COLUMBIA | SIPA
Center for Development Economics and Policy

CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 44

Does primary school duration matter? Evaluating the consequences of a large Chinese policy experiment

Alex Eble and Feng Hu

July 2018

COLUMBIA | SIPA
Center on Global Economic Governance

Does primary school duration matter? Evaluating the consequences of a large Chinese policy experiment

Alex Eble and Feng Hu*

July 2018

Abstract

Nearly all governments provide primary schooling, but surprisingly little is known about how changes to the duration of primary school affect educational attainment and performance in the labor market. We study a Chinese policy which extended the duration of primary school from five years to six but did not change the curriculum. We exploit its gradual rollout over space and time to generate causal estimates of its impact on educational attainment and subsequent labor market outcomes. We find that the policy has small, largely positive effects on post-primary educational attainment, and raises average monthly income by 2.6%. The policy is progressive, generating higher returns (5-8%) among both women and the least educated. We estimate the policy has already reallocated 450 million years of labor from work to schooling and we generate cost-benefit estimates to quantify this tradeoff, highlighting the large public finance implications of this policy decision.

*Eble (corresponding author): Teachers College, Columbia University. Address: Department of Education Policy and Social Analysis. 525 W 120th St, New York, NY 10027, USA. Email: eble@tc.columbia.edu; Hu: Dongling School of Economics and Management, University of Science and Technology Beijing. Email: feng3hu@gmail.com. We would like to thank Andrew Foster, Emily Oster, and John Tyler for extensive feedback and guidance, and Alexei Abrahams, Anna Aizer, Dionissi Aliprantis, Marianna Battaglia, Natalie Bau, Nate Baum-Snow, Ken Chay, Andrew Elzinga, John Friedman, David Glancy, Nate Hilger, Rob Jensen, Melanie Khamis, Eoin McGuirk, Bryce Millett-Steinberg, Sri Nagavarupu, Matthew Notowidigdo, Gareth Olds, Anja Sautmann, Rajiv Sethi, Jesse Shapiro, Miguel Urquiola, Felipe Valencia, Josh Wilde, Zach Sullivan, and many seminar audiences for helpful suggestions. Eble acknowledges support from the US National Science Foundation through Graduate Research and IGERT Fellowships, and from the Brown University PSTC. Hu acknowledges financial support from the National Natural Science Foundation of China (71373002, 71420107023). Previous versions of this paper were circulated as "The Importance of Educational Credentials: Schooling Decisions and Returns in Modern China." Keywords: compulsory education; human capital; redistribution; China. JEL codes: I24, I25, I26, J24.

1 Introduction

Nearly all countries regulate the duration of government-sanctioned education. The literature on compulsory schooling has studied the impacts of changes to the age at which students are permitted to leave schooling and enter the labor force. This work has found substantial gains in health, wealth, and longevity to increasing the age at which students are first allowed to leave school (e.g., Angrist and Krueger, 1991; Oreopoulos, 2007; Devereux and Hart, 2010). Another key aspect of this regulation is the policy decision setting the number of years of schooling needed to finish primary, secondary, and tertiary education.

Because of differentials in access and cost between levels of schooling, the duration of each level of schooling, and particularly primary school, has a large potential impact both on how many total years students spend in school and on their later earning power. Examining Demographic and Health Surveys (DHS) datasets for the 74 developing countries where years of schooling data is collected, we find bunching in the number of completed years of education at those years where a level of schooling is completed - usually primary and lower secondary - in 48 of these[1]. While recent work has begun to study the impacts of changes to school duration at the secondary and tertiary levels (Morin, 2013; Arteaga, 2018), no work that we are aware of studies this at the primary level, the most common stopping point we see in our DHS data.

In this paper, we study a policy in China which extended the duration of primary school from five years to six while holding the national curriculum and duration of all other levels of schooling unchanged. We exploit the fact the policy was rolled out gradually over space and time to generate three sets of empirical results. First, we measure how the extra year of primary school affects overall educational attainment. Second, we estimate how the policy affects performance in the labor market. Third, we generate estimates of the public finance implications of the policy which, to date, has reallocated more than 450 million years of labor[2] from employment to the pursuit of schooling.

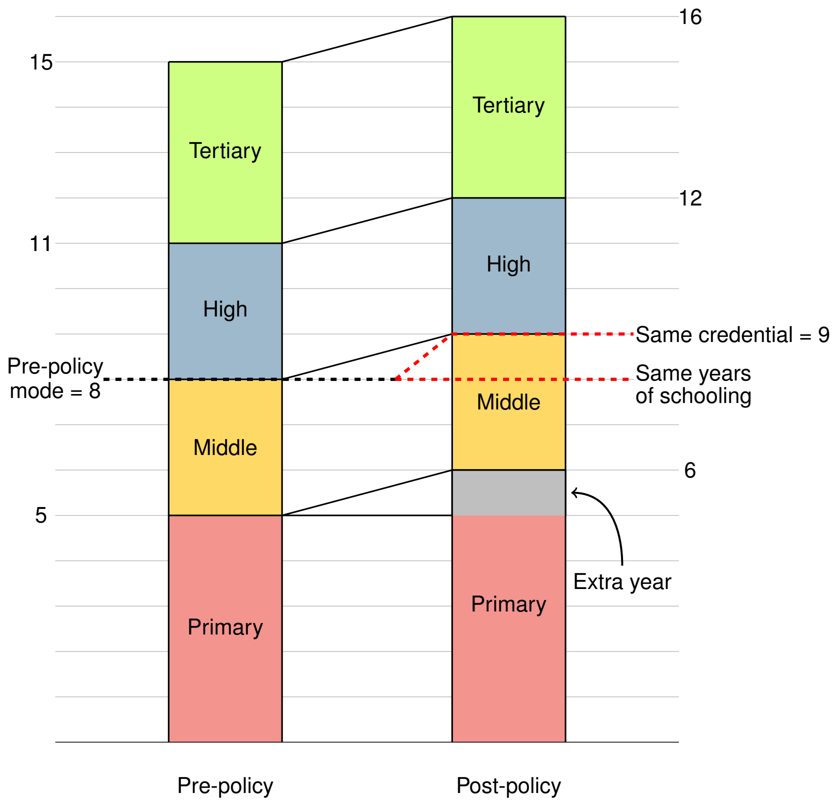
[1] These patterns are also common in developed countries. According to CPS data, 73 percent of 30-64 year olds in the US spend exactly as many years in school as is needed to earn either a secondary or post-secondary credential.

[2] Calculation given in footnote 5.

In 1980 the Chinese government announced a policy to increase the minimum number of years of schooling needed to complete primary school by one. The stated purpose of the policy was to better educate students in order to aid in China's transition towards a modern (i.e., market) economy and society. This policy left unchanged the national curriculum and duration of all other levels of schooling, meaning that the primary curriculum, originally designed to be taught over five years, was instead taught more gradually over a six year period. The policy thus changed two important aspects of education: one, it increased by 20% the amount of time students spent learning the primary curriculum; two, it changed the parameters of the schooling decision students faced. Specifically, to earn any given credential (primary, middle, high, or tertiary), affected students must spend an additional year in school and thus out of the labor market[3]. Figure 1 depicts the Chinese education system before and after the policy change.

