Overleaf Package Template File

Fan Wang *

January 16, 2019

Abstract

We estimate the impact of educational infrastructure consolidation on educational attainment using the case of China's rural primary school closure policies in the early 2000s. The rationale for consolidation policies is that economies of scale will enable higher quality education, but distance to school increases for children in villages with school closures. We use data from a large household survey covering 728 villages in 7 provinces in China, and exploit variation in villages' year of school closure and children's ages at closure to identify the causal impact of school closure. For children exposed to closure during their primary school ages, we find an average decrease of 0.60 years of schooling for girls after age 15, but no significant effect for boys. Negative effects strengthen with time since closure. There is no evidence of a differential trend in educational attainment between individuals in villages with school closure and those in villages without, as indicated by an analysis using older cohorts beyond primary school age when closure occurred. Different effects by gender may be related to greater sensitivity of girls' enrollment to distance and greater responsiveness of boys' enrollment to quality.

---

*See Tex4Econ for more latex examples.

1 INTRODUCTION

According to Becker and Tomes (1986), ipsum dolor sit amet, consectetur adipiscing elit. Integer placerat nunc orci, id pellentesque lacus ullamcorper at. Mauris venenatis gravida magna non dapibus. Nullam vel consequat purus, id luctus dui. Suspendisse vel auctor nulla. Proin ipsum felis, efficitur eu eleifend vitae, efficitur pellentesque mauris (Case, Fertig, and Paxson 2005; Conti and Heckman 2010).

2 DATA

Village closure information is taken from a village head survey, which was collected in conjunction with household surveys. Village heads were asked if the village currently had a primary school, and asked about the year of school closure if the village school had been closed. Based on the village heads survey, there are four categories of closure status. The first category includes 193 villages that did not have village schools in 2011 and experienced school closure between 1999 and 2010. In the second category, which included 22 villages, a school closure year between 1999 and 2010 was reported, but village heads also reported that the village currently had a school in 2011. In this case, it is plausible that new schools were built in these 22 villages after school closure.[1] In the third category, 430 villages had village schools in 2011 and did not experience school closure.[2] Finally, the fourth category includes 48 villages that had never had a primary school and 35 that do not currently have a school but had a village primary school at some point between 1954 and 1999. In the following analysis, we designate the first and second categories as school closure. The third and fourth categories are coded as non-closure.[3]

3 MODEL

Let τ be the fraction of poorest children receiving price discounts and δ be the percentage price discount that children receive. $Z(\tau, \delta)$ is the total cost of a subsidy in grams of protein for 1970, 1972, 1974 and 1976 cohorts, given reference point distribution Γ for each cohort:

---

1. Generally students went to schools in township centers after village school closure, but in these 22 villages, it is possible that a new consolidated school was built inside these villages.

2. We do not have survey information on the opening year of the schools. The vast majority of these schools should have been established in the 1980s and early 1990s when the central government aimed to have a school in each village to provide education to rural children.

3. We will test the robustness of regression results to dropping the second and fourth categories from the closure and non-closure groups.

$$Z\left(\tau, \delta\right) = \sum_{\substack{\text{cohort} \\ \in\{70,72,74,76\}}} \left\{ \delta \cdot \int_{\epsilon} \int_{Y_{min}}^{F_Y^{-1}(\tau)} \int_X N\left(\begin{smallmatrix} Y,X,\epsilon; \\ \delta,\Gamma_{\text{cohort}} \end{smallmatrix}\right) f\left(X|Y\right) f\left(Y\right) f\left(\epsilon\right) dX dY d\epsilon \right\} \quad (1)$$

4 ESTIMATION

In Section 3, we categorized villages by their school closure status. Among the 193 villages that did not have village schools in 2011 and experienced school closure between 1999 and 2010, there were 45 villages where the closed village school was a teaching-point. Among the 430 villages that had village schools in 2011 and did not experience school closure, there were 44 villages where the village school in 2011 was a teaching-point. Teaching-point village primary schools offer up to 4 years of within-village primary education partly with the aim of reducing travel distance for students. While teaching-points might offer proximity and small-school benefits for young village students, when quality is defined narrowly in terms of the physical facility quality and teacher qualifications, teaching-points are of lower quality compared to other primary schools. In this section, we analyze the heterogeneity of closure effects by teaching-points status.

