Automation and teenage school dropout: evidence from China

Ruilong Li[1,*] and Yanzhuo Li[2]

Abstract: Studies of developed economies find that automation increases schooling. We examine this relationship at the pre-college stages in China, where educational progression is rationed. Using four census waves (2000–2015) and a shift-share instrument built from European robot diffusion, we find that a one-standard-deviation increase in prefecture-level robot exposure lowers the in-school rate by 1.17 percentage points at primary-school ages, 2.15 at junior-high ages, and 4.33 at senior-high ages. The declines are larger among children with rural hukou and in areas with less access to higher education. Where progression is rationed, automation weakens intergenerational mobility in education.

Keywords: industrial robots; automation; human capital; school dropout; access to further education; China

Classification: JEL I25, J24, O33

[*]**Correspondence address:** ruilongtomli@gmail.com

1 Introduction

How does automation reshape the next generation's supply of human capital? The task-based literature on technological change has established that machines, software, and industrial robots, by taking over routine tasks, reallocate labor demand, employment, and wages across occupations and educational groups (Autor et al., 2003; Acemoglu and Restrepo, 2018, 2020). Acemoglu and Restrepo (2022) in particular link task displacement to the wage structure across given demographic and educational groups and show that automation accounts for an important part of the rise in the US education wage gap. But educational groups are not fixed across generations: teenagers can adjust their decisions about remaining in school and choosing an educational track in response to changes in the earnings structure. Automation therefore changes not only the price of existing skills but also, potentially, the future supply of skills through endogenous educational choices. This paper studies this pre-labor-market schooling margin: how robot exposure affects the decisions of school-age children and teenagers to remain in school before entering the labor market. In doing so, it extends the task-displacement framework from "the price of a given skill group" to "which skill group an individual enters," and examines the race between technology and education (Goldin and Katz, 2008) in an institutional setting where progression is rationed.

The direction of the schooling response is theoretically ambiguous. If robots depress low-skill wages and raise the skill premium, they lower the immediate opportunity cost of schooling and raise the relative return to more education, thereby increasing the time spent in school. If the shock first erodes family income and liquidity, educational investment may instead fall. Moreover, in an education system with selective progression at successive stages, what matters is not only whether college-educated workers earn higher wages but

whether the marginal student can reach college at all. Staying in secondary school is a sequential decision that preserves the option of continuing once more information about grades and progression is revealed (Stange, 2012); the value of that option depends on the tracks the student can actually reach. Where access to college is greater, an automation-induced increase in the relative value of abstract skills is more likely to encourage students to remain enrolled. Where access is limited, the value of continued secondary schooling depends primarily on the routine work available after high school, whose earnings prospects are themselves weakened by automation. The same robot shock can therefore push young people toward more education through lower opportunity costs and a higher skill premium, or push marginal students out through household constraints and a lower continuation value of feasible educational pathways. The empirical question is which force dominates and whether its relative importance varies with proximity to the educational gateway.

Recent studies directly estimate educational responses to automation, and the majority find increases in schooling and educational attainment in developed economies. Robot exposure and the decline of routine work raise college enrollment and degree attainment in the United States and shift German youth from vocational training toward universities (Dauth et al., 2021; Gunadi and Ryu, 2023; Di Giacomo and Lerch, 2026; Branco et al., 2025); at the high-school margin, they shift Norwegian students from vocational to academic tracks and on to college without increasing dropout rates, and reduce high-school dropout in declining US manufacturing areas (Bennett et al., 2026; Tuhkuri, 2023). The dominant channel is the decline in the opportunity cost of schooling as low-skill jobs disappear, rather than an increase in the college wage premium (Di Giacomo and Lerch, 2026; Branco et al., 2025),

and the gains are concentrated among lower-achieving students and those from families without college-educated parents, thereby increasing intergenerational mobility (Bennett et al., 2026). Counterexamples arise through household-level channels: exposure before adulthood in the United States temporarily lowers college enrollment before affected cohorts catch up (Kang, 2026), and in Sweden parental robot exposure lowers children's income and upward mobility and worsens their long-run educational outcomes. These Swedish effects are concentrated in low-income families and are driven by parental job loss and financial strain (Heyman and Olsson, 2026).

All of this evidence, however, comes from systems in which the marginal young person who seeks further schooling can obtain it: US colleges admit by application and community colleges have elastic capacity, Germany's system allows movement between vocational and tertiary tracks, and Norway has guaranteed every junior-high graduate a high-school place since the early 1990s, with a route from the vocational track to university since Reform 94 (Bertrand et al., 2021). This explains why the comparable educational transition at age 16 leads to track reallocation rather than school exit in Norway. In most developing economies, access to further schooling is instead rationed by selective examinations and quotas, and enrollment figures illustrate the scale of this difference. When the 1999 expansion was launched, China's tertiary gross enrollment ratio was 6.6% compared with 71.4% in the United States; by 2015, after the largest higher-education expansion in the world, it had reached 47.1%, still under six-tenths of the US level (88.9%).[1] In 2015 the ratio averaged 31.0% in low- and middle-income economies compared with 76.2% in high-income ones, with India (27.3%), Vietnam (29.7%), and Indonesia (32.5%) in China's range, suggesting that

[1]Tertiary gross enrollment ratios (indicator SE.TER.ENRR) are from the World Bank's *World Development Indicators* (World Bank, 2026), defined as tertiary enrollment divided by the population within five years of secondary-school completion age.

the United States is the exception rather than the rule. This difference determines whether a decline in the opportunity cost of schooling translates into more time in school. The more restricted the progression opportunities, the more the value of continuing depends on the routine occupations that secondary-school graduates have historically entered—precisely the occupations whose earnings prospects automation weakens most. The positive response observed in developed economies therefore need not extend to settings where progression is rationed.

This paper examines how robots affect educational choices at the pre-college stages, using the 2000 population census, the 2005 1% population survey, the 2010 census long form, and the 2015 1% population survey collapsed into prefecture–census-year–age cells, with robot exposure measured using the shift-share approach of Acemoglu and Restrepo (2020) as adapted to China by Giuntella et al. (2025). The results show a clear gradient across schooling stages: a one-standard-deviation increase in robot exposure lowers the in-school rate by 1.17 percentage points at primary-school ages (7–11), 2.15 at junior-high ages (12–14), and 4.33 at senior-high ages (15–17). The age-specific estimates become systematically more negative as children approach the senior-high gateway, and the 2000–2015 endpoint long-difference specification reproduces the same ordering. This does not mean automation reversed China's schooling expansion: the senior-high-age in-school rate rose from 55.1% in 2000 to 81.8% in 2015, while a prefecture with a one-standard-deviation larger increase in exposure experienced a smaller enrollment gain, forgoing 6.53 points of that rise. The senior-high coefficient remains negative under the standard robustness checks and diagnostics for shift-share designs (leave-one-sector-out estimates, controls for differential trends related to pre-period characteristics, migration restrictions, alternative weighting, Rotemberg weight

diagnostics, and shock-level inference), and destination shares show that those who leave school mainly go into agricultural employment. Two heterogeneity analyses link this gradient to access to further education. Across provinces, enrollment declines are larger where 1998 admission chances were lower. We also find that robot exposure weakens intergenerational mobility in education: where exposure is higher, whether a teenager stays in school at ages 15–17 depends more strongly on whether the parents have higher education.

The paper contributes to two literatures. First, it adds to the literature on automation and endogenous human capital. Existing work studies vocational training, high-school tracks, college enrollment, major choice, final attainment, and the intergenerational transmission of automation shocks (Dauth et al., 2021; Gunadi and Ryu, 2023; Di Giacomo and Lerch, 2026; Branco et al., 2025; Bennett et al., 2026; Kang, 2026; Heyman and Olsson, 2026). This paper shifts the focus to the pre-college schooling sequence at ages 7–17, where progression is still rationed at a gateway, and shows that the response becomes more negative as students approach the age at which school-leaving becomes feasible and is attenuated where ex ante progression opportunities were greater. Linking children to parental education further shows that the enrollment losses concentrate among children whose parents have less than a junior-college education, so that automation weakens intergenerational mobility in education already at the pre-college stage.

Second, the paper contributes to the literature on local labor-demand shocks and schooling choices. Educational investment responds quickly to perceived returns and immediate job opportunities (Jensen, 2010; Abramitzky and Lavy, 2014): declining opportunities for low-skilled workers raise high-school completion (Greenland and Lopresti, 2016; Tuhkuri, 2023), booms in low-skilled labor demand in coal, export manufacturing, and construction draw

students out of school (Black et al., 2005; Atkin, 2016; Charles et al., 2018), infrastructure-induced economic opportunities can raise educational investment by expanding household resources (Adukia et al., 2020), and in China the sign depends on the skill content and destination of the added demand (Li, 2018; Zhang and Zhou, 2023). This paper distinguishes the wage available upon leaving school from the future earnings value of the tracks accessible through continued schooling. Robots need not change a dropout's current wage to reduce the value of routine jobs historically entered by secondary-school graduates, and progression opportunities determine the extent to which the marginal student can avoid that devalued path.

2 A Two-Margin Framework for the School-Leaving Decision

Core intuition. Robots do not lower all returns to education proportionally; their effects differ across skill tracks. Automation substitutes for codifiable, repetitive routine tasks and primarily reduces earnings along the routine manufacturing track entered by senior-high graduates. The college track, which relies on abstract tasks, is affected much less adversely and may even gain in value. Junior-high leavers enter the labour market earlier, mostly in agriculture and other less-exposed work. The framework takes their earnings as the benchmark and assumes that students who enter the routine track after high school face larger wage losses relative to these earlier entrants. Staying in school provides an option whose payoff depends on the future track: with probability π, it leads to the college track; with probability $1 - \pi$, it leads to the routine track. When progression chances are limited,

the value of this option depends primarily on the routine track, which is most adversely affected, and the robot shock lowers the value of staying. The direction of the schooling response depends on the track underlying the marginal student's continuation option; its magnitude depends on the mass of students at the continuation margin and on how much of the marginal student's option rests on the routine track. More students are at this margin when they are closer to the educational gateway and opportunities for progression are more limited.

The individual continuation decision. The framework follows the self-selection model of educational investment developed by Becker (1975) and Willis and Rosen (1979), using the discrete-choice formulation that Charles et al. (2018) extend to multiple education types and Bennett et al. (2026) apply to high-school track choice: students self-select into educational tracks according to ability, which determines both the earnings on each track and the cost of schooling. The sole departure is that access to the college track in China is determined by examinations: whether a high-school graduate enters college depends on ability and on provincial admission quotas.

Consider a student of senior-high age deciding whether to stay in high school or drop out. Divide the student's remaining life into three periods: the high-school years, the college years, and working life thereafter. The student is risk neutral and compares the lifetime utility of the two choices: lifetime earnings discounted to the decision date with discount factor $\delta < 1$, net of the psychic cost of schooling. Earnings per period are w for a junior-high leaver, w_{high} on the routine track open to high-school graduates, and w_{univ} on the college track, with $w_{\text{univ}} > w_{\text{high}} > w$. A high-school diploma by itself provides only a small wage

premium in urban China (Li et al., 2012), so w_{high} is only slightly above w; the two paths also differ in the timing and sector of employment. The dropout works in all three periods; the student who stays earns nothing in high school and then either enters the routine track for the last two periods or, if admitted, spends the college years in school and earns w_{univ} in the final period only.