[3]Though middle school was made compulsory in 1986, in Appendix 4 we show evidence that the rollout of the compulsory middle school policy has little effect on whether or not individuals in our sample complete middle school or earn at least a middle school credential.

Figure 1: Years of schooling to earn credentials, pre- and post-policy



This figure depicts the duration and structure of the Chinese education system before the policy is implemented in the left column and after the policy is implemented in the right column. The y-axis represents the number of years needed to complete a credential. Middle, high, and tertiary refer to junior high school, senior high school, and university, respectively. The thick lines connecting the left and right columns depict the change in the total number of years it takes to earn each credential as a result of the policy. The dashed red lines show the trade-off that the modal student faces after the policy.

This policy differs importantly from those studied in prior work on compulsory education and the returns to schooling in the labor market. Most previous studies (e.g., those summarized in Card, 1999) use changes in compulsory schooling which also add to the set of skills students are meant to acquire. For example, both Angrist and Krueger (1991) and Oreopoulos (2006) study the impact of policies which induce students to spend an additional year in high school and, in so doing, advance further in the high school curriculum. This constitutes an increase in skills on the extensive margin. Our policy induces all students to spend six years learning the same material that unaffected students were made to learn in five, which instead constitutes an increase in skills on the intensive margin.

We identify the causal effect of the Chinese six year primary education policy on schooling and labor market outcomes using a stacked regression discontinuity (RD) design[4]. Our running/forcing variable is month-by-year-of-birth (as in Oreopoulos, 2006, and Erten and Keskin, 2018, among several others). We recenter this variable around the policy implementation year in a given locality, and stack cutoffs (as in Pop-Eleches and Urquiola, 2013, and Abdulkadiroğlu et al., 2014) across different policy implementation years.

We determine if, when, and how the policy is implemented in each of China's prefectures by collecting and coding thousands of official government documents, known as "educational gazetteers," which report implementation at the local level, and match this to new survey and census data from China. We restrict our sample to the 280 prefectures (of 345) where we are sure the policy was implemented. Our identification strategy compares outcomes of treated and untreated individuals within each of these prefectures, using only those leaving primary school within an optimal bandwidth around the year when the policy took effect. We also generate parallel results using a difference in differences specification. As we are unable to generate a satisfactory test of the parallel trends assumption and the language of the policy suggests that this assumption is likely to be violated (details discussed in section 3.3), we present these results in Appendix 5. They largely mirror the results we present in the body of the paper.

We first estimate the impact of this policy on years of schooling and the attainment of educational credentials. We find that the average number of years of completed primary schooling increased from 5.2 to nearly six for affected individuals. We find small changes in the proportion of individuals getting middle school and high school degrees, but no evidence that the policy changed the family background characteristics (parents' education, number of siblings) of who earns which credential.

We then estimate the impact of the policy on labor market performance. We find that the policy does not affect individuals' likelihood of being employed, but makes them slightly more likely (1.8

[4]This could also be called an event study; here we use the RD label, as we use the machinery for establishing causal inference in that literature (Imbens and Lemieux, 2008; Lee and Card, 2008; Lee and Lemieux, 2010).

percentage points, or 6.7 percent) to work in government. Our main labor market result is that the policy increases monthly income by 2.63 percent, with about 20 percent of this gain coming from the increase in credentials, and the other 80 percent coming from income gains conditional on highest credential earned. The magnitude and precision of our estimate of the policy's impact on income are robust to a battery of specification and robustness checks, including a permutation test. We then show that the magnitude of our estimate changes little after accounting for a set of potential mediators, including experience differentials between the treated and untreated and general equilibrium effects.

Despite these low overall returns, we find the policy is redistributive. While it induces nearly all to forgo a year of earnings, its benefits are concentrated among a key disadvantaged group in China: those with only a primary or middle school education, whose income increases by between five and eight percent as a result of the policy. We argue that the basic cognitive and non-cognitive/socio-emotional skills reinforced in this extra year are most likely to help those with lower skill levels. Furthermore, our finding is consistent with other work showing that extending the time children spend in school most helps struggling students (Meghir and Palme, 2005; Dobbie and Fryer, 2013; Clay et al., 2016).

A primary impact of this policy is the reallocation of human resources from work to school. We estimate that the policy has reallocated more than 450 million years, or more than 900 billion person-hours, from the labor market to the pursuit of schooling to date[5]. We perform a rough cost-benefit analysis to quantify the public finance implications of this massive reallocation of resources. We borrow the framework used in Duflo (2001) to estimate the lifetime value of the increase in monthly income conferred by the extra year of schooling and compare it to the estimated cost of the lost year of productive activity in the labor market for affected individuals, calculated over the period 1981-2050. The sign of these estimates hinges on how we assess the policy's effects in

[5]China's educational yearbooks estimate that 332,321,868 children graduated with six years of primary school between 1984 and 2009. The number of students leaving primary school between 2010-2017 under the six-year regime, assuming negligible drop-out from primary school, is 120,060,612. The number graduating under this regime between 1981 and 1984 is not listed in the yearbooks. Using the proportions given in Figure A.1, we estimate that it is likely to be no more than a few million students. We assume individuals spend 2,000 hours working in a year. We multiply the number of affected individuals to date by the year of lost labor hours each forgoes to generate an estimate of approximately 905 billion person-hours of labor reallocated to schooling between 1981 and 2017.

rural areas. Our preferred specification finds that the policy generates an overall net gain, though its benefits decrease over time with the decline in the per-cohort proportion of individuals with low levels of education, i.e., those who benefit most.

Our paper contributes to two active lines of inquiry. First, we add to the rich literature attempting to understand the education and labor market effects of changes to compulsory education policy (Card, 1999; Stephens and Yang, 2014). We advance this work, and the nascent literature on understanding the impact of changes to the duration of schooling, in a few key ways. First, as mentioned at the beginning of this section, the policy we study constitutes an increase in skills on the intensive margin, as opposed to on the extensive margin studied in the vast majority of previous work. Second, ours is the first paper we are aware of to study a change in the duration of primary school, as opposed to changes in the duration of secondary or tertiary schooling as in Morin (2013) and Arteaga (2018). Third, the policy changes the duration of school for all students, not just a subgroup, and has already affected the lives of hundreds of millions of individuals.

The second line of inquiry we contribute to is the wider set of studies using large changes in education policy from developing countries to assess the merits of different policy options and their distributional effects (e.g., Duflo 2001; Banerjee et al. 2007; Lucas and Mbiti 2012). To this work, we add the first analysis of a key policy lever - the duration of primary school - and show that it has potentially important redistributive properties. This is particularly pertinent for policymakers in the many developing countries with a sizable proportion of individuals who do not progress beyond primary or lower secondary school.