We estimate:

$$
\begin{aligned}
E_{pvia} &= \phi + \beta_v + \rho_{pa} + \rho_a^\tau \cdot TP_v \\
&\quad + \sum_{\tau \in \{0,1\}} \left(\sum_{z=1}^Z \tilde{\lambda}_z^\tau \cdot \mathbb{1}\{l_z \leqslant t_i \leqslant u_z\} \cdot c_v \right) \cdot \mathbb{1}\{TP_v = \tau\} \\
&\quad + X_i \cdot \gamma + X_i \cdot TP_v \cdot \gamma^\tau \\
&\quad + \epsilon_i
\end{aligned}
\quad (2)
$$

where $TP_v = 1$ if the village had a teaching-point school in 2011 or had a teaching-point school that was closed before 2001. Equation 2 allows teaching-point villages to have differential age patterns and covariate effects.

5 RESULTS

Table 1 presents estimates of λ in Equation 1. The first panel presents overall results, while the sex-specific results are shown in the second and third panels. In each panel, we compare three subsets of children below age 14 against baseline group – those between 14 and 21 at the time of school closure. Columns 1 and 2 include all individuals between 1 and 44 years of age in 2011, columns 3 and 4 restrict to individuals between 10 and 34, and columns 5 and 6 restrict further to individuals between 15 and 34 years

Table 1: Effect of School Closure on Educational Attainment

| | | | Outcome: grades completed by year 2011 | | | |
| | | | 10 ⩽ 2011 Age ⩽ 34 | | 15 ⩽ 2011 Age ⩽ 34 | |
	1	2	3	4	5	6
Panel Male and Female						
x Age at closure 0-5	-0.23	-0.30				
	(0.17)	(0.19)				
x Age at closure 6-9	-0.29*	-0.38**	-0.39**	-0.53***		
	(0.16)	(0.18)	(0.18)	(0.20)		
x Age at closure 10-13	-0.42***	-0.47***	-0.46***	-0.53***	-0.49***	-0.55***
	(0.14)	(0.16)	(0.14)	(0.15)	(0.15)	(0.16)
x Age at closure 22-29	0.11	0.18	0.027	0.12	-0.0062	0.079
	(0.17)	(0.18)	(0.17)	(0.19)	(0.17)	(0.19)
Observations	18804	15918	12072	10289	9998	8538
Exclusions and controls:						
Village FE, province-age FE and controls	Yes	No	Yes	No	Yes	No
Exclude villages that never had schools	Yes	Yes	Yes	Yes	Yes	No

†Includes village fixed effects, province-specific age FEs, controls for ethnicity, household size and relative household wealth.
‡Odd columns check robustness by dropping category 2 closure and category 4 non-closure status villages discussed in Section 3.
Statistical significance:* 0.10 ** 0.05 *** 0.01. Robust standard error clustered at village level. Each column (in each panel) represents a separate regression. Compare children with different starting age of exposure (and length of exposure) to children in baseline group who should not be impacted by school closure. Columns 3 to 6 restrict regressions to subsets of individuals who are closer together in 2011 age, testing the robustness of results to excluding younger and older individuals (see Table **??**). Regression sample includes villagers who were below 29 at year of closure for villages that experienced closure between 1999 and 2010. Individuals included in regressions from all columns are below 45 years of age in year 2011.

of age. The even-numbered columns drop villages that never had a school from the villages without closure group (category 4 as defined in data section). All regressions include several individual and household controls.[4] All standard errors are clustered at the village-level. Column 1 contains our focal main result, other columns contain results for robustness checks which we discuss later.

As noted earlier, school closures typically imply disruption, greater distance, and better quality school facilities for affected students. Our prior analyses showed that the impact of closure on attainment is not of short-term duration, indicating that there is limited short-term disruption effects of closure on attainment. To investigate mechanisms behind the impacts of school closure on educational attainment, we analyze how the two possible factors – distance to school and quality of school – are linked to enrollment status at the time of the survey. Here, we focus only on children between 5 to 12 years of age in 2011, ages at which nearly all children attend primary school in 2011.[5] For these children, we have information about distance to primary school and primary school quality.[6] We use children from both villages with and without school

---

4. All regressions include controls for households size, a dummy for if the individual is Han and a categorical variable for the relative wealth. The relative wealth variable is based on the survey question that asked households if they are better or worse off than village average. We do not have income measures for all families. The village income per capita variable shown in summary Table 1 is from the village-head survey and not based on household incomes.

5. Although previously in regressions on educational attainment we include children who were age 13 at year of closure into the group that may still attend primary school and therefore be exposed to school closure, we only include children who are in the age range that is definitely for primary school in the enrollment regression here.