Staying in school also carries a psychic cost. Models of educational self-selection allow students to differ in the cost of schooling (Willis and Rosen, 1979; Charles et al., 2018), and in Chinese high schools this cost is tied directly to ability: the three years are devoted to preparing for the *gaokao*, and abler students keep pace more easily and bear a lighter burden. We write the psychic cost of high school for a student of ability θ as

$$c(\theta) \;=\; c_0 - c_1\,\theta, \qquad c_1 \geq 0.$$

Labeled by the highest level of schooling completed, the lifetime utilities of the three paths are

$$U_{\text{mid}} \;=\; (1+\delta+\delta^2)\,w, \qquad U_{\text{high}}(\theta) \;=\; (\delta+\delta^2)\,w_{\text{high}} - c(\theta), \qquad U_{\text{univ}}(\theta) \;=\; \delta^2 w_{\text{univ}} - c(\theta).$$

The latter two paths both pass through high school and therefore both bear the psychic cost. The framework's key assumption concerns only the lowest-ability students: for them, the wage gain from the diploma does not compensate for the first period's forgone earnings and the psychic cost,

$$U_{\text{high}}(0) \;<\; U_{\text{mid}}, \qquad \text{that is,} \qquad (\delta+\delta^2)\,w_{\text{high}} - c(0) \;<\; (1+\delta+\delta^2)\,w.$$

Whether staying leads to college depends on ability and on capacity. Ability θ is uniformly distributed on $[0, 1]$; the *gaokao* score rises with ability, the student is admitted if the score exceeds the provincial cutoff, and the cutoff falls with the province's admission capacity K. We write the admission probability as

$$\pi \;=\; K\,\theta, \qquad 0 < K \leq 1 :$$

higher ability raises the probability of admission, and greater capacity raises that probability at each ability level. Academic and vocational high school both count as staying; differences in their subsequent educational pathways are captured by π. At the time of the decision, the admission outcome is not yet known, so the expected utility of staying is $\pi\,U_{\text{univ}}(\theta) + (1 - \pi)\,U_{\text{high}}(\theta)$, compared with U_{mid} from leaving. The student stays if and only if the former is at least the latter,

$$\pi\,U_{\text{univ}}(\theta) + (1-\pi)\,U_{\text{high}}(\theta) \;\geq\; U_{\text{mid}}. \tag{1}$$

The expected utility of staying rises with ability, because the admission probability rises and the psychic cost falls, whereas the utility of leaving does not depend on ability. There is therefore a unique ability threshold

$$\bar{\theta} \;=\; \frac{U_{\text{mid}} - U_{\text{high}}(0)}{K\,(U_{\text{univ}} - U_{\text{high}}) + c_1} \;=\; \frac{w + c_0 - (\delta + \delta^2)\,(w_{\text{high}} - w)}{K\,[\,\delta^2 w_{\text{univ}} - (\delta + \delta^2)\,w_{\text{high}}\,] + c_1} : \tag{2}$$

students above $\bar{\theta}$ continue, students below it leave, and the student exactly at $\bar{\theta}$ is the *marginal student*, the lowest-ability stayer, whose admission probability we denote $\bar{\pi} \equiv K\bar{\theta}$. The observed enrollment response comes only from students near $\bar{\theta}$: those far below had

already left before the robot shock arrived, and their binary in-school indicator cannot decrease further; those far above still continue, and any additional study effort does not register in a binary in-school indicator. The shock induces the lowest-ability students who had just chosen to continue to leave school.

The exposure gradient of the robot shock. Let local robot exposure be x. We express wages in units of the wage of an earlier-entering junior-high leaver, scaled so that this benchmark equals w. In these relative units, the per-period wage schedules are linear in exposure:

$$w_{\text{high}}(x) \;=\; w_{\text{high}} - \rho\, x, \qquad w_{\text{univ}}(x) \;=\; w_{\text{univ}} - \eta\, x, \qquad w_{\text{mid}}(x) \;=\; w.$$

The reference group consists of junior-high leavers who enter the labour market earlier. Relative to this group, ρ measures the additional wage loss per unit of exposure for students who enter the routine track after high school, and $-\eta$ measures the corresponding relative wage *gain* for those who enter after college. The timing assumption is that later entrants into routine work face a larger impact as automation changes the jobs available at entry. College education is built on the abstract tasks that automation complements rather than replaces, so the college wage rises relative to the benchmark: $\eta < 0$. Thus, $w_{\text{mid}}(x) = w$ fixes the earlier entrants as the wage benchmark; their wages in original units may also respond to exposure. The paths also differ in sectoral composition. In the 2000 census, 81% of working 15–17-year-olds, nearly all of whom have junior-high schooling or less, work in agriculture, a sector with essentially no robots; when exposure is measured using 2015 sectoral robot

densities, their jobs have less than one-third of the exposure of those held by 18–22-year-old high-school and *zhongzhuan* graduates.[2] Section 4.2 shows that the exits induced by exposure go to the same destination. In the same relative wage units, these responses imply that, per unit of exposure, U_{high} changes by $-(\delta + \delta^2)\rho$ and U_{univ} by $-\delta^2\eta$, while U_{mid} remains the fixed benchmark. The *exposure-gradient assumption* has only two parts:

$$\rho > 0, \qquad \eta < 0. \tag{3}$$

The first part is the central finding of the task-based literature on automation: robots substitute for routine tasks, so wage losses are concentrated among less-educated workers performing routine tasks (Autor et al., 2003; Acemoglu and Restrepo, 2020; Dauth et al., 2021; Giuntella et al., 2025). The second part is the complementary finding that automation raises the earnings of college-educated workers relative to the less educated; the same rising college premium is what draws students into college in the developed-economy evidence (Goldin and Katz, 2008; Di Giacomo and Lerch, 2026; Branco et al., 2025). Both responses are measured relative to earlier-entering junior-high leavers. The assumption is deliberately stronger than the ordering $\eta < \rho$: it requires the college track to gain, not merely to lose less, and it is what allows the framework to produce a positive schooling response at all.

[2] 2000 census, employed and not in school, five autonomous regions excluded. Working 15–17-year-olds (N = 137,981; 98% with junior high or less): agriculture 81%, manufacturing 10%, robot-exposed sectors (automobiles, electronics, metals, machinery, plastics and chemicals) 3%; the 2015 China robot stock per thousand 2000 workers of their sector averages 0.27. Working 18–22-year-olds with academic high school or *zhongzhuan* (N = 66,366): agriculture 25%, manufacturing 22%, robot-exposed sectors 13%, sector robot density 0.96. Junior-high leavers observed at 18–22 remain on the first path (agriculture 74%, density 0.36). Source: `census_sf_dropout_industry_2000.do`.

Substituting the utilities of the three paths into (2) gives the response of the ability threshold to exposure:

$$\frac{\partial \bar{\theta}}{\partial x} \;=\; \frac{L(\bar{\pi})}{K\,(U_{\text{univ}} - U_{\text{high}}) + c_1}, \qquad L(\bar{\pi}) \;\equiv\; (1-\bar{\pi})\,(\delta+\delta^2)\,\rho \;+\; \bar{\pi}\,\delta^2\,\eta. \tag{4}$$

The loading $L(\bar{\pi})$ weights the earnings losses associated with the two possible outcomes of continuing in school by the marginal student's progression probability: with probability $1-\bar{\pi}$ the option falls back to the routine track and loses $(\delta+\delta^2)\rho$ per unit of exposure; with probability $\bar{\pi}$ it reaches the college track and loses $\delta^2\eta$. This loading determines the direction of the schooling response: when $L(\bar{\pi}) > 0$, the shock raises the ability threshold and induces students just above the previous threshold to leave. Under (3) the two ends of the loading have opposite signs, $L(0) = (\delta+\delta^2)\rho > 0$ and $L(1) = \delta^2\eta < 0$, and L is linear in $\bar{\pi}$, so it crosses zero exactly once, at

$$\pi^* \;\equiv\; \frac{(1+\delta)\,\rho}{(1+\delta)\,\rho - \delta\,\eta} \;\in\; (0,1). \tag{5}$$

For $\bar{\pi} < \pi^*$ the option of staying is written mostly on the routine track, $L(\bar{\pi}) > 0$, and greater robot exposure pushes the marginal student out of school. For $\bar{\pi} > \pi^*$ the marginal student is nearly certain of progression, the option is written mostly on the college track, $L(\bar{\pi}) < 0$, and the same shock keeps the marginal student in school. This is the margin on which the US and Norwegian evidence operates: admission capacity K is higher there, so the marginal student's admission probability $\bar{\pi}$ is higher as well, and we assume that it exceeds π^*; in China the opposite holds: capacity is limited, the marginal student is far from certain of progression, and $\bar{\pi}$ lies below π^*.

From individual choice to the in-school share. The in-school share S is the fraction of students with ability at or above the threshold; with uniform ability, it is the length of the interval above it:

$$S \;=\; 1-\bar{\theta} \;=\; 1-\frac{U_{\text{mid}}-U_{\text{high}}(0)}{K\left(U_{\text{univ}}-U_{\text{high}}\right)+c_1}. \tag{6}$$

The robot shock raises the ability threshold by $\partial\bar{\theta}/\partial x$, and students who fall below the new threshold leave school:

$$\frac{\partial S}{\partial x} \;=\; -\,\frac{\partial\bar{\theta}}{\partial x} \;=\; -\,\frac{L(\bar{\pi})}{K\left(U_{\text{univ}}-U_{\text{high}}\right)+c_1}, \tag{7}$$

which is strictly negative when $L(\bar{\pi})>0$. Equation (7) decomposes the schooling response into two factors: the loading $L(\bar{\pi})$ determines the direction of the response and how much the utility of staying changes; the denominator is the slope of the expected utility of staying in ability, and it scales the response by determining how many students a given change in utility sweeps across: the steeper the slope, the fewer students change their decision.

This delivers *attenuation*: greater capacity reduces the response to the same shock, through two channels. First, the denominator rises with K: where capacity is larger, the admission probability rises faster with ability, the same fall in the utility of staying corresponds to a smaller rise in the threshold, and fewer students are pushed out of school. Second, the marginal student's admission probability $\bar{\pi} = K\bar{\theta}$ rises with K, $\partial\bar{\pi}/\partial K = \bar{\theta}c_1/(K(U_{\text{univ}}-U_{\text{high}})+c_1) \geq 0$: where capacity is larger, the marginal student's option is written more on the college track and the loading $L(\bar{\pi})$ is smaller. This

conclusion can readily be extended to other ability distributions. The other comparative-static result used in the analysis is the level effect of progression chances: from (6), $\partial S/\partial K = \bar{\theta}\,(U_{\text{univ}} - U_{\text{high}})/(K(U_{\text{univ}} - U_{\text{high}}) + c_1) > 0$, so where admission capacity expands faster, the share choosing to continue into senior high school rises more—consistent with the evidence that the expansion raised academic-high-school completion and narrowed the urban–rural secondary gap (Zhang, 2026).