The rest of the paper proceeds as follows. In Section 2, we discuss the history of education in modern China and describe the policy we study. In Section 3, we describe the data we use and our identification strategy. Section 4 contains empirical results related to educational attainment and Section 5 provides our empirical results relating to the labor market, a series of robustness checks, a discussion of potential mechanisms, and our cost-benefit analysis. Section 6 concludes.

2 Setting

This paper studies the education system in China after the Cultural Revolution ended in 1976. China's system resembles those of the US and many other developed and developing countries, with primary school, middle school, high school, and then tertiary education[6]. The structure of the system prior to and after the policy we study is shown visually in Figure 1. In this section, we provide more detail on our setting and on the purpose and details of the policy we study.

In the aftermath of the chaos caused by China's Cultural Revolution (1966-1976), the country moved to standardize its education system. To this end, the *Full-Time Ten-Year Primary and Middle Education Teaching Plan (Draft)* was passed in January 1978. This mandated national harmonization of the duration and structure of primary, middle, and high school in all provinces, setting the duration of primary school to be five years in schools across the country. At the end of 1980, the Central Committee of the Communist Party of China and State Council issued the *Decision on Several Problems Relating to Universal Primary Education*, the policy whose changes we use for our analysis. This policy mandated that the total years of primary and secondary education be extended to twelve years, including a shift from five to six years of primary school[7]. This policy was announced early on in Deng Xiaoping's time as China's de facto leader and at the beginning of the country's transition from a planned to a more market-oriented economy. One of Deng's early directives was that education should "face the demands of the new era and meet head-on the challenges of the technological revolution," part of his larger move in the late 1970s and early 1980s to prepare China's labor force to adapt to this new economic arrangement (Vogel, 2011).

The letter of the law allowed gradual adoption of the primary school duration change across localities, with the central government putting more initial pressure on urban schools (Liu, 1993). In practice, roughly 60 percent of localities switched to a six year system between 1981 and 1993, relatively few made the change in the mid-1990s, and the rest shifted in the late 1990s and 2000s,

[6] In China, the levels of schooling known as middle school and high school in the US are referred to as junior middle school and senior middle school. We refer to them here as middle school and high school for ease of exposition.

[7] In Shanghai and a few other localities, this policy was implemented instead by requiring that middle school last four years instead of the usual three.

reaching near-universal adoption in 2005. Figure A.1 plots national data on the proportion of students in six year (or equivalent) primary school systems, showing this gradual adoption of the policy over time.

This policy usually did not change the age at which children entered school, nor did it change the primary school, middle school, or high school curricula. Rather, the intent of the change was that primary school students be taught the same material over a longer period of time to ensure mastery of the curriculum. To collect a richer account of what happened during this extra year and how it was perceived, we conducted a focus group discussion in Shandong Province with a team of middle school teachers and former students who were affected by the policy. Participants reported that in the first year or two after the reform, the content of the extra year consisted of a review of what was covered in the fifth year of primary school and the addition of elective courses, such as physical education and music. The teachers reported that, after this adjustment process, the primary school curriculum previously covered in five years was extended more smoothly over six years. In practice, this meant more time allowed for review and ensuring the foundational concepts of the primary curriculum were mastered by all students. Generally, respondents felt the extra year was most likely to have helped those of lower ability. More than half the respondents mentioned the loss of a year of productive work as the main downside of the policy.

The extra year of schooling posed logistical and personnel challenges. The policy required primary schools to hold and teach an additional cohort of children, but these schools were given no additional resources to do so. Gazetteer records and our interviews indicated that the burden of housing and schooling this extra cohort in a given primary school involved assigning more work to existing teachers and dividing up existing facilities, as opposed to building new structures and hiring new staff. When asked about the effect of this extra burden on the quality of education, respondents generally thought it unlikely to have a substantial impact. This claim is consistent with the rote nature of Chinese primary education during this time, which we argue is likely to dampen a possible negative relationship between class size and learning. As we document later in the paper, this additional burden was gradually offset by a secular decline in cohort sizes over time.

The gazetteers document that the transition from five to six years of primary school was carried out in a number of ways. In Table A.1, we show six examples of gazetteers reports' of how the policy was enacted in different implementing cities and counties. In some cases, the transition was accomplished by enforcing the policy fully, forcing all students in affected cohorts to remain in primary school an extra year. In other instances, a portion of the exiting cohort of students was sent on to middle school after their fifth year of primary school while the rest remained to finish a sixth year. This practice was explained in the gazetteers as a method to smooth the flow of students during the first year or two of transition, after which all subsequent cohorts would then take six years.

The decision of when to implement the policy was made at the local level. Though upper-level pressure certainly played a factor, as we discuss in Section 3, most counties had the ultimate say on the year in which the switch was made[8]. In the next section, we describe how our identification strategy avoids the need to have random timing of implementation, and we address the other issues surrounding discretion in timing of implementation and the attendant concerns of omitted variable bias.

3 Data and identification strategy

This section describes the data sources and empirical methods of the paper. We show evidence that the main identifying assumptions for our research design are satisfied, and address a set of issues which could confound causal interpretation of our results.

3.1 Data sources

Our sources of data are listed in Table A.2. There are two main sources of observational data: the 2005 China 1% Inter-censal National Population Sample Survey (also known as the "mini-census") and the 2010 wave of the China Family Panel Studies (CFPS). The 2005 Chinese mini-census collects basic data on family structure, highest educational credential attained, health, and

[8]Local educational gazetteers document that, in most cases, counties within a prefecture implemented the policy in the same year or within a few years of each other.

income, and contains 2.6 million observations[9]. The CFPS is a nationally representative panel data set containing information from over 30,000 individuals in rural and urban China across 25 provinces, representative of 94.5% of China's population[10]. Summary statistics on demographic, education, and employment characteristics of our sample population are given in Table A.3 for each data set, separately for rural and urban residents.

Most of our main results come from the Census data. We include the CFPS as a complement, as it allows us to show a more detailed picture of how the policy was implemented within a given locality. We then use the large sample size of the census data to generate precise estimates of the coefficients we are after: the effect of the policy on educational attainment and labor market outcomes.

3.2 Determining treatment status

We also collect our own data from two sets of national archives to determine which observations in the census data were affected by the policy we study. The shift to six year primary school was implemented at different times in different places both across and within China's provinces, as shown in Figure A.2. We hired a team of research assistants to read through prefecture and county educational gazetteers stored in the Chinese National and Peking University Archives, to determine if, when, and how the policy was implemented in each locality[11]. Figure A.3 shows a page from one of these gazetteers.

In China, the largest geographical unit is the province, followed by the prefecture, followed by the county, then the township and, finally, the village. The census data is at the resolution of the prefecture, and we determine the year the policy was implemented, separately for rural and urban residents, in 280 of the 345 prefectures in the Census. The gazetteers document that

[9]Though the full sample is approximately 13 million observations, researchers are granted access to 20% sub-samples of the parent dataset.