6. For older individuals in the survey who are included in the earlier attainment regressions, we do

Figure 1: Effect of School Closure on Educational Attainment (Number of Grades Completed by 2011) by 15 Age-at-Closure Group.
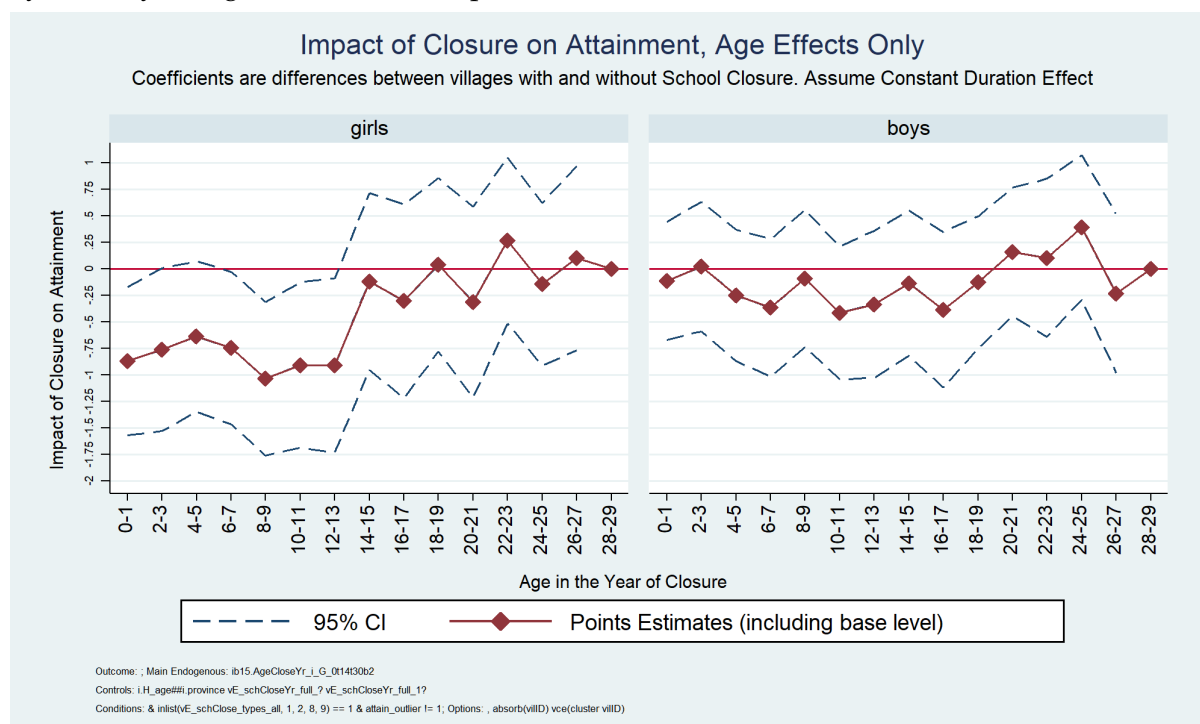
Each a–b group shows impact of closure on grades completed by 2011 for children who were between a and b years of age at the time of school closure. These results correspond to results from column 1 of Table 1 which had 5 age-at-closure groups.

closure in these regressions. The data for these analyses come from columns 2 and 3 of Table 1.

Primary school enrollment in 2011 is high but not full. For children at 5 years of age, the enrollment rate is 10 percent. At age 6 and 7, the enrollment rate increases to 47 percent and then 83 percent. Enrollment peaks at 94.4 percent at age 11.[7] In the following regressions, we analyze the relationship between distance, quality and school enrollment.

We regress enrollment on distance to closest primary school and quality of these schools. Regressions control for county fixed effects, province-specific age fixed effects, village per capita income, village per capita land size, village population size, household relative wealth, the number of household members and household ethnicity. Despite the controls, the coefficients we obtain for distance to school and quality of school would not be causal if there are unobserved village-level attributes that affect enrollment and that are also correlated with distance and quality. Our inclusion of village-level controls and county fixed effects, however, seeks to reduce the risk of omitted variables bias our estimates.

---

not have measures for the quality of school when they were attending primary school.

7. After primary school, at age 13, 15 and 17, enrollment rates drops to 92, 82, and 55 percent respectively.

6 CONCLUSION

In solving and estimating this model, I developed a set of solution and estimation algorithms. The solution algorithm captures the impacts of variations in the four fixed costs, three interest rates, and one collateral constraint on the household-specific feasible risky and safe asset choice sets, the average costs (returns) of borrowing (savings) across the seven credit market participation categories, and conditional and marginal asset distributions. The estimation algorithm chooses initial values for estimation by exploring the global properties of the estimation objective function through polynomial approximations. The solution and estimation algorithms can reduce estimation time by taking advantage of scalable cloud computing resources.

REFERENCES

Becker, Gary S, and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3): S1–39.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The lasting impact of childhood health and circumstance." *Journal of health economics* 24, no. 2 (March): 365–389.

Conti, Gabriella, and James J Heckman. 2010. "Understanding the Early Origins of the Education–Health Gradient." *Perspectives on Psychological Science* 5, no. 5 (September 1): 585–605.