The framework thus yields two empirical predictions when $L(\bar{\pi}) > 0$, that is, when the marginal student's progression chance lies below π^* in (5). First, greater robot exposure reduces the in-school share by inducing marginal students to leave, and the response is stronger where admission capacity is more limited. Second, as K rises, the negative response approaches zero without changing sign: both channels only shrink the magnitude of the response, and $L(\bar{\pi})$ remains positive as long as $\bar{\pi}$ stays below π^*.

Farther from the gate, smaller response: the dropout ladder. The same decision framework applies at the two stages preceding the gateway, but two factors attenuate the response at earlier stages. The first is information. The farther a student is from the *gaokao* gateway, the less information the student has about the prospects for progression: information on grades, rankings, and cutoffs arrives grade by grade and is most complete only by the end of junior high school. Leaving school is (nearly) irreversible, and each additional year in school provides new information, so the value of staying contains a waiting option: students who stay preserve the right to choose again after uncertainty is resolved. This is "schooling as experimentation" in Manski (1989), Altonji (1993), and Stinebrickner and Stinebrickner (2012), and the option that Stange (2012) measures at college entry.

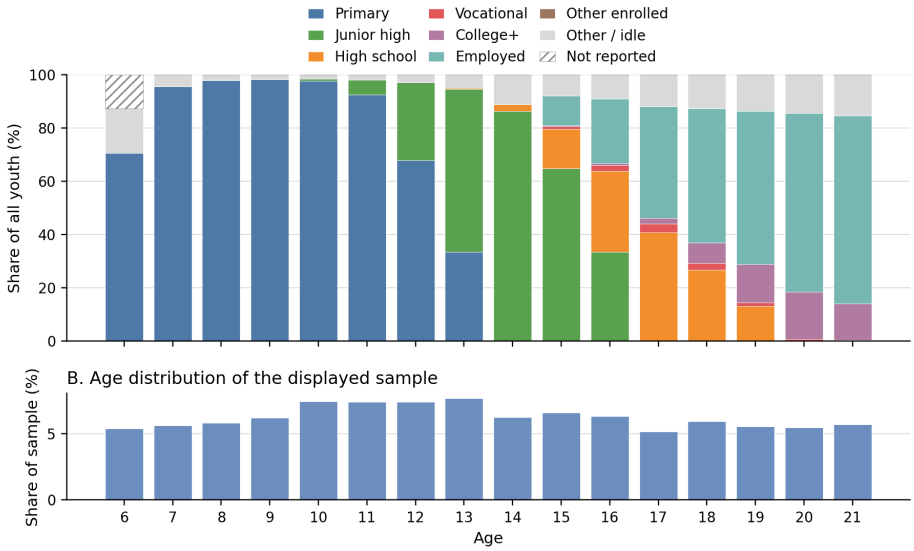
The value of the waiting option increases with unresolved uncertainty and vanishes at the gateway. It lowers the ability threshold required to remain in school, especially at earlier stages. At primary-school ages, the decision is close to a corner solution in which continuing is worthwhile for everyone and no students are induced to leave by the shock. The second factor is access to employment after leaving school. The Labor Law prohibits employing those under 16, an age threshold close to the transition into senior high school. Younger students who would otherwise leave school have limited opportunities to enter employment.

3 Data and Empirical Strategy

Census samples. The main analysis uses four waves of census data from 2000, 2005, 2010, and 2015. The 1990 census data are used to construct control variables. We define primary-school ages as 7–11,[3] junior-high-school ages as 12–14, and senior-high-school ages as 15–17.[4] At the province level, the sample covers 26 of mainland China's 31 provincial units—all 22 provinces plus every prefecture-level unit of the four provincial-level municipalities of Beijing, Tianjin, Shanghai, and Chongqing—excluding only the five ethnic autonomous regions, whose education-policy environment (minority education policy, bilingual schooling, and preferential admissions) is not comparable to that of the rest of the country.

[3]Children begin entering primary school at age 6, but many enter later, especially in rural areas. Starting the sample at age 7 reduces the risk of classifying delayed school entry as school leaving.

[4]Although some students still attend high school at age 18, many begin college at that age (Figure 1). College students may move their residence and transfer their hukou to the university location, making their census location an unreliable measure of where they received their pre-college education. We therefore exclude ages 18 and above to limit location misclassification associated with college entry.



Notes: Ages 6--21 are shown for description; the regressions use ages 7--17. Hatching indicates unreported status.
Employed denotes non-enrolled employed youth.
The descriptive sample excludes age-14/15 primary, age>=17 junior-high, and age>=20 academic-high-school cells.

Figure 1. *Current school level, employment status, and age distribution.*

Notes: Top panel: status shares by age, with enrolled youth classified by current school level and non-enrolled employed youth shown as employed. Bottom panel: age distribution of the displayed descriptive sample (ages 6–21). Hatching denotes unreported status. The regression sample remains ages 7–17.

Shift-share exposure. The exposure construction follows Acemoglu and Restrepo (2020), adapted to China as in Giuntella et al. (2025). For sector s, prefecture p, and census year t,

$$\begin{aligned} \exp_{p,t} &= \sum_s \ell_{p,s}^{2000} \frac{R_{s,t}}{L_{s,2000}} \times 1000, \\ \exp_{p,t}^{IV} &= \sum_s \ell_{p,s}^{2000} G_{s,t}^{EU}, \\ G_{s,t}^{EU} &= \frac{1}{|C_{s,t}|} \sum_{c \in C_{s,t}} \frac{R_{c,s,t}^{IFR}}{L_{c,s,t}^{EUKLEMS}} \times 1000, \end{aligned} \tag{8}$$

where $\ell_{p,s}^{2000}$ is the employment share of sector s in prefecture p in 2000, $R_{s,t}$ is the IFR robot stock in sector s in China at time t, and $L_{s,2000}$ is employment in that sector in 2000. We use China's current stock directly rather than subtracting the 2000 stock, because China's 2000 robot stock is approximately zero (930 robots installed nationwide); the paper therefore treats China's 2000 exposure as a zero-exposure pre-robot baseline. The 2000–2015 endpoint long difference, which uses the change in exposure rather than its level, gives the same result (Table 6).[5] The IV shifter $G_{s,t}^{EU}$ is the average robot density by sector-year across the 14 EU15 countries for which the IFR reports robot stocks (all but Luxembourg, which the IFR does not report separately): IFR operational robot stock divided by EU KLEMS employment, in robots per thousand workers, using the observed 2000 European density for the base wave. The IFR manufacturing robot stock with unspecified sector assignments in the early years is allocated across sectors using the shares of stocks with identified sectors, and 2005 is retained as a low-exposure wave. The coefficient in the levels specification is reported for a one-standard-deviation increase in the exposure **level** ($\widetilde{\exp}$), standardized to zero mean and unit variance on the four-wave estimation sample; the long-difference estimate is reported for a one-standard-deviation increase in the five-year exposure **change**. Figure 2 shows the timing and sectoral concentration of robot adoption: adoption in China is near zero in 2000 and remains low in 2005, before accelerating after 2010. It is concentrated in automobiles, electronics, plastics/chemicals, and metals. Figure 3 maps the resulting 2000–2015 change in prefecture exposure: exposure growth is concentrated in the Pearl River

[5] On the 3,239 senior-high cells matched to the sector-exposure source, replacing zero China exposure in 2000 with observed sector stocks changes the IV estimate from −4.80 to −4.81 percentage points, about 0.012 points. The endpoint long difference changes by about 0.0014 points. This comparison holds the matched sample fixed; the difference between this estimate and the headline estimate based on all 3,497 cells is not attributed to the stock convention.

Delta, the Yangtze River Delta, and the automotive and electronics hubs of the northeast, Beijing–Tianjin, Wuhan, and Chongqing, with most of the interior near zero.

Cell-level specification. For the core results, we collapse each census outcome to prefecture–census-year–age cells and relate the in-school rate to the level of robot exposure across the four waves. For age a in prefecture p and wave $t \in \{2000, 2005, 2010, 2015\}$,

$$\bar{Y}_{p,t,a} = \beta \widetilde{\exp}_{p,t} + \phi_p + \eta_t + \alpha_a + u_{p,t,a}, \tag{9}$$

where $\widetilde{\exp}_{p,t}$ is robot exposure in China standardized to unit variance, instrumented with the corresponding European shifter (both constructed in equation (8) above), and ϕ_p, η_t, α_a are prefecture, wave, and age fixed effects. Identification uses within-prefecture exposure changes relative to common national time trends; differential provincial trends may still contribute to identification. We estimate (9) separately within each of three schooling stages—primary (ages 7–11), junior high (ages 12–14), and senior high (ages 15–17)—so β is the stage-specific in-school-rate coefficient in the levels specification. The primary-stage response may also reflect household budget constraints, migration, and expectations about future progression; the paper's concern is not whether the primary slope equals zero, but whether the secondary-stage coefficients are substantially more negative than the primary-stage coefficient. The main specification weights cells equally; the robustness section compares weighting by raw sample cell counts and by counts normalized within each year and age.

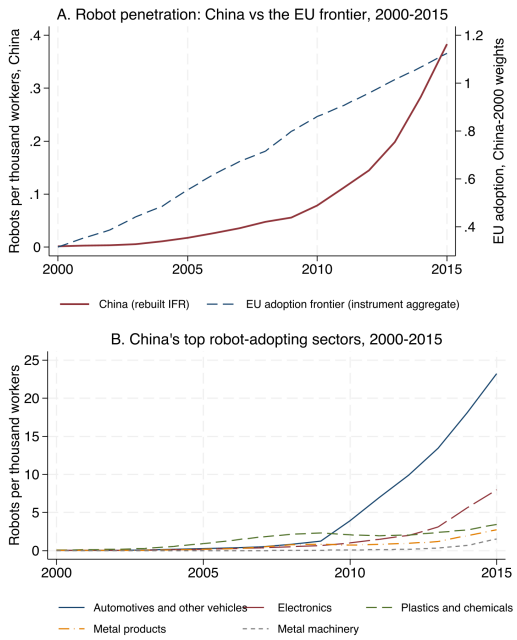


Figure 2. *Industrial robot penetration, 2000–2015.*

Notes: Panel A: robots in China per thousand workers employed in 2000 and the EU shifter aggregated with China's 2000 sector shares. Panel B: penetration in the five Chinese IFR sectors with the highest robot penetration.

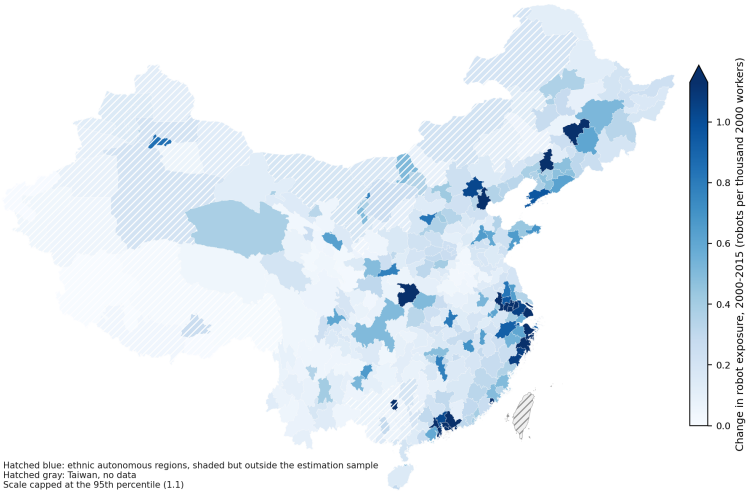


Figure 3. *Change in prefecture-level robot exposure, 2000–2015.*

Notes: Change in the exposure measure of equation (8) between 2000 and 2015, in robots per thousand 2000 workers; because China's 2000 stock is essentially zero, the change equals the 2015 level. Darker shading indicates higher exposure; the scale is capped at the 95th percentile. The five ethnic autonomous regions are shaded but hatched because they lie outside the estimation sample; Taiwan has no data.