[10]The data include all provinces but Tibet, Xinjiang, Inner Mongolia, Hainan, and Ningxia. The CFPS is conceived of as a panel, with six waves planned, taking place in 2010, 2012, 2014, 2016, 2018, and 2020. For this analysis, we use only the 2010 wave. The project is organized by a team of economists and sociologists at Peking University and collects a rich set of data on family structure, income, expectations, and several other social and economic indicators. Detailed information about the sampling structure and overall plan for CFPS is available in Lv and Xie (2012).

[11]Recent work by Almond et al. (Forthcoming) uses Chinese gazetteers to identify when land-reform policy was implemented in different counties across the country.

implementation sometimes occurred in a different year in the rural regions of a given county or prefecture than it did in that locality's urban areas. We match these gazetteer accounts to the census data on residence status to generate different implementation years for rural and urban residents in a given prefecture. In most cases, the policy is implemented contemporaneously or nearly contemporaneously across urban (and, separately, rural) parts of different counties within a prefecture[12]. In around 20 cases, the policy is rolled out so gradually across counties within a prefecture that it is implausible to assign a single policy year appropriate for the entire prefecture. We exclude these cases from our analysis. We also exclude prefectures that changed to a system of five years of primary school plus four years of middle school and prefectures that have no record of policy implementation in the currently available educational gazetteers. Of those 65 prefectures in the census we exclude, 45 either implemented the policy too gradually or instead changed to the 5+4 system. The remaining 20 had no record of implementing the policy in the currently available educational gazetteers.

In the CFPS dataset, the location of observations is anonymized to the provincial level, which prevents use of the gazetteers to determine treatment status. Instead, we use the detailed data it collects on how many years individuals spend in each level of schooling to identify the year in which the policy was implemented within each (anonymized) county. Among individuals living in a given county, we apply a mean-shift algorithm (Fukunaga and Hostetler, 1975) which generates the most likely cohort in which the number of years spent in primary school jumps from five to six[13]. Its implementation in our context is straightforward; for observations in a given county, we regress individual-level years of primary school on a constant and an indicator function for having graduated in or after a given year as given in Equation 1:

$$s_i = \gamma_0 + \gamma_1 \cdot \mathbf{1}\{t_i \geq t^*\} + \varepsilon_i \tag{1}$$

We chose a functional form with only a break and no pre- or post-trends based on both the de-

[12] There are on average nine counties per prefecture. In some cases, one or two counties implement the policy one to three years before or after the majority of other counties. In these cases, we code the prefecture's policy implementation year as that of the majority of its counties.

[13] This mean-shift approach is similar to that used Munshi and Rosenzweig (2013).

scriptions of the policy implementation we read in the gazetteers and the patterns we saw in the county-specific histograms plotting years of primary school by cohort (such as that depicted in Figure A.4). We estimate 27 regressions for each county, corresponding to every possible treatment year in our data, $t^* \in [1981, 2007]$. In this equation, s_i is the number of years individual i spent in primary school, t_i is the year in which she graduated from primary school, and ε_i is an i.i.d. error term. The year (t^*) with smallest sum of squared residuals (ssr) is the predicted treatment year for that county[14]. This exercise generates a treatment year for each county in our estimation sample[15]. In Appendix 3, we use national statistics and the application of both archival and mean shift methods to a third observational data set, the China Labor-Force Dynamics Survey[16], to corroborate the reliability of the mean shift method's identified treatment years.

3.3 Empirical strategy

Our identification strategy is a stacked regression discontinuity (RD) design with a discrete running/forcing variable (Lee and Card, 2008). We define our running variable as the locality-specific distance to policy implementation in years and months, that is, the number of years and months[17] between an observation's birth year and month and the birth year and month of the first affected children in the locality. This choice of running variable is identical to that of Clark and Royer (2013) and Erten and Keskin (2018), among several others, and the method of stacking cutoffs similar to Pop-Eleches and Urquiola (2013) and Abdulkadiroğlu et al. (2014). This strategy compares outcomes of individuals finishing primary school just before the policy is implemented in a given locality (county or prefecture) to those in the same locality finishing primary school just after im-

[14]An example of this process is shown in Figure A.4.

[15]Beginning with 162 counties in the CFPS, we exclude the 18 counties from Shanghai, as they implemented the policy by extending middle school instead of primary school. Of the remaining 144, we include only those 112 counties in which we can detect a clear policy change - in the gazetteers and histograms, we see that some counties in other provinces, especially Shandong, also use the 5+4 model. Appendix 3 lists the inclusion criteria used to determine this sample. Our empirical results are robust to using data from all 144 non-Shanghai counties.

[16]The China Labor-Force Dynamics Survey (CLDS) is a panel survey similar to CFPS. We use it to corroborate the reliability of the Mean Shift method we use on CFPS. We do not use the CLDS for our main regressions as it has neither the large sample size of the census nor the fine-grain locality information of the CFPS (CLDS indicates only which prefecture individuals are in, not which county).

[17]Year of primary school entry + (month of primary entry -1)/12, using the annual school entry cutoff month, September, to differentiate cohorts.

plementation. The gradual rollout of the policy across time and space allows us to make this comparison while controlling flexibly for locality and cohort-by-province fixed effects.

Following Imbens and Lemieux (2008) and Lee and Lemieux (2010), our main estimating equation is an ordinary least squares regression of Y_{iclp}, the outcome of interest for individual i in birth cohort c and locality[18] l in province p, on a short set of key regressors:

$$Y_{iclp} = \beta_0 + \beta_1 * Treated_{cl} + \beta_2(t_{cl}|t_{cl} \geq 0) + \beta_3(t_{cl}|t_{cl} < 0) + \beta_4 V_i + \xi_m + \mu_l + \eta_{cp} + \varepsilon_{iclp} \qquad (2)$$

Here $Treated_{cl}$ is an indicator variable equal to 1 if the individual belongs to a cohort finishing primary school in or after the first affected cohort in her locality. t_{cl} is the locality-specific distance-to-treatment (in month-by-year units) for cohort c. We estimate the coefficient on distance-to-treatment separately for treated and untreated groups to account for pre- or post-policy trends, e.g., the possibility that the effect may differ as time since policy implementation increases and counties get better at implementing the policy. This ensures that β_1 captures only the difference between pre- and post-policy means[19] (Gelman and Imbens, 2014). V_i is a vector of predetermined characteristics which includes, at the individual level, gender, ethnicity, residence permit status, and urban/rural residence, which can vary within a county or prefecture. Birth month (ξ_m), locality (μ_l) and cohort-by-province (η_{cp}) fixed effects are also included in all specifications.

