born after1966 who were older than 15 at the time of the CHNS survey. The regression sample in Panel B includes individuals between ages 15 and 60 in the 2005 mini census. Columns 1 and 2 include individuals with more than 6 years of education, while columns 3 and 4 include individuals with less than 6 years of education. Columns 1 and 3 and columns 2 and 4 present the village sample and the town sample, respectively.

Table 11 examines the impacts of school consolidations on child labor and the heterogeneous treatments across households that have different assets. Panel A column 1 shows that after school consolidations, treated girls were 10.8 percentage points less likely to be child laborers. Columns 2 and 3 divided children into two groups on the basis of whether their household assets were more (rich) or less (poor) than 750 yuan, which is the first quartile. Girls were less likely to be child laborers in both rich and poor households after school consolidation. The coefficient is larger for poor households, although the difference between poor and rich households is insignificant. On the contrary, boys did not have less child labor after school consolidations. As shown in Panel B, the coefficients for boys from both poor and rich households were negligible and insignificant. The comparison of Panels A and B suggests that girls might receive more human capital investment from parents after school consolidations.

Table 11: Females are Less Likely to be Child Labor after School Consolidations

	(1)	(2)	(3)
	Probability of working before 16		
	All households	Poor households	Rich households
Panel A: girls			
School consolidation	-0.108**	-0.208*	-0.0850*
	(0.0418)	(0.111)	(0.0471)
Dependent Variable Mean	0.0390	0.0758	0.0198
Observations	2,206	725	1,439
R-squared	0.348	0.435	0.327
Panel B: boys			
School consolidation	0.0248	0.0667	0.00788
	(0.0153)	(0.0502)	(0.0189)
Dependent Variable Mean	0.0252	0.0442	0.0163
Observations	2,665	816	1,811
R-squared	0.154	0.242	0.153
Cohort FE + Village FE	√	√	√
Province-Year FE	√	√	√

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. Robust standard error clustered at the village level. The dependent variable is the dummy of whether an individual works before age 16. The regression sample in column 1 includes individuals born after 1966 who at the time of the survey were between 7 and 15 years of age. Column 2 includes individuals from poor households (whose assets were less than 750 yuan), while Column 3 includes individuals from rich households. Panels A and B are for females and males, respectively. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0.

Table 12 provides evidence for a reduction in gender inequality among individuals between ages 0 and 18 at the time of the survey. Columns 1 and 2 focus on physical health, while columns 3 and 4 investigate attitudes and preferences. Panel A shows that after school consolidations, treated females gained 5.35% more weight and 5.33% more height, suggesting that they had better nutrition than the control group[13]. As columns 3 and 4 show, treated females cared more about being liked by friends and getting good grades.[14] In many developing countries parents view girls as less important and investment-worthy than boys. As a result, girls might care less than boys about their personal development. Girls exposed to school consolidation gained access to people outside their traditional community, which exposed them to different values and led to more self-recognition.

Table 12 Panel B shows the placebo test results with the subsample of males.

[13] There are two potential channels. First, parents might have invested more in girls because they knew that the return on females' human capital was higher after school consolidations. This channel is consistent with my assumption that the urban labor market for high-educated workers is more gender-equalized. Second, new schools large enough to have dormitories and dining halls provided school lunches with government subsidies; thus, students at these schools might have had better nutrition.

[14] There are fewer observations in columns 3 and 4. This is so because the questions about priority were only asked after 2004. Thus, individuals surveyed before 2004 are not included.

Compared to similarly treated females, males do not have significantly more weight. They gained 3.35% more height after consolidation, which is smaller than their female counterparts. The coefficients on the priority of being liked by friends and achieving good grades were also smaller and less precisely estimated among males than among females. Although in columns 1 and 3 the differences in coefficients between Panel A and B are marginally insignificant, the comparison between Panels A and B suggests that there was a reduction in gender inequality among individuals younger than 18.