Pre-expansion admission chances. The provincial heterogeneity test (Section 6.3) uses a predetermined measure of access to further education: the 1998 provincial admissions-to-graduates ratio, defined as provincial regular-college admissions divided by provincial regular senior-high graduates in the 1998 volume of the China Education Statistical Yearbook. 1998 is the last year before the 1999 expansion, and the expansion increased admissions in proportion to each province's existing capacity, so the ratio captures both pre-expansion access to college and the room for post-expansion growth in college access. Provincial admissions sum to 1,083,627, matching the printed yearbook.

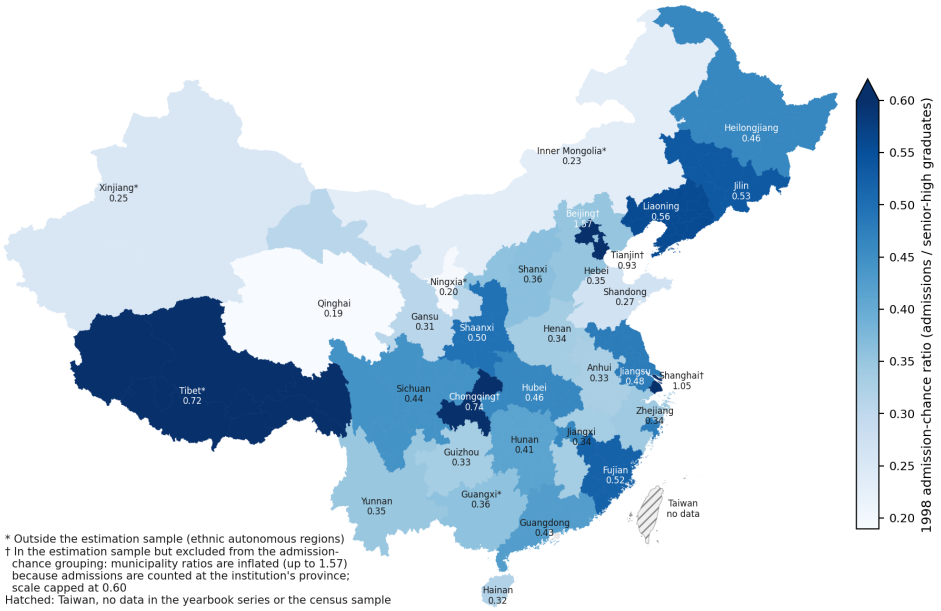


Figure 4. *Provincial admission chances on the eve of the 1999 expansion: the ratio of regular-college admissions to regular senior-high graduates in 1998 (Section 3).*

Notes: Starred: outside the estimation sample; daggered: municipalities excluded from the grouping; hatched: Taiwan, no data.

Individual-level overview. Table 1 summarizes the pooled sample at the individual level: roughly 3.11 million children aged 7–17 with a known hukou type across the four census waves. Nearly eighty percent hold rural hukou; the overall in-school rate is 88.6%, falling from 97.8% at primary ages to 66.0% at senior-high ages; among 15–17-year-olds, 24.6% are already employed. Among children living in family households, only 2.5% live with a household member holding a bachelor's degree or above (6.8% when the threshold is a junior-college education or higher).

Table 1. *Individual-level summary statistics, pooled 2000–2015 censuses.*

	Mean	SD	N
Age in years	11.98	3.00	3,113,188
Female	0.472	0.499	3,113,188
Rural hukou	0.783	0.412	3,113,188
Han ethnicity	0.928	0.259	3,113,188
Lives in a collective household	0.016	0.125	3,113,188
In school	0.886	0.317	3,113,188
In school, ages 7–11	0.978	0.147	1,408,094
In school, ages 12–14	0.939	0.239	922,725
In school, ages 15–17	0.660	0.474	782,369
Employed, ages 15–17	0.246	0.431	782,369
Child with a linked parent record	0.870	0.336	3,150,814
Linked as grandchild of the head (share of linked)	0.073	0.260	2,742,691
Parents' education: senior high or above	0.231	0.422	2,742,691
Parents' education: junior college or above	0.053	0.224	2,742,691
Household size	4.35	1.60	3,150,814

Notes: Children aged 7–17 with nonmissing hukou type. Rows below the rule use the file of children living in family households of Section 6.4; the grandchild row and the two parental-education rows report shares among children with a linked parent record. Statistics are unweighted.

Baseline schooling levels. Table 2 reports prefecture-year cell means pooled across hukou. Robot exposure is essentially zero in 2000, still low in 2005, and rises sharply by 2015.

Table 2. *Summary statistics by census wave.*

	2000	2005	2010	2015
Robot exposure (raw)	0.000	0.014	0.064	0.309
Youth observations per prefecture-year	6,925.2	1,401.1	1,613.8	732.6
In-school rate, ages 7–11	0.970	0.973	0.983	0.955
In-school rate, ages 12–14	0.916	0.941	0.959	0.938
In-school rate, ages 15–17	0.551	0.715	0.785	0.818
In-school rate, ages 18–21	0.103	0.260	0.368	0.472
Prefecture-year cells	292	292	290	292
Underlying youth observations	2,022,156	409,116	468,009	213,907

Notes: Prefecture–year cells pooled across hukou. Schooling rows are in-school rates for ages 7–11, 12–14, 15–17, and 18–21.

4 Main Results: The Decline in School Enrollment Is Concentrated at the Senior-High Gateway

4.1 Levels evidence: the stage gradient

Age profile. Figure 5 shows age-specific coefficients from separate levels regressions, pooling observations across hukou groups. The coefficient is −1.58 percentage points at age 7 and −5.05 at age 17. The primary-age estimates are not monotonic, but the declines generally become larger across secondary-school ages and are largest at senior-high ages.

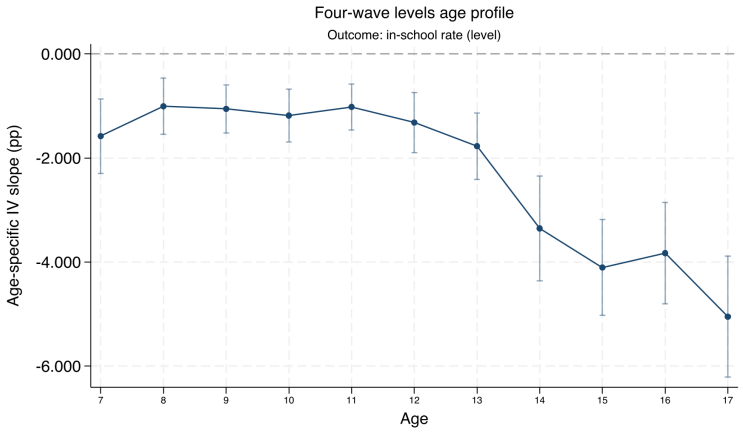


Figure 5. *Robot exposure and the in-school rate by age: levels estimates pooled across hukou groups.*

Notes: Each age-a coefficient $\hat{\beta}_a$ is estimated in a separate 2SLS regression $\bar{Y}_{p,t,a} = \beta_a \exp_{p,t} + \phi_p + \eta_t + \varepsilon_{p,t,a}$, instrumented with the EU shifter. Whiskers indicate 95% confidence intervals.

Stage-by-stage levels. Table 3 aggregates ages into schooling stages. The 2SLS coefficient is −1.17 percentage points for primary (ages 7–11), −2.15 for junior high (ages 12–14), and increases in absolute magnitude to −4.33 for senior high (ages 15–17). OLS gives the same stage ordering.

This age gradient has a clear economic interpretation. Younger children are still far from immediate school-leaving and labor-market entry; compulsory-schooling constraints and uncertainty about academic prospects keep most of them away from the continuation-versus-exit margin. As age rises, students learn more about their own performance and progression prospects, and leaving school for work becomes progressively more feasible. The same labor-market shock to the future returns to education and work is therefore more likely to translate into observable changes in school enrollment at older ages. The end of compulsory schooling and the minimum working age both fall in this age range: the closer

teenagers are to the actual school-leaving margin, the stronger the negative relationship. This is the intuition of the dropout ladder in Section 2: the farther a student is from the gate, the less is known about whether progression is reachable, the larger the waiting option contained in the value of staying, and the fewer students can exit into work; as the gate approaches, information arrives, the waiting option is exhausted, and work becomes accessible, so the same exposure shock induces more students to leave school.

Table 3. *Effects on the in-school rate by schooling stage: four-wave levels estimates pooled across hukou groups.*

	Robot exposure (1 SD), pp			
Schooling stage	OLS	2SLS	First-stage F	Obs
Primary 7–11	-0.81***	-1.17***	485	5,828
	(0.16)	(0.24)		
Junior high 12–14	-1.43***	-2.15***	484	3,495
	(0.25)	(0.35)		
Senior high 15–17	-2.70***	-4.33***	484	3,497
	(0.35)	(0.44)		
Fixed effects	Prefecture, wave, and age FE			
Cell unit	Prefecture × wave × age (pooled across hukou)			
Waves	2000, 2005, 2010, and 2015			
Weights	Equal cells			

Notes: Each row is a separate regression of the in-school rate on the level of robot exposure using prefecture × wave × age cells; 2SLS instruments the China exposure level with the European shifter level. Coefficients are reported in percentage points per one-standard-deviation increase in exposure; standard errors are clustered by prefecture. Later tables follow these conventions unless noted.

4.2 Post-schooling destinations: weaker routine employment prospects rather than factory hiring

Destinations after exposure. Table 4 decomposes the population in the ages 15–17 cells into four mutually exclusive categories: in school, out of school in agriculture, out of school

in non-agricultural work, and out of school not working. Operator occupations and manufacturing are both included in non-agricultural employment. The in-school decline of −4.33 percentage points is accompanied by an increase in the agricultural employment share of 5.33 and a change in the non-agricultural employment share of −1.16. The non-employment (0.16) and operator (0.24) coefficients are not statistically significant; manufacturing employment falls significantly (−0.85). The increase in agricultural employment is the clearest destination result; an increase in non-employment is not precisely estimated.

Table 4. *Post-exit destinations at ages 15–17.*

		Out of school, by destination			Of which (non-agricultural work)	
	In school	Agricultural work	Non-agricultural work	Other / idle	Operator (line work)	Manufacturing
	(1)	(2)	(3)	(4)	(5)	(6)
Robot exposure (per 1 SD)	-4.33***	5.33***	-1.16***	0.16	0.24	-0.85***
	(0.44)	(0.50)	(0.33)	(0.25)	(0.18)	(0.30)
Mean outcome	0.705	0.126	0.073	0.096	0.011	0.035
Observations	3,497	3,497	3,497	3,497	3,497	3,497
First-stage F	484.4	484.4	484.4	484.4	484.4	484.4
Ages	15–17; pooled across hukou and sex					
Fixed effects	Prefecture, year, and age FE					
Weights	Equal cells					

Notes: Each column is a separate 2SLS regression of a status share on the level of robot exposure, using the cells in Table 3. The three out-of-school shares in columns (2)–(4) sum to one minus the in-school share, so their coefficients sum to the negative of the coefficient in column (1); columns (5)–(6) are subsets of column (3).