All regression results we present use robust standard errors clustered at the level of the running variable bin (month-by-year distance to policy implementation), as in Lee and Card (2008). We restrict our potential sample to cohorts leaving primary school between 1976 and several years before the sample is drawn (1995 in the census data and 2003 in the CFPS data) to give most individuals sufficient time to finish their schooling career before being observed[20]. We then conduct an optimal bandwidth calculation using the cross-validation procedure recommended by Imbens and Lemieux (2008) and used in Clark and Royer (2013). This generates a suggested bandwidth

[18] As mentioned earlier, this is at the county-level in the CFPS data and prefecture in the census.

[19] In Table A.1 we give six examples from gazetteers of how the policy is implemented which speak to the need to control for the possibility of implementation varying over time.

[20] 1976 marked the end of the Cultural Revolution and the end of the chaos it brought to the educational system of China.

of six years and six months on either side of the treatment threshold, which is the sample we use in all of our main results. For our main results (Tables 1 and 3) we also report results from the a smaller bandwidth - three years and five months - which was the second-best optimal bandwidth generated by the cross-validation procedure.

Migration poses two potential problems for our analysis. One is that people are observed somewhere other than where they are born. The second is that the policy may affect the likelihood of migrants to be missing from our dataset entirely. There is little evidence of the first potential problem in our dataset, as there are relatively few individuals observed away from the place where their hukou is registered. We exclude these observations, as there are so few of them and their treatment status is more difficult to pin down (because we do not have data on precisely when they moved or how long they have been staying in their current locality). We include individuals who have ever migrated and who are observed in the county/prefecture that their hukou is assigned to (in other words, migrants within localities and migrants who have either temporarily or permanently returned).

We conduct two tests for the second possible migration-related problem - that the policy induced more (or different) people to migrate and, as a result, be missing from our sample. First, we use Equation 2 to estimate the effect of the policy on the propensity to have ever migrated among those in CFPS sample, and we find no evidence of an effect ($\beta_1 = 0.00986$, $\sigma = 0.00726$). Second, we test for the possibility that the policy affects the density of individuals we observe in a given bin, e.g., it induces migration, which would show up as attrition from our sample. We find no evidence of this either; these results are described below and shown in Table A.4.

Another common strategy to use in this type of setting is a difference-in-differences (DD) design. This design compares the difference across individuals before and after the policy change in one locality to the same difference (across individuals over time) in other localities where there was no policy change during the period in question. As discussed in Angrist and Krueger (1999) and Abadie (2005), this method requires strong assumptions. One of these is the parallel trends assumption, which stipulates that in the absence of the policy or intervention being studied, the average outcomes of the treated and untreated should follow parallel paths over time. Abadie

(2005) raises the concern that this assumption is unlikely to hold when there is an imbalance between the two groups in pre-treatment characteristics that are likely to be associated with the dynamics of the outcome variable. In our setting, the stated reasons for differences between localities in the year of policy implementation appear to fit this description and, thus, likely violate the parallel trends requirement. Specifically, the law recommends the policy be implemented earlier in localities with "better economic conditions." The likelihood that the economic conditions of the locality may affect educational decisions and, all the more so, labor market outcomes, make the identifying assumptions of the DD design difficult to defend. Ostensibly, it would be possible to use the strategy of Abadie (2005) to semi-parametrically control for these differences. That approach, however, restricts its attention to differences in observed characteristics. Given the vagueness of what "economic conditions" were used in determining timing of policy implementation, we think it is highly likely that important unobservable determinants of policy implementation (e.g., government strategic plans regarding investment in infrastructure and industry) would be correlated with the labor market outcomes we measure. Nonetheless, we present DD results in Appendix 5 for thoroughness, and the results we generate using a balanced panel largely mirror what we find using the RD.

3.4 Testing assumptions behind our research design

For causal interpretation of our results, we require that within our geographical unit of interest, there is continuity in the conditional expectation of the outcome variable across the assignment threshold[21] (Lee and Lemieux, 2010). This requirement contains two main conditions - one, that the policy did not cause a change in the predetermined characteristics of who appears in our sample (i.e., selection or attrition bias), and two, that no other policy occurred concurrently with the one we study.

First, we address the potential for policy-related selection and attrition biases. This can be summarized by two questions: one, did the policy induce affected individuals to migrate or oth-

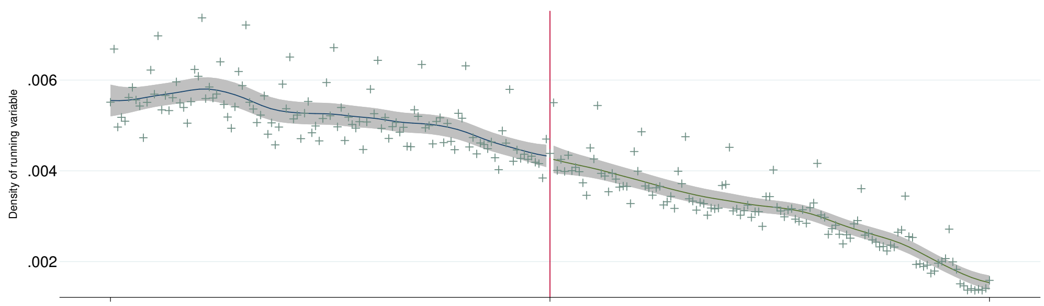
[21] As we are comparing within prefectures and counties, we do not need the timing of the policy to be randomly assigned across localities (Black et al., 2005; Meghir and Palme, 2005).

erwise attrit from our sample to a lesser or greater extent than the pre-policy group; two, if so, was this attrition concentrated among certain groups, for example, men or women? To address these questions, we test for continuity in the conditional expectation across the threshold of treatment assignment of three fundamental predetermined characteristics: the density of the running variable, gender composition of cohort, and proportion of individuals with a household registration certificate (hukou) from an urban area. As recommended by McCrary (2008), in Figure 2 we plot the density of the running variable, testing for bunching on one side of discontinuity and failing to reject the null of no bunching. Furthermore, as recommended by Lee and Card (2008), we use our main regression equation to estimate the "effect" of the treatment on the three predetermined variables for each dataset. These results are given in Table A.4; in all cases we fail to reject a zero effect.

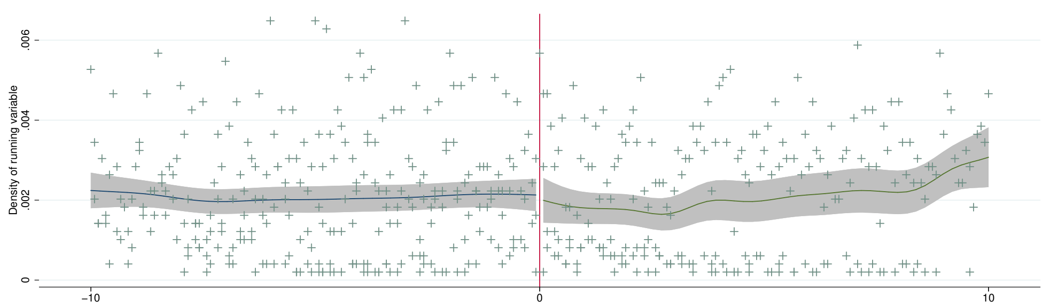
The second concern is that one or more other policies are implemented at the same time as the one we study. If this were the case, our treatment effect estimates would conflate their impacts. While many economic and policy changes occurred in the period we study, the staggered rollout of the policy over 25 years protects against this risk. For our estimates to be biased by another confounding policy, such a policy would have to be implemented in a way that was coordinated across counties and prefectures with the same time pattern of implementation as the primary education expansion.