Table 12: Suggestive Evidence on Reduction in Gender Inequalities (Age <18)

	(1)	(2)	(3)	(4)
	ln(weight)	ln(height)	Priority of being liked by friends	Priority of getting good grades
Panel A: females				
School consolidation	0.0535***	0.0533***	0.487***	0.820***
	(0.0196)	(0.00854)	(0.183)	(0.209)
Dependent Variable Mean	3.248	4.825	2.508	2.700
Observations	2,782	2,739	506	502
R-squared	0.903	0.871	0.323	0.365
Panel B: males				
School consolidation	0.0155	0.0335***	0.174	0.159
	(0.0168)	(0.00679)	(0.152)	(0.167)
Dependent Variable Mean	3.259	4.829	2.441	2.503
Observations	3,640	3,598	640	620
R-squared	0.909	0.931	0.277	0.305
Cohort FE + Village FE	√	√	√	√
Province-Year FE	√	√	√	√

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. Robust standard error clustered at the village level. The dependent variables in columns 1 to 4 are individuals' log of weight, log of height, and how often individuals care about being liked by friends and getting good grades, respectively. The regression sample in column 1 and 2 includes individuals born after1966 who were 0 to 18 years of age at the time of the survey. The regression sample in column 3 and 4 includes individuals born after1966 who were 6 to 18 years of age at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0.

Table 13 provides evidence for the reduction in gender inequality among individuals older than 18 at the time of the survey. Column 1 shows that females exposed to school consolidation spent 15.8 minutes less time on cooking, buying food for the family, or doing other housework per week than females not exposed to school consolidation. This is consistent with my previous finding that school consolidation increased the labor force participation of treated females. Column 2 shows that when exposed to school consolidation, females were 12.7 percentage points more likely to migrate for work after they grew up. Columns 3 and 4 show that treated females had more control over TV programs at home than non-treated females.

As a comparison, Panel B shows that the time males spent on housework and the probability of migrating for work were not affected by school consolidations. The large ex ante gender gaps in housework load and migration were closed after school consolidation. The results of Panels A and B suggest that treated females had more bargaining power at home and enjoyed more choices and opportunities in the labor market than non-treated females.

Table 13: Suggestive Evidence of Reduction in Gender Inequalities (Age > 18)

	(1)	(2)	(3)	(4)
	Time on housework	Migration for work	Only wives choose TV programs	Only husbands choose TV programs
Panel A: females				
School consolidation	-15.83***	0.127***	0.591**	-0.385**
	(5.977)	(0.0433)	(0.235)	(0.165)
Dependent Variable Mean	22.85	0.352	0.157	0.108
Observations	3,592	2,856	842	842
R-squared	0.304	0.275	0.142	0.156
Panel B: males				
School consolidation	1.864	-0.00178		
	(2.459)	(0.0398)		
Dependent Variable Mean	6.841	0.467		
Observations	3,660	2,814		
R-squared	0.210	0.265		
Cohort FE + Village FE	√	√	√	√
Province-Year FE	√	√	√	√

Note: Statistical significance: * 0.10 ** 0.05 *** 0.01. Robust standard error clustered at the village level. The dependent variables in columns 1 to 4 are the times that individuals spent doing housework every week; dummies of whether an individual once migrated for work; and whether TV programs are chosen only by the husband or only the wife, respectively. The regression sample includes individuals born after1966 who were older than 18 at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0.

Examining Tables 11, 12, and 13 together reveals suggestive evidence that households invested more in girls' human capital, which impacts females both in their early adolescence and when they grow up. This pattern is consistent with the story that females had an incentive to get more education because the high-education labor market in towns discriminated less against females. The reduction in gender inequality during the school consolidation process explains why treated females gained more education while treated males did not.

8. THE DISTRIBUTIONAL EFFECTS

8.1 Distributional effects of education, wage, and incomes

The overall distributional impacts of consolidation remain inconclusive. In this subsection, I follow Carrell, Hoekstra, and Kuka (2018) and investigate two questions: (1) Did school consolidations lead to more education for females at the upper end or lower end? (2) Which individuals along the earnings distribution were most affected? To this end, I estimate my main specification in equation (5), except that I define the dependent variable to be an indicator of whether the individual's education level or earning level exceeded a given amount.