These results do not support the creation of new non-agricultural jobs as the main explanation for students leaving school. Unlike the employment-pull mechanisms in Atkin (2016) and Charles et al. (2018), agricultural employment shares rise while non-agricultural employment falls. This is consistent with a weakening of the routine employment route reached through further schooling. The coefficients are net population-share changes, however: they neither trace individual transitions nor establish robots as the only cause.

These destinations are cross-sectional statuses at the census date, not individual transitions. The operator and manufacturing shares are small on average (1.1% and 3.5%), so small

positive coefficients cannot support a "factory pull" interpretation; and age 15 is below the legal minimum working age, so "agricultural work" at that age is mostly within-household labor that the census records as employment.

5 Identification Threats and Robustness

The exclusion restriction of Section 3 requires that 2000 sector shares affect schooling only through robots. This section addresses, in turn, pre-existing differential trends correlated with the geographical distribution of sectors, sample composition and migration, and the mechanics of the shift-share instrument itself; the main tests are reported in Table 5, and the supplementary migration and inference tables are in Appendix A. The section closes with the endpoint long-difference confirmation. The senior-high estimates remain negative across these specifications, although their magnitude and precision vary with the controls and weights used.

Fixed effects and timing. Column (1) of Table 5 reproduces the baseline with prefecture, wave, and age fixed effects (-4.33); column (2) adds province-by-wave fixed effects (-3.75). Both use 3,497 cells, isolating the fixed-effect change from sample composition. Province-by-wave fixed effects absorb common provincial education-policy and economic changes, which the baseline does not absorb. Changing the baseline specification does not by itself distinguish the effects of robots from those of other industrial shocks or subsequent offsetting policies.

Sector exclusions and industrial trends. Table 5 focuses on the senior-high ages (15–17). Columns (3)–(4) reconstruct the exposure measure and instrument after excluding electronics and automobiles, respectively, testing whether the gradient is driven by a single major sector. Column (5) adds 1990 prefecture characteristics interacted with year indicators. All seven characteristics come from the 1990 census micro-data and are determined before the arrival of robots: log population, the male share, the working-age (15–64) share, college attainment and secondary-or-higher attainment among those aged 25 and over, and the manufacturing and agricultural shares of employment; each is standardized across prefectures and interacted with year dummies, absorbing differential trends correlated with pre-period demographics, education stocks, and industrial structure. Column (6) addresses convergence directly: it adds the prefecture's own 2000 in-school rate at ages 15–17 interacted with year indicators. Prefectures where robot exposure rose most started from higher enrollment, and enrollment converged across prefectures over the period, so the two-way fixed-effects design attributes part of the smaller enrollment gains in high-exposure prefectures to exposure; the interactions allow each year's trend to depend on the baseline level. The coefficient falls to -1.29 percentage points (standard error 0.47) and remains significant. Column (6) is the lagged-dependent-variable analogue of the fixed-effects design in column (1), and Angrist and Pischke (2009, Section 5.3) show that the two estimates bracket the causal effect: the fixed-effects estimate overstates it when exposure responds to the baseline level of the outcome, and the baseline-control estimate understates it when the fixed-effects model is correct. We therefore read the senior-high effect as lying between 1.29 and 4.33 percentage points per standard deviation of exposure, with column (1) as the upper bound.

Table 5. *Robustness of the pooled in-school-rate coefficient at senior-high ages, 15–17.*

	(1)	(2)	(3)	(4)	(5)	(6)
In school, ages 15–17	-4.33***	-3.75***	-4.86***	-4.21***	-3.22***	-1.29***
	(0.44)	(0.39)	(0.63)	(0.48)	(0.61)	(0.47)
Observations	3,497	3,497	3,269	3,269	3,107	3,353
First-stage F	484.4	520.7	1781.4	149.3	699.6	354.0
Mean outcome	0.705	0.705	0.702	0.702	0.701	0.704
SD outcome	0.214	0.214	0.216	0.216	0.216	0.214
Prefecture, year, and age FE	Yes	Yes	Yes	Yes	Yes	Yes
Province × year FE	No	Yes	No	No	No	No
Excluding electronics from exposure	No	No	Yes	No	No	No
Excluding automotives from exposure	No	No	No	Yes	No	No
1990 characteristics × year	No	No	No	No	Yes	No
2000 in-school rate (15–17) × year	No	No	No	No	No	Yes
Weights	Equal cells					

Notes: Column (1) uses prefecture, wave, and age fixed effects; column (2) adds province × wave fixed effects. Columns (3)–(4) reconstruct the exposure measure and instrument excluding electronics and automobiles; columns (5)–(6) add 1990 prefecture characteristics × year and the prefecture's 2000 in-school rate at ages 15–17 × year. Each column changes one feature of column (1). Columns (1) and (6) bracket the causal effect in the sense of Angrist and Pischke (2009, Section 5.3).

Migration. Appendix Table A2 restricts the sample to youth whose residence five years earlier was local or in a comparable location within the same province, testing sensitivity to migration restrictions: the primary-age response is close to zero, while the in-school-rate coefficient at senior-high ages (15–17) remains −3.94 percentage points.

Instrument mechanics. Three diagnostics probe the shift-share construction itself. Rotemberg weights (Appendix Table A3) show that the identifying variation is concentrated in automobiles and electronics (about 52.4%), with the top five sectors accounting for 87.1%. Leave-one-sector-out estimates (Appendix Table A4) reconstruct both the exposure measure and the instrument after dropping each of the 18 IFR sectors in turn; the senior-high slope stays negative throughout. Cell-weighting choices do not change the conclusion either: the

senior-high stage shows the largest decline under equal cell weights (-4.33), raw sample cell counts (-7.03), and counts normalized within each year and age (-3.13) alike (Appendix Table A1).

Shock-level and province-level inference. Appendix Table A5 implements sector-shock inference following Borusyak et al. (2022). Both this analysis and the Rotemberg decomposition use the same sample of five-year changes for ages 15–17, with period fixed effects, yielding -3.72 percentage points. The heteroskedasticity-robust shock-level standard error is 1.82, with $p = 0.046$; the estimate remains statistically significant when standard errors are clustered by sector. The estimate therefore only just reaches statistical significance at the 5% level under the most conservative test, indicating substantially less precision than prefecture-clustered inference suggests. These are difference-based diagnostics, not an exact decomposition of the levels coefficient in Table 3; they also differ from the endpoint change and fifteen-year scaling of Table 6. Residuals may be correlated across prefectures in a province under either design. Appendix Table A6 therefore uses province-clustered standard errors and a wild cluster bootstrap with Webb weights and 9,999 replications. For the senior-high levels specification, the province-clustered standard error is 0.52 and the bootstrap p-value is < 0.001; for the endpoint long difference they are 0.86 and < 0.001. Bootstrap p-values across all three stages and both designs are at most 0.02.

5.1 The 2000–2015 long-difference confirmation

To place the levels result alongside more conservative change-based evidence, we re-estimate the same stage gradient with a 2000–2015 endpoint long difference. Let $\Delta\bar{Y}_{p,a} \equiv \bar{Y}_{p,2015,a} -$

$\bar{Y}_{p,2000,a}$ be the change in the in-school rate at age a in prefecture p, and let $\Delta\widetilde{\text{exp}}_p$ be the 2000–2015 change in robot exposure, standardized to unit variance across prefectures. On prefecture–age cells, separately within each schooling stage, we estimate

$$\Delta\bar{Y}_{p,a} = \beta^{\text{LD}}\,\Delta\widetilde{\text{exp}}_p + \psi_{\text{prov}(p)} + \alpha_a + u_{p,a}, \tag{10}$$

where $\Delta\widetilde{\text{exp}}_p$ is instrumented with the standardized 2000–2015 change in the European shifter, $\psi_{\text{prov}(p)}$ and α_a are province and age fixed effects, cells are weighted equally, and standard errors are clustered by prefecture. Identification comes from exposure changes across prefectures within the same province. This design tests the same gradient with fifteen-year changes rather than four-wave levels, a conservative confirmation of the levels result.

Table 6 reports estimates of (10) by schooling stage. The differenced estimates reproduce the same ordering: -0.66 percentage points for primary (ages 7–11), -1.73 for junior high, and -6.53 for senior high.

Table 6. *Cumulative effects on the in-school rate, 2000–2015: long-difference estimates pooled across hukou groups.*

	Δ robot exposure, 2000–2015 (1 SD), pp			
Schooling stage	OLS	2SLS	First-stage F	Obs
Primary 7–11	-0.61**	-0.66**	463.7	1,320
	(0.24)	(0.26)		
Junior high 12–14	-1.46***	-1.73***	458.7	792
	(0.41)	(0.41)		
Senior high 15–17	-6.18***	-6.53***	458.7	792
	(0.66)	(0.70)		
Fixed effects	Province FE and age FE, within stage			
Cell unit	Prefecture × age (pooled across hukou)			
Difference	2015 minus 2000			
Weights	Equal cells			

Notes: Each row regresses the 2015-minus-2000 change in the in-school rate on the 2015-minus-2000 change in robot exposure, using prefecture × age cells, with the exposure change instrumented by the change in the European shifter. Coefficients are in percentage points per one-standard-deviation increase in the exposure change.

Dose response. Tables 3 and 6 constrain the response to be linear in exposure. Figure 6 relaxes that constraint: Panel A sorts prefectures into quintiles of the 2000–2015 change in the European shifter and estimates, stage by stage, the change in the in-school rate of each quintile relative to the bottom one; Panel B is a binned scatter of the same relationship. The senior-high ages (15–17) show a clear dose-response gradient: the second and third quintiles are statistically indistinguishable from the bottom quintile, the fourth-quintile coefficient is −5.88 percentage points, and the top-quintile coefficient falls sharply to −16.33 points, with the bin means in Panel B falling monotonically along the linear fit. The pattern differs at primary-school ages (7–11): the coefficients for the third through fifth quintiles are about 3 percentage points below the bottom-quintile reference (top quintile −3.00), but there is no gradient across them—the small negative primary coefficient of Table 6 reflects a discrete step relative to the least-exposed prefectures, rather than a response that becomes more

negative with exposure. At junior-high ages, the response becomes more negative across the first three quintiles (top quintile −7.80) and then levels off.