We conduct a series of analyses to examine the possibility and scope for such bias, and conclude from the evidence summarized below that it is highly unlikely that there was concurrent, coordinated implementation of another policy which would violate our exclusion restriction. First, in our main estimating equation we include cohort-by-province fixed effects to flexibly control for potential province-specific time trends. The geography of the timing of implementation in each of China's prefectures according to archival records displays no such pattern - Figure A.2 provides a heat map of prefecture implementation years, with lighter shades indicating earlier implementation. Second, our coding of the gazetteers also uncovered no mention of a policy or external influence that was regularly coincident with implementation of the six year primary policy. Given the official nature of these documents, this strongly suggests absence of a consistent, officially-sanctioned

Figure 2: Density of running variable



Panel A: Census data



Panel B: CFPS data

This figure plots the density of the running variable (distance in time, measured by months and years, to treatment year/month) for the CFPS and census data. Corresponding regressions testing for continuity in the expectation of the treatment variable, proxied for by predetermined characteristics, are provided in Table A.4 and fail to reject the null of continuity.

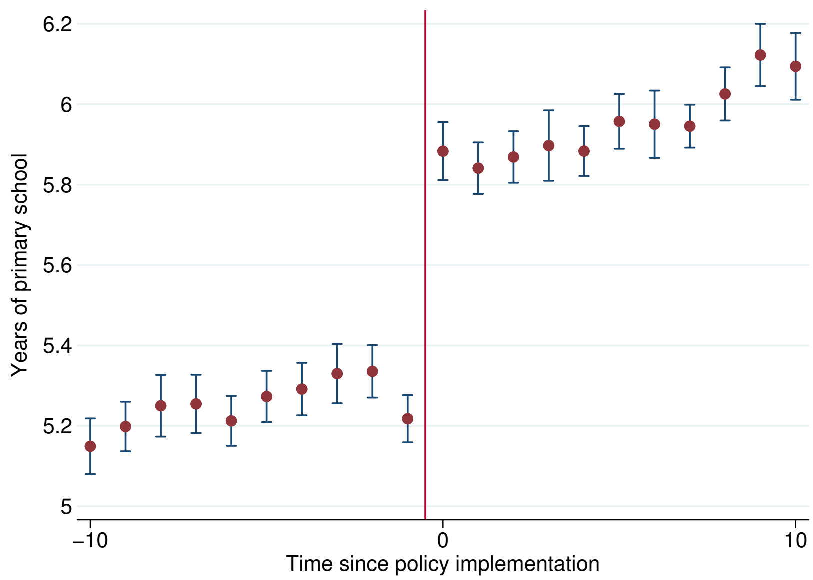
confounding policy.

Finally, two major education policy changes occurred during this period and deserve separate attention. First, a separate policy issued in April 1981 by the Ministry of Education mandated that the duration of high school to be extended from two years to three by the end of 1985. This implementation occurred over a much shorter time frame than the extension of primary school from five to six years: by 1984, 90% of students in high school were in three year programs; in contrast, it was not until 2003 that more than 90% of primary school students were in six year programs (National Institute of Education Sciences, 1984). Second, in 1986, the Chinese government made middle school compulsory. We show in Appendix 4 that this law appears to have little impact on the middle school attainment of observations in our estimation sample. We argue this is likely due to two main factors. One, there is widely documented porous enforcement of the law in rural areas (Fang et al., 2012). Two, in urban areas education levels are already high at the time of the policy announcement and so the law is likely to bind only for relatively few urban residents.

4 Empirical results - educational attainment

In this section, we estimate the impact of the policy on educational attainment. First, we show that the policy was indeed effective at extending the number of years individuals spent in primary school from five years to six. We then estimate the impact of this change on subsequent schooling outcomes, including years spent in post-primary schooling, whether or not an individual attains primary, middle, and high school credentials, and drop-out. We finish this section looking at the effect of the policy on subgroups of interest and on the characteristics of individuals by the highest educational credential earned. In this section and the next, we estimate results separately for urban and rural areas, following convention from previous work on schooling and the labor market in China and reflecting the fact that the labor market and other relevant parameters for schooling decisions differ systematically between the two (Liu and Zhang, 2013).

Figure 3: Event study of years of primary school before and after policy change



This plot uses CFPS data to show distance-to-treatment-year bin means of the number of years affected and unaffected individuals spent in primary school. The vertical line separates the affected (to the right of the line) and unaffected (to the left) cohorts, stacking the different policy implementation years at 0. We plot this in year bins instead of month-by-year bins for ease of viewing. Figure A.5 is an analogous figure using month-by-year bins instead, consistent with what we use in the regressions.

4.1 Primary schooling

We first examine whether the policy achieved its desired effect of increasing primary school for affected individuals. Figure 3 plots distance-to-treatment bin means and confidence intervals for the number of years spent in primary school among individuals in our CFPS sample. Prior to implementation of the policy, the mean is between 5.2 and 5.3, that is, between 20 and 30% of this sample spends more than the required five yeas in primary school. This comprises mainly individuals performing poorly in school who were made to repeat a grade[22]. At the policy implementation year the mean jumps to nearly 5.9, increasing to just over six years by the end of the ten year period included here. Results from the regression analog to this exercise are presented

[22]Data from the baseline wave of the China Education Panel Survey (CEPS), a new, nationally representative dataset collected by scholars at Renmin University of China, corroborate this claim. In the CEPS dataset, approximately 16% of surveyed individuals repeated at least one year in school.

Table 1: Effects of the policy on schooling outcomes

	CFPS		*Census*	
	(1)	(2)	(3)	(4)
Outcome	Urban	Rural	Urban	Rural
Years of primary schooling	0.498***	0.625***		
	(0.066)	(0.067)		
Years of post-primary schooling	0.057	0.168		
	(0.498)	(0.191)		
Highest credential: at least primary school	0.0451	-0.0087	-0.0022*	-0.0010
	(0.0480)	(0.0344)	(0.0012)	(0.0027)
Highest credential: at least middle school	0.0333	0.0347	0.0008	0.0008
	(0.0306)	(0.0397)	(0.0031)	(0.0048)
Highest credential: at least high school	0.0048	0.0041	0.0211***	0.0027
	(0.0712)	(0.0385)	(0.0075)	(0.0028)
Dropped out of school, any level			0.0012	-0.0020
			(0.0011)	(0.0021)
Number of observations	1,164	2,240	107,422	199,126

Each cell presents a treatment effect estimate from a separate regression (specification given in Equation 2) with the relevant robust standard error below, in parentheses. Standard errors are clustered at the county (CFPS) or prefecture (census) level. Columns 1 and 2 show results using CFPS data and 3 and 4 show results using census data.