Figure 8 shows the resulting estimates and the 95% confidence interval across the education distribution, where the two sub-figures refer to education for females and males, respectively. Results suggest that while the average impact on education was not significant, there was a salient difference across genders. That is, while exposure to school consolidation did not increase education for most low-educated females, the females from the middle to the upper end of the education distribution gained more education. This might have been caused either by ex-ante heterogeneities in learning capacity or budget constraints, indicating that school consolidations might have enlarged the education inequalities among females.

In contrast, throughout the distributions the consolidation had negligible and insignificant impacts on males. The only exception is that the probability of males who had more than four years of education decreased after school consolidation, which is consistent with my finding that treated individuals are more likely to drop out before the 3rd grade. The education for high-educated males also decreased, although insignificantly. The heterogeneities across genders are consistent with the argument that the opportunity costs of education for males increased after school consolidation. Increased opportunities for low-educated male workers in town might have induced males to drop out early; hence, the increased male earnings were more likely to have been driven by proximity effects than by education effects. Females might have had higher earnings because they had more education and the high-education labor market in town was less gender discriminatory.

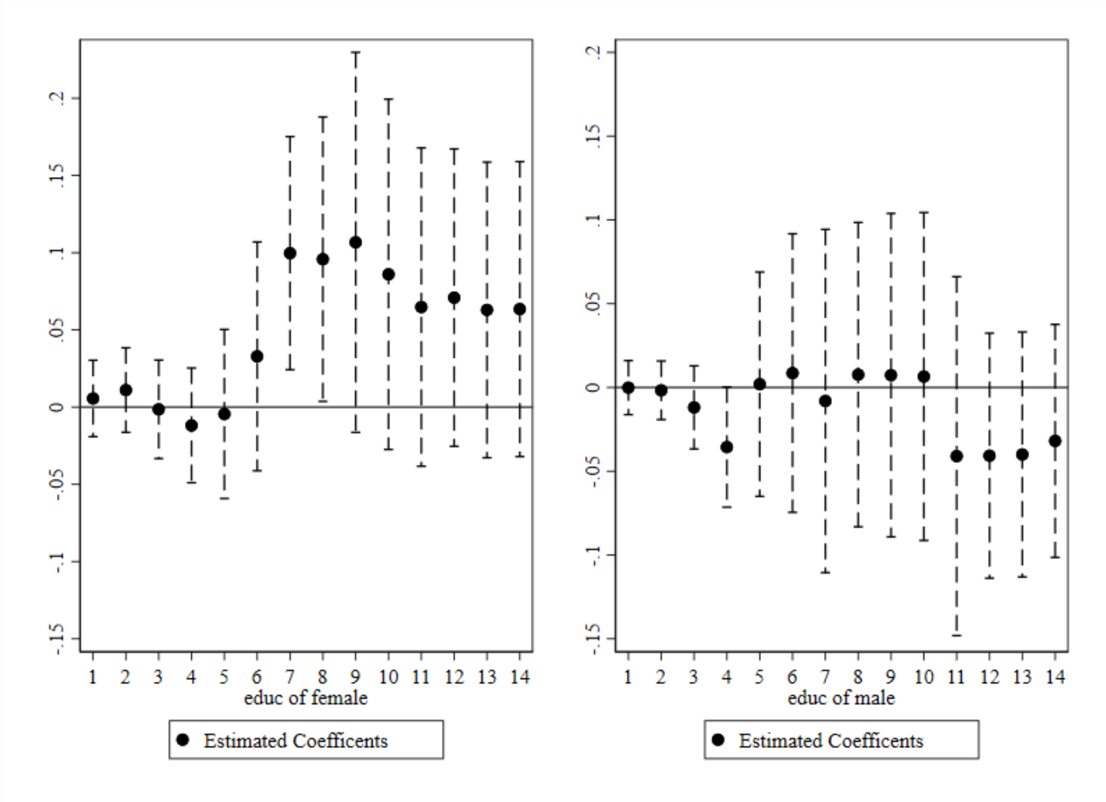

Figure 8: Effects of School Consolidations on the Distribution of Education

Note: I estimate the generalized DID specification, and the dependent variable is a dummy variable for whether the individual's educational attainment exceeded a given amount. The regression sample includes individuals born after 1966 who were older than 15 at the time of the survey. All regressions include community and birth cohort fixed effects, and province-by-year fixed effects. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0. I report the resulting estimates and 95 percent confidence intervals. The robust standard error is clustered at the village level.