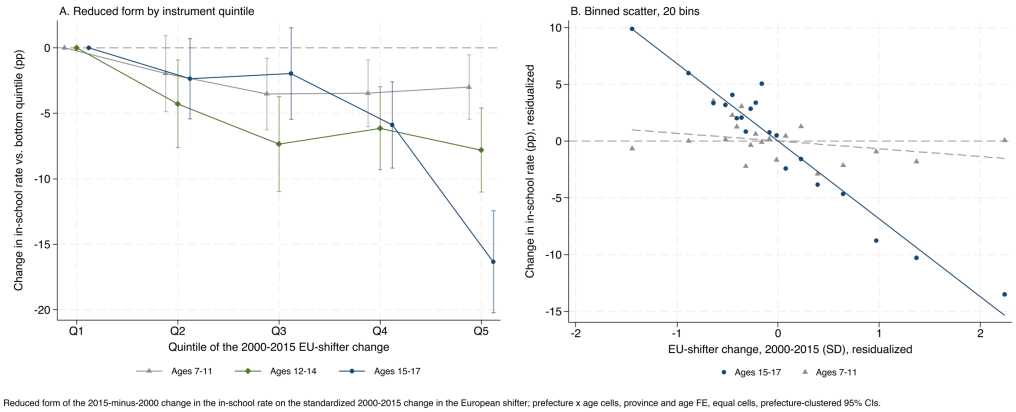


Figure 6. *Dose-response patterns by schooling stage: 2000–2015 long differences.*

Notes: Panel A: reduced-form coefficients on quintiles of the standardized 2000–2015 change in the European shifter (bottom quintile omitted), one regression per stage with province and age fixed effects; whiskers indicate 95% confidence intervals. Panel B: binned scatter of the same outcome on the same instrument change, with both variables residualized with respect to the fixed effects and grouped into 20 equal-count bins.

6 Heterogeneity in the Enrollment Response: Hukou and Sex

This section reports 2SLS comparisons across the three schooling stages for each grouping. Each table shows the two group coefficients, their signed difference, and the p-value for a test of equality. Parentheses contain prefecture-clustered standard errors from the joint regression.

6.1 Hukou: larger rural declines

To estimate differences by hukou, we re-estimate the levels specification separately for each schooling stage using prefecture–census-year–hukou–age cells, interacting both the exposure variable and the fixed effects with hukou:

$$\bar{Y}_{p,t,h,a} = \beta \widetilde{\exp}_{p,t} + \rho \left(\widetilde{\exp}_{p,t} \times \text{rural}_h \right) + \phi_{p \times h} + \eta_{t \times h} + \alpha_{a \times h} + u_{p,t,h,a}. \tag{11}$$

For each stage, β is the non-rural in-school-rate coefficient and ρ is the rural-minus-non-rural gap. Table 7 reports both group coefficients and their difference at primary, junior-high, and senior-high ages. The rural decline is significantly larger at all three stages, with differences already present during compulsory schooling.

Table 7. *Effects on the in-school rate by schooling stage and hukou: levels estimates.*

	2SLS: robot exposure (1 SD), pp		
	Primary 7–11 (1)	Junior high 12–14 (2)	Senior high 15–17 (3)
Non-rural	-0.33**	-0.46**	-1.70***
	(0.13)	(0.19)	(0.37)
Rural	-1.20***	-2.39***	-3.60***
	(0.32)	(0.61)	(0.53)
Rural–non-rural	-0.87**	-1.93***	-1.89***
	(0.37)	(0.63)	(0.53)
Difference p-value	0.019	0.002	< 0.001
Observations (cells)	11,629	6,964	6,977
First-stage F	242	241	242
Fixed effects	Prefecture × hukou, wave × hukou, age × hukou		
Cell unit	Prefecture × wave × hukou × age		
Waves	2000, 2005, 2010, and 2015		
Weights	Equal cells		

Notes: Prefecture × wave × hukou × age cells; exposure slopes and prefecture, wave, and age fixed effects are interacted with hukou. The difference is rural minus non-rural.

Both the non-rural (-1.70 percentage points) and rural (-3.60 points) senior-high coefficients are large and significant; the rural–non-rural gap is -1.89 points and is statistically significant at the 1% level. One possible explanation is that rural students' parents are more likely to work in occupations vulnerable to automation, so parental employment and earnings losses may tighten household financial constraints more severely. Another is that rural students may have lower chances of entering high-quality universities, making the expected returns to continued schooling more dependent on the routine employment pathway and their school-leaving decisions more responsive to robot exposure.

6.2 Sex: similar senior-high declines but larger junior-high declines among girls

The senior-high response is broadly similar for boys and girls; at junior-high ages the female decline is larger, and the difference is statistically significant at the 1% level. Table 8 re-estimates the stage-by-stage levels design on prefecture–census-year–sex–age cells, using the analogue of (11) for sex:

$$\bar{Y}_{p,t,s,a} = \beta\,\widetilde{\exp}_{p,t} + \rho\left(\widetilde{\exp}_{p,t} \times \text{female}_s\right) + \phi_{p\times s} + \eta_{t\times s} + \alpha_{a\times s} + u_{p,t,s,a}, \tag{12}$$

where female_s is an indicator for females, the exposure slope and all fixed effects are interacted with sex, and both $\widetilde{\exp}_{p,t}$ and its interaction are instrumented with the European shifter and its interaction. For each stage, β is the male coefficient and ρ is the female-minus-male difference. At senior high (ages 15–17), the male slope is -4.29 percentage points and

the female slope −4.32, with a female-minus-male difference of −0.03 points that is not statistically significant.

Table 8. *Effects on the in-school rate by schooling stage and sex: levels estimates.*

	2SLS: robot exposure (1 SD), pp		
	Primary 7–11 (1)	Junior high 12–14 (2)	Senior high 15–17 (3)
Male	-1.08***	-1.80***	-4.29***
	(0.25)	(0.34)	(0.45)
Female	-1.25***	-2.51***	-4.32***
	(0.27)	(0.38)	(0.47)
Female–male	-0.17	-0.71***	-0.03
	(0.15)	(0.18)	(0.31)
Difference p-value	0.242	< 0.001	0.919
Observations (cells)	11,649	6,987	6,991
First-stage F	242	242	242
Fixed effects	Prefecture × sex, wave × sex, age × sex		
Cell unit	Prefecture × wave × sex × age		
Waves	2000, 2005, 2010, and 2015		
Weights	Equal cells		

Notes: Prefecture × wave × sex × age cells; the exposure slope and all fixed effects are interacted with sex.

The intuition that "automation should hit boys harder" draws on the male-dominated automobile plants of Western labor markets. However, the sectors accounting for China's robot adoption during this period—besides automobiles, also electronics and textiles (Appendix Table A3)—employ many young women in routine assembly-line work. The "secondary-schooling–routine-work" pathway depressed by automation is therefore not confined to men, and the enrollment response is correspondingly present for both sexes rather than concentrated among boys facing factory employment prospects.

6.3 Pre-expansion admission chances: weaker exposure responses where access to further education was greater

Provinces with higher 1998 admissions-to-graduates ratios also saw larger post-expansion growth in college access (Section 3), so their teenagers faced greater access to further education (K higher). By the attenuation result in Section 2, exposure-driven school-leaving should concentrate in low-chance provinces, with the response in high-chance provinces approaching zero rather than turning positive.

Table 9 splits the 22 provinces into the top and bottom quartiles of the 1998 admissions-to-graduates ratio (six per tail) and compares their slopes separately at primary, junior-high, and senior-high ages. Let low_p equal one for prefectures in a bottom-quartile province; on the prefecture–census-year–age cells of the twelve provinces in these two quartiles we estimate, within each stage,

$$\bar{Y}_{p,t,a} = \beta\,\widetilde{\exp}_{p,t} + \rho\left(\widetilde{\exp}_{p,t} \times \text{low}_p\right) + \left(\widetilde{\exp}_{p,t} \times Z_p\right)'\gamma + \phi_p + \eta_{t\times\text{low}_p} + \alpha_{a\times\text{low}_p} + u_{p,t,a}, \quad (13)$$

where the prefecture fixed effect ϕ_p absorbs group membership, the wave and age fixed effects are interacted with group indicators, and exposure and each of its interactions are instrumented with the European shifter and the corresponding interactions. β is the top-quartile slope, $\beta + \rho$ the bottom-quartile slope, and ρ the bottom-minus-top difference. Panel A sets $\gamma = 0$; Panel B lets Z_p contain four 1990 prefecture characteristics. Each panel reports both group coefficients and the bottom-minus-top difference with its standard error and p-value.

Table 9. *Effects on the in-school rate by schooling stage and 1998 provincial admission opportunities: top and bottom quartiles, levels estimates.*

	2SLS: robot exposure (1 SD), pp		
	Primary 7–11 (1)	Junior high 12–14 (2)	Senior high 15–17 (3)
Panel A: baseline			
Top-quartile provinces	0.03	-0.62*	-3.96***
	(0.15)	(0.36)	(0.83)
Bottom-quartile provinces	-4.63***	-6.51***	-5.22***
	(1.67)	(1.70)	(1.04)
Bottom–top difference	-4.67***	-5.89***	-1.26
	(1.68)	(1.73)	(1.33)
Difference p-value	0.006	< 0.001	0.346
Observations (cells)	2,460	1,476	1,476
First-stage F	818	817	817
Panel B: 1990 characteristics × exposure			
Top-quartile provinces	1.61	1.33	-3.17**
	(1.31)	(1.51)	(1.51)
Bottom-quartile provinces	-4.25***	-5.79***	-4.54***
	(1.57)	(1.59)	(1.17)
Bottom–top difference	-5.86***	-7.12***	-1.37
	(2.21)	(2.33)	(1.66)
Difference p-value	0.009	0.003	0.411
Observations (cells)	2,460	1,476	1,476
First-stage F	50	50	50
Fixed effects	Prefecture, wave × group, age × group		
Cell unit	Prefecture × wave × age		
Waves	2000, 2005, 2010, and 2015		
Weights	Equal cells		

Notes: The 22 provinces are classified by the top and bottom quartiles of the 1998 admissions-to-graduates ratio (six per tail; municipalities excluded, Section 3); the exposure slope and all fixed effects are interacted with the group. Panel B adds exposure interactions with four 1990 prefecture characteristics (agricultural employment share, college and secondary attainment stocks, log population), with all exposure interactions instrumented.

The two groups of provinces may also differ in other characteristics. To keep such differences from driving the contrast, Panel B adds the term $\widetilde{\exp}_{p,t} \times Z_p$ in (13), where Z_p

contains the four 1990 characteristics (agricultural employment share, college and secondary attainment stocks, log population).

Both specifications yield consistent estimates: robot exposure has a smaller negative effect on school enrollment in provinces with greater pre-expansion admission opportunities, which also experienced larger subsequent increases in college access. The difference between the two groups is statistically significant at primary and junior-high ages, although it is not statistically significant at senior-high ages.

6.4 Effects on intergenerational mobility

The previous subsections split the sample by hukou, sex, and provincial admission chances. This subsection turns to the intergenerational dimension within the family: does robot exposure change the extent to which a child's enrollment depends on parental education? By the logic of Section 2, better-educated parents are better positioned to help their children access the college track, so the option value of continuing rests less on the routine work that automation devalues. If so, school-leaving should concentrate among children of less-educated parents, and the enrollment gap by parental education should widen with exposure.

Our research design follows Black and Devereux (2011), Chetty et al. (2014), and Bennett et al. (2026). We use a two-step approach: within each area and child cohort, estimate the coefficient from a regression of the child's educational outcome on the father's education, and then regress that coefficient on the labor-demand shock. Bennett et al. examine college completion by age 30; here we examine enrollment at ages 15–17 rather than completed education, capturing the first observable intergenerational enrollment gap at the senior-high gateway.