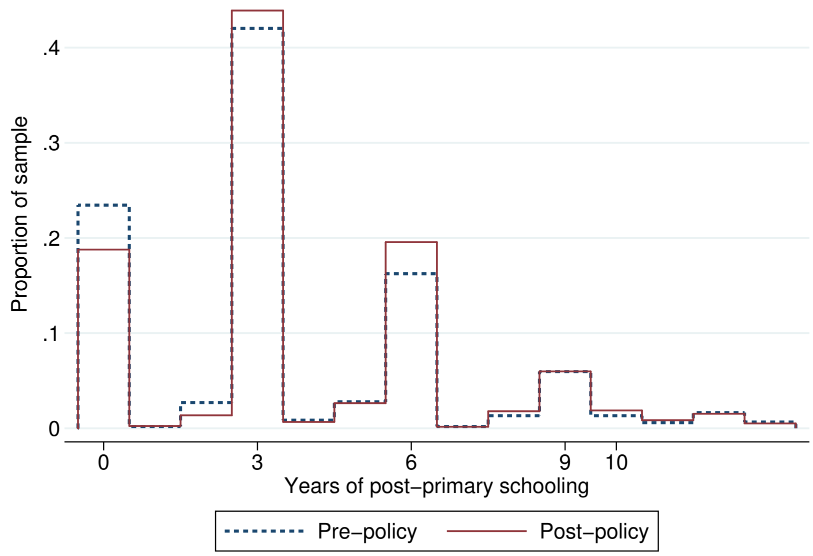
in the first row of Table 1. We estimate that the policy causes a 0.498 year increase in the number of years spent in primary school for urban residents who finish primary school within our optimal bandwidth sample restriction. For rural residents, the increase is 0.625. The smaller bandwidth sample generates nearly identical results.

There are several reasons why we do not estimate an immediate jump to six years of primary school. First, according to data from gazetteers and interviews, the policy was often rolled out in a way that split one or two cohorts of students in half to ease the transition, e.g., in the first year, sending half of fifth graders on to middle school, while retaining the other half in primary school for an additional year. The second reason is that we retain in our sample some counties where the fidelity of implementation appears to be less than perfect. If we include only those counties where the mean shift test precisely identifies a break, this estimate comes close to 0.8. Finally, as we are using recall data asking about an event that occurred between 5 and 25 years prior to the time of data collection, we anticipate some measurement error in estimates of when the shift happened.

To be conservative, we present reduced form estimates of the effect of the policy on subsequent outcomes as our central results, and give the instrumental variable (IV) coefficient estimates in parenthetical notes for the main labor market comparisons of interest. Our motivation for this is as follows: a sixth year of primary schooling post-policy is a deliberate expansion of the primary curriculum, as opposed to a forced repetition of the fifth year of primary school. Even for those who would have spent six years of primary school under the old system, the nature of the sixth year of primary school changes dramatically with the implementation of the policy. Finally, taking the first stage from the CFPS data and applying it to the Census, which includes data from a more representative sampling of counties as well as from five provinces not included in the CFPS, would involve a greater deal of uncertainty than we prefer.

To generate these parenthetical IV calculations, we divide the urban reduced form estimates by an adjusted version of the "first stage" given in the first row of Table 1. We show in Table 2 that the first stage estimates do not differ meaningfully between men and women, and so when estimating IV coefficients we use the same first stage estimate for both. We assume the vast majority of children reporting six years of primary school after our imputed policy implementation date are under the new system and the vast majority of those reporting six years of primary school before this date are under the old system. This suggests that in order to generate our first stage estimate, we should add the pre-policy rate of completing six years of schooling, which is 0.278 in urban areas (0.272 in rural areas), to our estimated coefficient of the impact of the policy on completed years of primary school. This yields a first stage of 0.776 in urban areas and 0.897 in rural areas or, equivalently, guidance to multiply the urban reduced form estimates by 1.29 and the rural estimates by 1.11 to generate the relevant IV estimates.

Figure 4: Distribution of post-primary schooling before and after policy change



This figure shows the probability mass functions of post primary schooling for observations in the CFPS data restricted to our optimal bandwidth/estimation sample. Note that the bunches at 3, 6, 9, and 10 correspond to middle school, high school, technical college, and university credential attainment years, respectively.

4.2 Post-primary schooling

This policy was implemented in each locality at a time when over 75% of students went on to get at least some post-primary schooling. The majority of affected individuals could potentially hold total years of schooling constant by offsetting the additional year of primary school with one less year of post-primary school, as depicted for the model student in Figure 1. In this subsection, we study this choice and, more broadly, how the policy affected completion of post-primary schooling.

Figure 4 shows, separately for untreated and treated observations in our optimal bandwidth sample, the distribution of post-primary schooling. This figure summarizes our main empirical results related to the effect of the policy on post-primary schooling. We see extensive bunching at credential attainment years and a small but visible difference between the treated and untreated groups in the location of this bunching.

The regression results for our schooling outcomes are given in the rest of Table 1, providing a formal test for the differences we observe in Figure 4. The second row shows our estimate of the effect of the policy on post-primary years of schooling to be 0.057 in urban areas (0.168 rural), which we interpret as evidence of a zero effect of the policy on years of completed post-primary schooling. This implies that the vast majority of Chinese citizens induced by the policy to attend an extra year of primary school chose not to offset this with less post-primary schooling.

We next use the census data to examine the effect of the policy on credential attainment. The census has coarser data on educational achievement (only highest credential attained, not years spent in each level of schooling) but is two orders of magnitude larger than the CFPS data. In the third, fourth, and fifth rows of Table 1, we estimate the effect of the policy on whether or not an individual earns at least a primary, middle, or high school credential, using both census and CFPS data. The effect of the policy on primary school completion is very close to zero and largely insignificant, with marginally significant evidence that it may have caused a slight decrease in primary school completion in urban areas. Our estimates show no evidence of a change in completion of middle school, and for high school, the census data suggests that there was a slight increase in completion in urban areas (but not rural) as a result of the policy. Using the census data, we find no effect of the policy on the pooled probability of dropping out of any level of schooling, shown in the final row of Table 1. Here again, the estimates we generate using the sample restricted to the alternative optimal bandwidth of three years and five months approximate those presented in Table 1.

4.3 Effects by gender and testing for changes in composition

Next we estimate the effects of the policy by gender of affected individuals and test for a change in the composition of the types of individuals who earn each credential. Here, we test for the possibility that the small estimates in Table 1 could mask two countervailing phenomena: first, some individuals advancing further than they would by virtue of the skills gained in the extra year, and second, others reducing post-primary schooling by an entire credential. We perform two exercises to investigate this possibility. First, we present gender-specific estimates of the results in Table 1; second, we explicitly test for changes in composition of background characteristics at each level of schooling.