Similarly, I examine whether exposure to school consolidation affected individuals equally across the earnings distribution. Figure 9 shows the results for income and wage distribution. School consolidation increased incomes more for individuals at the middle-to-upper end of the income distribution. In addition, individuals at the bottom end of income distribution were worse off, although the magnitude is small. Exposure to school consolidations increased wages predominantly for individuals at the lower-to-middle end of the distribution and had negligible impacts on the top end of the distribution. However, this does not conflict with the results for income distribution. Because the average income was higher for wage-earning individuals, the individuals who benefitted more in wages (lower-to-middle end) largely overlapped with those who benefited more in incomes (middle-to-upper end).[15] I also observe a small but significant decrease in wages at the very bottom part of the wage distribution. This was probably due to a higher labor force participation rate after school consolidations.

[15] One may presume that individuals who earn wages occupy a higher part of the income distribution. The majority (5%-95%) of individuals who earned log wages between 7.5 and 9 had log earnings between 8.2 and 9.8.

The newly-entered groups were less competitive than those already in the sector and earned lower wages. Overall, school consolidations narrowed the inequalities between individuals from the middle and the top ends of the income distribution, but meanwhile it enlarged inequalities between individuals from the bottom and the middle part of the income distribution.

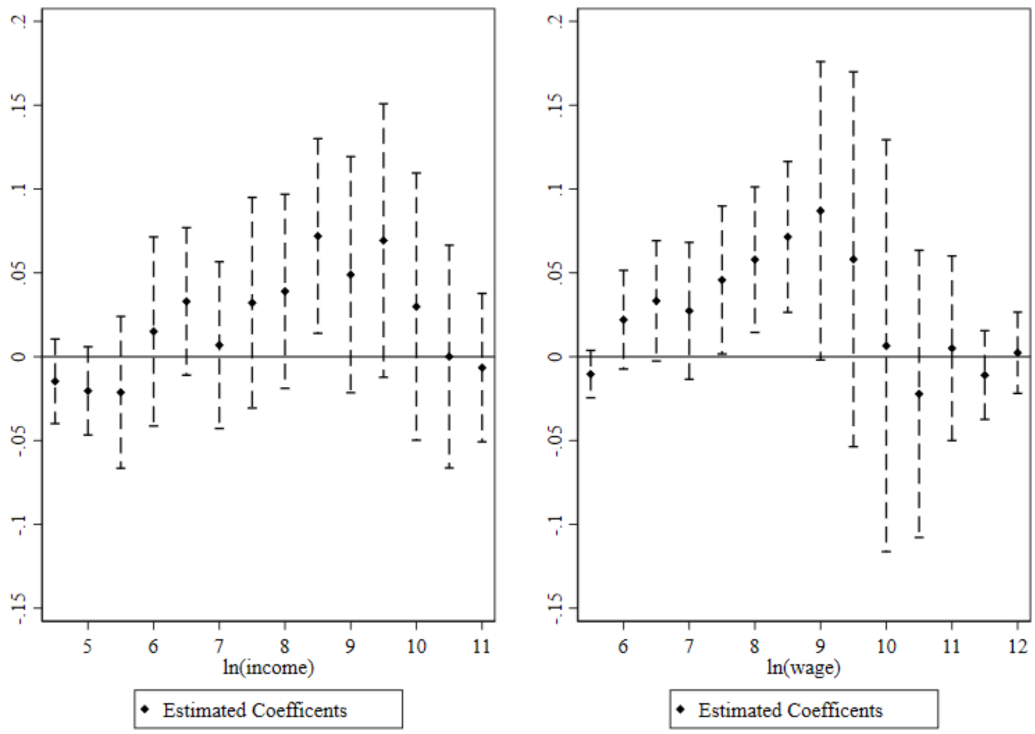

Figure 9: Effects of School Consolidations on the Distribution of Income and Wage

Note: I estimate the generalized DID specification, and the dependent variable is a dummy variable for whether the individual's log of income or wage exceeded a given amount. The regression sample includes individuals born after 1966 who were older than 15 years old at the time of the survey. All regressions include community and birth cohort fixed effects, province-by-year fixed effects, and a gender dummy. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0. I report the resulting estimates and 95 percent confidence intervals. The robust standard error is clustered at the village level.