Parents are matched inside the household roster (children and grandchildren of the household head), and parents' highest education is the higher of the father's and the mother's attainment on a six-step ladder harmonized across waves; collective households and children whose parents cannot be identified are excluded. The indicator of interest is $D_i = \mathbf{1}\{\text{parents' highest education is junior college or above}\}$. The persistence this indicator captures is strong: pooling the four waves, 95.9% of 15–17-year-olds whose parents have junior college or above are in school, against 63.2% of the remaining children, and the gap is positive in 96.8% of the 1149 prefecture–waves. Whether the parents reached higher education thus strongly predicts whether the child is still in school at the senior-high gateway; the question is whether robot exposure widens this gap.

Within each prefecture p and wave t, the first step estimates

$$\text{InSchool}_{ipt} = \alpha_{pt} + \gamma_{pt} D_i + \sum_{a=16}^{17} \xi_{a,pt} \mathbf{1}\{\text{age}_i = a\} + e_{ipt}.$$

γ_{pt} is the age-adjusted difference in enrollment probabilities between the two groups, capturing the component of children's enrollment associated with parental tertiary education; it is the parent–child regression slope in binary form, and estimating this slope area by area as a measure of relative mobility is standard in the intergenerational-mobility literature. The second step tests whether γ_{pt} moves with exposure using a panel 2SLS regression with prefecture and wave fixed effects:

$$\gamma_{pt} = \mu_p + \nu_t + \beta_P \widetilde{R}_{pt} + u_{pt},$$

where $\widetilde{R}_{pt}$ is standardized robot exposure with the same instrument as before, and $\beta_P > 0$ means exposure widens the enrollment advantage conferred by parental education. Each wave observes children aged 15–17, so the comparison is across parent–child associations for different child birth cohorts. The regressions use equal weights for each prefecture–wave cell, and standard errors are clustered by prefecture; every prefecture–wave in which both groups are observed is included. To see which group drives any change in the gap, the same second-step regression is estimated using the two unadjusted group enrollment rates S^0_{pt} and S^1_{pt} as outcomes.

Table 10. *Parents' higher education and children's enrollment at ages 15–17.*

	Age-adjusted enrollment gap	In-school rate, ages 15–17	
	γ_{pt}	$D_i = 0$	$D_i = 1$
	(1)	(2)	(3)
Parents' highest education is junior college or above			
Junior college or above		No	Yes
Robot exposure (1 SD), 2SLS	1.44***	-3.40***	-1.97***
	(0.47)	(0.40)	(0.39)
Robot exposure (1 SD), OLS	1.15***	-2.63***	-1.47***
	(0.38)	(0.33)	(0.32)
2000 level of the dependent variable (pp)	40.6	54.9	93.9
Prefecture–waves	1,130	1,130	1,130
First-stage F	479	479	479
Step 1	Column (1): in-school on D_i and age dummies		
Fixed effects	Prefecture, wave		
Unit	Prefecture × wave		
Weights	Equal cell weights (prefecture–wave)		
Waves	2000, 2005, 2010, and 2015		

Notes: The indicator D_i equals one if the parents' highest education (the higher of the father's and the mother's attainment, matched inside the household roster) is junior college or above, and zero otherwise. Column (1) uses the within-prefecture–wave enrollment gap adjusted for age dummies; columns (2)–(3) use unadjusted group enrollment rates. Each outcome is regressed on robot exposure with prefecture and wave fixed effects, using equal prefecture–wave cell weights, with standard errors clustered by prefecture.

The baseline gap in 2000 is already large: the average prefecture-level enrollment rate is 93.9% among 15–17-year-olds whose parents' highest education is junior college or above, compared with 54.9% among the remaining children. The age-adjusted gap is 40.6 percentage points. Column (1) of Table 10 shows that a one-standard-deviation increase in exposure widens this gap by a further 1.44 points (standard error 0.47, significant at 1%), about 4% of the baseline. Columns (2)–(3) show which group drives the widening gap: enrollment falls significantly in both groups, by 3.40 points among children whose parents have less than junior-college education and by 1.97 points among those whose parents have junior college or above. The gap widens not because better-educated families gain but because children from less-educated families experience larger enrollment declines.

7 Conclusion

This paper examines how robots affect educational choices at the pre-college stages and finds a negative relationship between robot exposure and the in-school rate that becomes stronger toward the transition from junior to senior high school: −1.17 percentage points at ages 7–11, −2.15 at ages 12–14, and −4.33 at ages 15–17, with the endpoint long difference reproducing the same ordering. This is not a reversal of the schooling expansion: the senior-high-age in-school rate rose from 55.1% to 81.8% over the same period, and more-exposed prefectures experienced smaller enrollment gains, forgoing 6.53 points of that rise. Destination shares primarily show an increase in agricultural employment and a decline in non-agricultural employment. This does not support new factory jobs as the main factor drawing students out of school. The senior-high coefficient remains negative across leave-one-sector-out estimates, controls for differential trends related to pre-period characteristics, migration restrictions, alternative weighting, Rotemberg diagnostics, and province-level inference, but the estimate only just reaches significance at the 5% level under the most conservative shock-level test, and the wave-specific timing pattern is sensitive to province-by-wave controls.

Two complementary heterogeneity analyses link the gradient to access to further education. At the province level, access is measured by admission chances in 1998, just before the 1999 expansion. The school-leaving response is stronger on average where access to further education was more limited, although both groups show significant declines at senior-high ages. At the family level, children whose parents have less than junior-college education experience larger enrollment losses: each standard deviation of exposure widens the age-adjusted enrollment gap at ages 15–17 between children whose parents have junior college

or above and the rest by 1.44 percentage points. Rural-hukou teenagers also have larger declines, with differences already significant at junior-high ages. These patterns are consistent with access to further education cushioning automation-induced school-leaving.

These findings highlight the importance of expanding educational opportunities and easing household financial constraints to help disadvantaged students adapt to skill-biased technological change. Higher education expansion can strengthen the incentive to remain in secondary school by making further study a more attainable option. Combining broader access to higher education with support for disadvantaged families can help students acquire the skills needed to benefit from technological change and reduce the risk that automation reinforces educational disadvantage across generations.

Acknowledgments

We thank participants at the CUHK-Shenzhen Summer School and the CUHK Applied Microeconomics Reading Group for valuable discussions, especially Yiheng Liu, Ziyu Peng, Anqi Wu, and Haocheng Yong. We also thank Julius Ilciukas, Junsen Zhang, Yifan Zhang, and Ce Matthew Shi for their insightful comments and generous feedback. All errors are our own.

Data availability statement. The code that reproduces every table, figure and in-text number, together with the public inputs and a synthetic version of the census microdata, is deposited in the Economic Journal replication repository on Zenodo. The Chinese population census and one-percent survey microdata for 2000, 2005, 2010 and 2015 (NBS, 2000, 2005, 2010, 2015) are restricted and were obtained for research use; they cannot be redistributed, and

the replication README describes how to obtain them. The 1990 prefecture characteristics come from an IPUMS International extract (IPUMS International, 2024), the robot stocks from the IFR World Robotics database (IFR, 2024), the European employment denominators from the EU KLEMS 2019 release (EU KLEMS, 2019), the 1998 admission chances from the China Education Statistical Yearbook (Ministry of Education, 1999), and the 2000 sector employment shares, trade controls and map boundaries from the replication package of Giuntella et al. (2024).

A Appendix: Robustness Tables

Table A1. *Sensitivity of pooled levels estimates to weighting: effects on the in-school rate by schooling stage.*

	Equal cells	Raw counts	Within-year
	(1)	(2)	(3)
Primary 7–11	-1.17***	-0.97***	-0.77***
	(0.24)	(0.18)	(0.16)
Junior high 12–14	-2.15***	-2.23***	-1.46***
	(0.35)	(0.33)	(0.24)
Senior high 15–17	-4.33***	-7.03***	-3.13***
	(0.44)	(0.83)	(0.57)
Estimator	2SLS pooled levels in-school slope; pp per 1 SD exposure level		

Notes: Each entry reports a stage-specific 2SLS coefficient under the indicated weighting scheme. Let n_{pat} denote the number of individuals with nonmissing schooling status in prefecture p, age a, and census year t, after sample restrictions. Equal cell weights set $w_{pat} = 1$. Raw counts set $w_{pat} = n_{pat}$, giving greater weight to cells with more sampled individuals; these are observed sample counts, not population expansion weights. Within-year normalization sets $w_{pat} = n_{pat} / \sum_q n_{qat}$, where the denominator sums over prefectures in the same census year and at the same age. Thus, each year–age group has total weight one, while prefectures within that group remain weighted in proportion to their sample counts. This prevents larger census samples from mechanically receiving greater total weight across years. All three schemes enter the regressions as analytic weights.

Table A2. *Robustness to migration restrictions by schooling stage: levels estimates pooled across hukou groups.*

	Primary	Junior high	Senior high
	7–11	12–14	15–17
	(1)	(2)	(3)
Migration-restricted	-0.08**	-1.28***	-3.94***
	(0.03)	(0.21)	(0.37)
Observations	5,398	3,238	3,239
First-stage F	448.9	448.8	448.8
Fixed effects	Prefecture, wave, and age FE		
Cell unit	Prefecture × wave × age (pooled across hukou)		
Estimator	2SLS levels; coefficients in pp; negative means lower enrollment		
Weights	Equal cells		

Notes: The sample includes youth whose residence five years earlier was local or in a comparable location within the same province.

Table A3. *Rotemberg weights for the European-adoption Bartik instrument: pooled five-year-difference estimates for the in-school rate at ages 15–17, residualized with respect to period fixed effects.*

Sector	Rotemberg weight α_k	2000 emp. share (%)	EU shock change ending 2015	$\hat{\beta}_k$ (pp)
Automotive and other vehicles	0.346	0.69	11.1	-4.32
Electronics	0.178	1.00	1.5	-4.94
Industrial machinery	0.149	0.99	4.0	0.19
Textiles	0.108	2.45	0.8	-3.64
Plastics and chemicals	0.091	1.26	1.2	-6.95
Sum of top-5 weights		0.871		
Sum of negative weights (1 sectors)		-0.043		
Effective number of sectors ($1/\sum_k \alpha_k^2$)		4.9		
IV estimate $\sum_k \alpha_k \hat{\beta}_k$ (pp)		-3.72		
Weights		Equal cells		

Notes: α_k is the sector weight and $\hat{\beta}_k$ is the just-identified IV estimate.

Table A4. *Leave-one-sector-out robustness of the pooled in-school-rate coefficient at senior-high ages, 15–17.*

Specification	Coef. (pp)	SE	First-stage F
All sectors	-4.33***	(0.44)	484.4
Agriculture, forestry, fishing	-4.90***	(0.43)	455.1
Mining and quarrying	-4.83***	(0.43)	454.4
Construction	-4.80***	(0.43)	449.5
Food and beverages	-4.73***	(0.42)	466.1
Textiles	-4.75***	(0.44)	412.5
Wood and furniture	-4.78***	(0.43)	443.1
Paper and printing	-4.80***	(0.43)	445.9
Plastics and chemicals	-4.59***	(0.45)	370.8
Glass and ceramics	-4.78***	(0.43)	451.6
Basic metals	-4.77***	(0.43)	438.3
Metal products	-4.71***	(0.45)	396.0
Industrial machinery	-5.00***	(0.47)	418.9
Electronics	-4.86***	(0.63)	1781.4
Automotive and other vehicles	-4.21***	(0.48)	149.3
Electricity, gas, water	-4.80***	(0.43)	448.5
Education and R&D	-4.80***	(0.43)	448.4
Other manufacturing	-4.84***	(0.44)	475.3
Other non-manufacturing	-4.80***	(0.43)	448.2
Negative leave-one-out estimates	18 of 18 sectors		
Fixed effects	Prefecture, year, and age FE		
Weights	Equal cells		

Notes: The first row is the all-sector baseline; each following row re-estimates the slope after dropping the named IFR sector from both the China exposure and the European shifter. Prefecture-clustered standard errors are in parentheses; the last column reports the Kleibergen–Paap first-stage F-statistic.

Table A5. *Shock-level inference for the pooled five-year-difference effect on the in-school rate at ages 15–17, residualized with respect to period fixed effects.*

	In school
Senior-high-window exposure effect (pp)	-3.72
Prefecture-clustered SE (268 clusters)	(0.36)***
Shock-level robust SE (53 shocks)	(1.82)**
Shock-level sector-clustered SE (18 sectors)	(0.80)***
Prefecture-period cells	796
Sector × wave shocks (nonzero)	53
Weights	Equal cells

Notes: Shock-level standard errors follow Borusyak et al. (2022).

Table A6. *Province-level inference for the stage gradient.*

	Coefficient (pp)	SE, prefecture clusters	SE, province clusters	Wild cluster bootstrap p (province clusters)
	(1)	(2)	(3)	(4)
Panel A: four-wave levels (Table 3); 311 prefectures, 26 provinces				
Primary 7–11	-1.17	(0.24)***	(0.44)**	0.008
Junior high 12–14	-2.15	(0.35)***	(0.63)***	0.005
Senior high 15–17	-4.33	(0.44)***	(0.52)***	<0.001
Panel B: 2000–2015 long difference (Table 6); 264 prefectures, 22 provinces				
Primary 7–11	-0.66	(0.26)**	(0.32)*	0.017
Junior high 12–14	-1.73	(0.41)***	(0.59)***	0.004
Senior high 15–17	-6.53	(0.70)***	(0.86)***	<0.001
Bootstrap	Wild cluster restricted-efficient bootstrap, Webb weights, 9,999 replications			
Fixed effects	Panel A: prefecture, year, age; Panel B: province, age			
Weights	Equal cells			

Notes: Rows report re-estimated 2SLS coefficients by schooling stage of Table 3 (Panel A) and Table 6 (Panel B) with standard errors clustered by prefecture and by province, along with the p-value from a province-level wild cluster bootstrap (Webb weights, 9,999 replications). Stars refer to the corresponding clustered t-test.

Affiliations

[1]Department of Economics, The Chinese University of Hong Kong

[2]Department of Economics and Finance, City University of Hong Kong

References

Abramitzky, R. and V. Lavy (2014). How Responsive Is Investment in Schooling to Changes in Redistributive Policies and in Returns? *Econometrica* 82(4), 1241–1272.

Acemoglu, D. and P. Restrepo (2018). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review* 108(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy* 128(6), 2188–2244.

Acemoglu, D. and P. Restrepo (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica* 90(5), 1973–2016.

Adão, R., M. Kolesár, and E. Morales (2019). Shift-Share Designs: Theory and Inference. *Quarterly Journal of Economics* 134(4), 1949–2010.

Adukia, A., S. Asher, and P. Novosad (2020). Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction. *American Economic Journal: Applied Economics* 12(1), 348–376.

Altonji, J. G. (1993). The Demand for and Return to Education When Education Outcomes Are Uncertain. *Journal of Labor Economics* 11(1), 48–83.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Atkin, D. (2016). Endogenous Skill Acquisition and Export Manufacturing in Mexico. *American Economic Review* 106(8), 2046–2085.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics* 118(4), 1279–1333.

Becker, G. S. (1975). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, 2nd ed. New York: National Bureau of Economic Research.

Bennett, P., K. Liu, and K. G. Salvanes (2026). Structural Change, Education Investment, and Intergenerational Mobility. Working paper, June 30, 2026 version.

Bertrand, M., M. Mogstad, and J. Mountjoy (2021). Improving Educational Pathways to Social Mobility: Evidence from Norway's Reform 94. *Journal of Labor Economics* 39(4), 965–1010.

Black, D. A., T. G. McKinnish, and S. G. Sanders (2005). Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust. *Industrial and Labor Relations Review* 59(1), 3–16.

Black, S. E. and P. J. Devereux (2011). Recent Developments in Intergenerational Mobility. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 4B, 1487–1541.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-Experimental Shift-Share Research Designs. *Review of Economic Studies* 89(1), 181–213.

Branco, D., B. Carrillo, and W. Iglesias (2025). Routine-Biased Technological Change and Endogenous Skill Investments. *American Economic Journal: Economic Policy* 17(3), 236–280.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2018). Housing Booms and Busts, Labor Market Opportunities, and College Attendance. *American Economic Review* 108(10), 2947–2994.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics* 129(4), 1553–1623.

Chuan, A. and W. Zhang (2025). Non-College Occupations, Workplace Routinization, and Female College Enrollment. *Journal of Human Resources* 60(3).

Dauth, W., S. Findeisen, J. Südekum, and N. Wößner (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association* 19(6), 3104–3153.

Di Giacomo, G. and B. Lerch (2026). Automation and Human Capital Adjustment: The Effect of Robots on College Enrollment. *Journal of Human Resources* 61(3), 733–760.

Stehrer, R., A. Bykova, K. Jaĝer, O. Reiter, and M. Schwarzhappel (2019). *EU KLEMS 2019 Release: Statistical National Accounts* [dataset]. Vienna: The Vienna Institute for International Economic Studies (wiiw) for the European Commission. `https://euklems.eu`

Giuntella, O., Y. Lu, and T. Wang (2024). Replication Package for: "How Do Workers Adjust to Robots? Evidence from China" [dataset]. Zenodo. `https://doi.org/10.5281/zenodo.13760908`

Giuntella, O., Y. Lu, and T. Wang (2025). How Do Workers Adjust to Robots? Evidence from China. *Economic Journal* 135(666), 637–652.

Goldin, C. and L. F. Katz (2008). *The Race between Education and Technology*. Cambridge, MA: Belknap Press of Harvard University Press.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review* 110(8), 2586–2624.

Greenland, A. and J. Lopresti (2016). Import Exposure and Human Capital Adjustment: Evidence from the U.S. *Journal of International Economics* 100, 50–60.

Gunadi, C. and H. Ryu (2023). How Do People Respond When They Know That Robots Will Take Their Jobs? *Oxford Bulletin of Economics and Statistics* 85(5), 939–958.

Hannum, E. (1999). Political Change and the Urban–Rural Gap in Basic Education in China, 1949–1990. *Comparative Education Review* 43(2), 193–211.

Heyman, F. and M. Olsson (2026). Long-Run Effects of Technological Change: The Impact of Automation on Intergenerational Mobility. *Economic Journal*, forthcoming. `https://doi.org/10.1093/ej/ueag093`

International Federation of Robotics (2024). *World Robotics: Industrial Robots, Operational Stock by Country, Industry and Year, 1993–2023* [dataset]. Frankfurt am Main: IFR Statistical Department.

Minnesota Population Center (2024). *Integrated Public Use Microdata Series, International: Version 7.5, China 1990 Census Extract* [dataset]. Minneapolis, MN: IPUMS. `https://doi.org/10.18128/D020.V7.5`

Jensen, R. (2010). The (Perceived) Returns to Education and the Demand for Schooling. *Quarterly Journal of Economics* 125(2), 515–548.

Kang, R. (2026). The Dynamic Effects of Robot Exposure on Human Capital Investment and Labor Market Outcomes. Job market paper. https://sites.google.com/view/rihyeon-kang/research

Li, B. (2018). Export Expansion, Skill Acquisition and Industry Specialization: Evidence from China. *Journal of International Economics* 114, 346–361.

Li, H., P. W. Liu, and J. Zhang (2012). Estimating Returns to Education Using Twins in Urban China. *Journal of Development Economics* 97(2), 494–504.

Liu, Y., Y. Guo, Z. Li, B. Wang, and J. You (2025). The Human Capital Promotion Effect of Industrial Intelligence: Theory and Evidence. Working paper, SSRN No. 5553525. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5553525

Loyalka, P., C. Liu, Y. Song, H. Yi, X. Huang, J. Wei, L. Zhang, Y. Shi, J. Chu, and S. Rozelle (2013). Can Information and Counseling Help Students from Poor Rural Areas Go to High School? Evidence from China. *Journal of Comparative Economics* 41(4), 1012–1025.

Manski, C. F. (1989). Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon. *Economics of Education Review* 8(4), 305–312.

Ministry of Education of the People's Republic of China, Department of Development and Planning (1999). *China Education Statistical Yearbook 1998* [Zhongguo jiaoyu tongji nianjian 1998]. Beijing: People's Education Press.

National Bureau of Statistics of China (2000). *Fifth National Population Census of China, Long-Form Microdata Sample* [dataset]. Beijing: NBS. Restricted access.

National Bureau of Statistics of China (2005). *2005 One-Percent National Population Sample Survey, Microdata* [dataset]. Beijing: NBS. Restricted access.

National Bureau of Statistics of China (2010). *Sixth National Population Census of China, Microdata Sample* [dataset]. Beijing: NBS. Restricted access.

National Bureau of Statistics of China (2015). *2015 One-Percent National Population Sample Survey, Microdata* [dataset]. Beijing: NBS. Restricted access.

Stange, K. M. (2012). An Empirical Investigation of the Option Value of College Enrollment. *American Economic Journal: Applied Economics* 4(1), 49–84.

Stinebrickner, T. and R. Stinebrickner (2012). Learning about Academic Ability and the College Dropout Decision. *Journal of Labor Economics* 30(4), 707–748.

Tuhkuri, J. (2023). The Surprising Intergenerational Effects of Manufacturing Decline. Working paper, February 2023 version. `https://joonastuhkuri.com/pdf/children.pdf`

Willis, R. J. and S. Rosen (1979). Education and Self-Selection. *Journal of Political Economy* 87(5), S7–S36.

World Bank (2026). School Enrollment, Tertiary (% Gross), Indicator SE.TER.ENRR. *World Development Indicators*. `https://data.worldbank.org/indicator/SE.TER.ENRR?end=2024&locations=US-CN&start=1970&view=chart`, data version 2026-07-13, accessed 2026-08-18.

Yi, H., Y. Song, C. Liu, X. Huang, L. Zhang, Y. Bai, B. Ren, Y. Shi, P. Loyalka, J. Chu, and S. Rozelle (2015). Giving Kids a Head Start: The Impact and Mechanisms of Early

Commitment of Financial Aid on Poor Students in Rural China. *Journal of Development Economics* 113, 1–15.

Zhang, J. and K. Zhou (2023). Quota Removal, Destination-Specific Export Shocks, and Skill Acquisition in China. *Journal of Development Economics* 165, 103149.

Zhang, R. (2026). Higher Education Expansion and the Rural–Urban Gap in Secondary Schooling and Labor Outcomes: Evidence from China's 1999 Reform. *Journal of Development Economics* 181, 103721.