8.2 Weighted average treatment effects using the Atkinson's approach

Subsection 8.1 showed large heterogeneous treatment effects for individuals at different ranks in the distribution. However, we still do not know whether the change in average earnings is large enough compared with the change in inequality, which depends largely on the weights we put on individuals. Figure 9 suggests that in an extreme case, if we only care about the people at the very bottom end of the income distribution, the aggregate impact on utility could be negative. In this subsection, I follow Maitra et al. (2022), who evaluate efficiency-equity trade-offs in the utilitarian tradition of Atkinson (1970).

I assume the social utility function is symmetric, additively separable in individual outcomes, and homothetic. The impact of school consolidations is an increasing, concave function of the incomes. Specifically, assume the individual utility function is given by

$$U(y_{ijct}) = \frac{{y_{ijct}}^{1-\theta}}{1-\theta} \tag{9}$$

where y_{ijct} is individual i's income. θ is the inequality-aversion parameter. When $\theta = 0$, the weighted average treatment effect represents the change in average individual income; when $\theta = 1$, U is a logarithmic utility function that represents the proportional changes in income. A higher value of θ indicates greater weight on the well-being of worse-off individuals in the social utility function.

In each column of Table 14, I present the treatment effects for different values of θ. At higher values of θ, the incomes of low-income individuals receive greater weight in the calculation. Therefore, the same increase in income would have a larger impact on utility if it were an increase among lower-income individuals rather than among higher-income individuals. Moreover, the utility reduction caused by an increase in the inequality of income distribution is more likely to overpower the increased utility caused by higher average incomes. The weighted treatment effects of school consolidations were positive and statistically significant for θ equals 0.5, 1, and 1.5, and were positive but marginally insignificant for θ equals 0 and 2. Hence, the increase in inequality was relatively small compared to the efficiency increase. The direction and magnitude of the weighted average treatment effects depend on the inequality-aversion parameter θ, which reflects the efficiency and equity objectives.

Table 14: Weighted Average Treatment Effects of the School Consolidations

	(1)	(2)	(3)	(4)	(5)
Value of θ	θ=0	θ=0.5	θ=1	θ=1.5	θ=2
School consolidation	2,561	18.61**	0.234**	0.00411**	0.000101
	(1,564)	(9.215)	(0.0910)	(0.00184)	(6.61e-05)

Notes: Statistical significance: * 0.10 ** 0.05 *** 0.01. The robust standard error is clustered at the village level. Weighted average treatment effects are estimated following the procedure outlined in Section 8, where θ indicates the value of the inequality-aversion parameter. The regression sample includes individuals born after 1966 who were older than 15 at the time of the survey. Consolidation Exposure is a dummy that takes 1 if the individual was younger than 13 when there was a primary school closure in the village and otherwise is 0.

9. CONCLUSION

A thorough investigation of the long-term impacts of school consolidation is significant because the policy has occurred not only in China but also in many other countries across the world. Many developing countries have recently experienced rapid demographic transitions. The fertility decline and rural-urban migration have caused an unprecedented shrinkage of rural populations, and many primary schools in rural areas are not fully filled. Thus, how to reallocate education resources efficiently without sacrificing wide coverage is a question of broad interest, and school consolidation is likely to become a popular policy in many countries. The long-term effects and distributional effects of school consolidation are important in economic research and are also critical to human capital formation and structural transformation in a broad geographical context. Yet to date few studies of the impacts of school consolidations in developing countries have generated plausible inferences about the mechanisms that precipitate those impacts.

By exploiting a generalized DID identification strategy, I study the plausible causal effect of school consolidation on education and the labor market in rural China. I find that school consolidations increased females' education but not males' education. In contrast, treated individuals of both genders had higher earnings than the non-treated. An important characteristic of school consolidation understudied in the literature is that conditional on attending the new schools, students will be exposed to a more prosperous and gender-equalized labor market for high-educated workers because those schools usually locate in towns. In this paper, I provide suggestive evidence consistent with the mechanisms of enhanced labor market proximity and reduced gender inequality after school consolidations. School consolidations also have important implications for inequalities. First, they narrowed the education inequalities across genders. Second, they narrowed the earning inequalities across villages, as individuals from villages far from economic zones had larger increases in wage and income. Third, inequalities were potentially enlarged across households, as early dropouts increased after consolidations. The analysis of the distributional effects indicates that individuals in the middle part of income distribution benefited more from school consolidations, and the increase in inequality was relatively small compared to the efficiency increase.

The prerequisites of the two main mechanisms discussed—limited labor market proximity in remote villages and discrimination against females in the low-education labor market—can be widely applied to other developing countries. However, we should be mindful of the influence on the enrollment rate when analyzing the impacts of school consolidations. Most of the positive effects of consolidation are conditional on attending a new school; if the dropout rate increases significantly, the overall effects could be negative. Overall, the trade-offs initiated by the school consolidations analyzed in this paper could apply to many other contexts, but the actual impacts of school consolidations can vary considerably depending on the program implementation details.

References

Andrabi, Tahir, Jishnu Das and Asim Ijaz Khwaja. 2013. "Students today, teachers tomorrow:

Identifying constraints on the provision of education." *Journal of Public Economics* 100:1–14.

Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn. (2020). Negotiating a better future: How interpersonal skills facilitate intergenerational investment. *The Quarterly Journal of Economics*, *135*(2), 1095-1151.

Ashraf, Nava, Erica Field and Jean Lee. 2014. "Household bargaining and excess fertility: An experimental study in Zambia." *American Economic Review* 104(7):2210–2237.

Atkinson, Anthony B. 1970. "On the measurement of inequality." *Journal of Economic Theory* 2(3):244–263.

Baland, J. M. and J. A. Robinson. 2000. "Is child labor inefficient?" *Journal of Political Economy* 108(4):663–679.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul and Munshi Sulaiman. 2020. "Women's empowerment in action: Evidence from a randomized control trial in Africa." *American Economic Journal: Applied Economics* 12(1):210–259.

Baranov, Victoria, Sonia Bhalotra, Pietro Biroli and Joanna Maselko. 2020. "Maternal depression, women's empowerment, and parental investment: Evidence from a randomized controlled trial†." *American Economic* Review 110(3):824–859.

Bau, Natalie, Martin Rotemberg, Manisha Shah and Bryce Steinberg. 2021. "Human capital investment in the presence of child labor.". Working paper.

Beuchert, Louise, Maria Knoth Humlum, Helena Skyt Nielsen and Nina Smith. 2018. "The short-term effects of school consolidation on student achievement: Evidence of disruption?" *Economics of Education Review* 65:31–47.

Burde, Dana and Leigh L Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Con- trolled Trial of Village-Based." *American Economic Journal: Applied Economics* 5(3):27–40.

Callaway, Brantly and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225(2):200–230.

Carrell, Scott E., Mark Hoekstra and Elira Kuka. 2018. "The long-run effects of disruptive peer." *American Economic Review* 11:3377–3415.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner and Ben Zipperer. 2019. "The effect of minimum wages on low-wage jobs." *Quarterly Journal of Economics* 134(3):1405–1454.

China Ministry of Education. 2000. "Statistical communiqué on the development of national education in 2000.". Report.

de Chaisemartin, Clément and Xavier D'Haultfoeuille. 2022. "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey.". Working paper.

de Chaisemartin, Clement, Xavier D'Haultfoeuille and Antoine Deeb. 2020. "TWOWAYFEWEIGHTS: Stata module to estimate the weights and measure of robustness to treatment effect heterogeneity attached to two-way fixed effects regressions".

Dhar, Diva, Tarun Jain and Seema Jayachandran. 2022. "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India." *American Economic Review* 112(3):899–927.

Duflo, Esther. 2001. "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." *American Economic Review*

91(4):795–813.

Duflo, Esther. 2012. “Women empowerment and economic development.” *Journal of Economic Literature* 50(4):1051–1079.

Engberg, John, Brian Gill, Gema Zamarro and Ron Zimmer. 2012. “Closing schools in a shrinking district: Do student outcomes depend on which schools are closed?” *Journal of Urban Economics* 71(2):189–203.

Goodman-Bacon, Andrew. 2021. “Difference-in-differences with variation in treatment timing.” *Journal of Econometrics* 225(2):254–277.

Haan, Monique De, Edwin Leuven and Hessel Oasterbeek. 2016. “School consolidation and student achievement.”*Journal of Law, Economics, and Organization* 32:816–839.

Hannum, Emily, Xiaoying Liu and Fan Wang. 2021. “Estimating the effects of educational system consolidation: The case of china’s rural school closure initiative.” *Economic Development and Cultural Change* 70(1):485–528.

Howley, Craig, Jerry Johnson and Jennifer Petrie. 2011. “Consolidation of schools and districts: What the research says and what it means.”. *National education policy center.*

Kahn-Lang, Ariella and Kevin Lang. 2019. “The promise and pitfalls of Differences-in-Differences: Reflections on 16 and pregnant and other applications.” *Journal of Business & Economic Statistics* pp. 1–14.

Kazianga, Harounan, Dan Levy, Leigh L Linden and Matt Sloan. 2013. “The effects of "girl-friendly" schools: Evidence from the BRIGHT school construction program in Burkina Faso.” *American Economic Journal: Applied Economics* 5(3):41–62.

Lee, Jin and Christopher Lubienski. 2017. “The impact of school closures on equity of access in Chicago.” *Education and Urban Society* 49(1):53–80.

Lei, Wanpeng and Lu Xu. 2011. “Government responsibility for rural school bus development: Compulsory education layout adjustment as background.” *Journal of the Chinese Society of Education* 1:6–9.

Liang, Yucheng. 2011. “Research on the success tracking rates in panel survey: Sample attrition in the context of social transition.” *Sociological Studies (Chinese)* 6:132–153.

Xiang, Ma. 2017. “Economies of Scale and Heterogeneity in the Provision of Public Goods: Evidence from School Consolidation in China.”. Working paper.

Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee and Sujata Visaria. 2022. “Evaluating the distributive effects of a micro-credit intervention.” *Journal of Development Economics*.

Mei, Hong, Quanbao Jiang, Yuanyuan Xiang and Xiaoping Song. 2015. “School consolidation: Whither China’s rural education?” *Asian Social Work and Policy Review* 9(2):138–150.

Muralidharan, Karthik; and Nishith Working Prakash. 2018. “Cycling to school: Increasing secondary school enrollment for girls in India.” *American Economic Journal: Applied Economics* 9(3):321–350.

Patrinos, Harry A. and Eduardo Velez Bustillo. 2018. “Four education trends that countries everywhere should know about.”. Report.

Roth, Jonathan and Pedro H. C. Sant’Anna. 2021. “Efficient Estimation for Staggered Rollout Designs.”. Working paper.

Secondo, Noah. 2020. “Rural school consolidation is not the answer.”. Report.

Sicular, Terry and Yaohui Zhao. 2004. “Earnings and labor mobility in rural China: Implications for China’s WTO entry”. In *China and the WTO: Accession, policy reform and poverty* *reduction Strategies*. World Bank pp. 239–260.

Sun, Liyang and Sarah Abraham. 2021. “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.” *Journal of Econometrics* 225(2):175–199.

Taghizadeh, Jonas Larsson. 2020. “Are students in receiving schools hurt by the closing of low-Performing schools? Effects of school closures on receiving schools in Sweden 2000–2016.” *Economics of Education Review* 78.

Xinhua News Agency, China. 2009. “Glory for 60 years: the disappearance of the words in "Literacy" and a backward era.”. Report.

Yang, Dongping. 2012. “Rural primary schools reduced by over 50% in 10 years of School Closure and Consolidation.”. Report.