Table 2 shows treatment effect estimates for the same outcomes examined in Table 1 and in the same format; the only departure is that in each set of results we added an interaction term - treated x female - to Equation 2 to capture the gender-specific effect of the policy on women and report this below the coefficient estimate on the treated variable.

We find that the policy has a slight negative effect on the completion of primary school for boys in both urban and rural areas, from high completion levels: the baseline completion rates of primary school for boys are 0.99 and 0.94 in urban and rural areas, respectively. For middle school, they are 0.96 and 0.66. For urban girls, we see small positive impacts on completion of all credentials, again from a high baseline (0.99, 0.94, and 0.55 for primary, middle, and high school completion, respectively). For girls in rural areas, we see gains in primary and middle school (baseline rates 0.91 and 0.55), but not in high school. We see a similar pattern for the impact on dropout, with a 0.38 percentage point decrease in the dropout rate for females in rural areas but no other detectable effects.

While the CFPS and Census estimates largely align, in three cases - urban female middle school graduation, rural male middle school graduation, and urban female high school graduation - the coefficients generated using the CFPS differ in sign from those generated using the Census. In all of these cases, the Census point estimate is statistically significant at the one percent level and the CFPS estimate is insignificant. We reconcile these results by observing that the point estimate generated using the Census data falls well within the confidence interval around the CFPS point estimate in each instance; we place more weight on the census estimate because of the data set's higher level of representativeness and the greater precision of the estimate. In a fourth case - urban male primary school graduation - there is a similar difference, but the Census estimate is only significant at the 10% level and, again, falls well within the confidence interval around the CFPS estimate.

Next, we estimate a version of our main empirical specification to test for compositional changes. We replace the single treatment variable with four dummy variables for the treatment interacted

Table 2: Heterogeneity, by gender, in effect of policy on education outcomes

	CFPS		*Census*	
	(1)	(2)	(3)	(4)
Outcome	Urban	Rural	Urban	Rural
Years of primary schooling				
Treated	0.5143***	0.6591***		
	(0.0695)	(0.0786)		
Treated x female	-0.0332	-0.0755		
	(0.0654)	(0.0550)		
Years of post-primary schooling				
Treated	0.0135	0.1352		
	(0.5458)	(0.2146)		
Treated x female	0.0912	0.0729		
	(0.4196)	(0.1927)		
Graduated from primary school				
Treated	0.0383	-0.0126	-0.0024*	-0.0079***
	(0.0540)	(0.0376)	(0.0013)	(0.0028)
Treated x female	0.0143	0.0086	0.0004	0.0133***
	(0.0452)	(0.0339)	(0.0009)	(0.0045)
Graduated from middle school				
Treated	0.0533	0.0238	-0.0040	-0.0230***
	(0.0334)	(0.0450)	(0.0034)	(0.0057)
Treated x female	-0.0418	0.0241	0.0091***	0.0460***
	(0.0301)	(0.0443)	(0.0024)	(0.0045)
Graduated from high school				
Treated	0.0240	0.0104	-0.0023	0.0037
	(0.0820)	(0.0432)	(0.0085)	(0.0032)
Treated x female	-0.0401	-0.0137	0.0445***	-0.0021
	(0.0688)	(0.0362)	(0.0060)	(0.0023)
Dropped out of school, any level				
Treated			0.0012	-0.0001
			(0.0011)	(0.0023)
Treated x female			0.0001	-0.0038**
			(0.0009)	(0.0017)
Number of observations	1,164	2,240	107,422	199,126

The reporting of results follows the organization in Table 1, only adding the coefficient estimate of the treatment variable interacted with an indicator variable for female gender for each of the analyses.

with an individual's highest educational credential (primary, middle, high, or tertiary). For outcome variables, we use a set of predetermined characteristics to proxy for household resources allocated to the child and scholastic ability, the most likely predictors of adjustment on the credential margin. We use the child's number of siblings to proxy for resources allocated to the child, and her/his parents' highest credential (mother's and father's separately) to proxy for scholastic ability. Wald tests of the equality of the treatment-by-credential level coefficients tell us whether the proportion of individuals holding a given credential with the predetermined characteristic of interest, relative to that proportion among those holding other credentials, changes across the treatment threshold. We use the CFPS and CLDS data for these tests (the census does not contain these data points), and fail to reject equality of the treatment-by-credential coefficients on any of the comparisons: number of siblings, mother's highest educational credential, and father's highest credential. We conclude from these analyses that the characteristics of who earns which credential are unlikely to have changed substantially as a result of the policy.

5 Empirical results - the labor market

In this section, we estimate the effects of the six year primary education policy on affected individuals' labor market outcomes. We then perform a series of robustness checks for these estimates, investigate potential mechanisms for our estimated effect on income, and conclude with a cost-benefit analysis of the policy over the period 1981-2050.

In the main results of this section, we restrict our attention to urban residents. This restriction follows the vast majority of empirical work on the returns to schooling in China, for example, Li (2003), Li et al. (2012b), and most of the studies included in Liu and Zhang's 2013 meta-analysis. This body of work motivates the focus on urban results with several claims, including 1) in rural areas treatment effect estimates would be muddied by the difficulty of measuring productivity in agriculture using existing household survey data, 2) in rural areas, schooling is likely to also affect labor allocation decisions on the intensive margin, i.e., how much time to spend working in agriculture vs. in the non-agricultural labor market, and 3) in rural areas, there is greater concern about selective loss to migration. For the sake of completeness, we present our results for rural residents in Table A.5.

Since the intervention had small but detectable effects on the proportion of people who earn middle school and high school credentials, we present two versions of our main results. The first uses equation 2 as written; the second adds fixed effects for the highest credential earned so that β_1 captures the change in income, conditional on highest credential, that accrues as a result of the policy.

5.1 Main labor market results

We use Equation 2 to estimate the effect of the policy on three labor market outcomes: one, employment status; two, if working, whether the individual is employed by the government; and three, monthly income. Though China was strictly a command economy as recently as 1978, reforms enacted in the 1980s and 1990s pushed the Chinese labor market to more closely resemble that of a market economy as early as the late 1990s (Cai et al., 2008). By the time we observe individuals in 2005, we expect most workers to earn wages that are at least strongly correlated with their relative productivity (Zhang et al., 2005). We use the 2005 mini-census data for all of the analyses in this section due to its large sample size. Our main dependent variable of interest is the natural logarithm of monthly income[23]. We also investigate the effects of the policy on whether the individual is employed and, if employed, whether she works for the government.

We present regression results for the urban sample in Table 3. We find no evidence that the policy had any effect on whether or not an individual is working, with a treatment effect very close to zero (0.4 percent, from a treated group mean of 77.6 percent; estimate shown in column 1) and a confidence interval which rejects anything larger in magnitude than a 1.6 percentage point increase or a 0.8 percentage point decrease in this probability. In column 2, we present our estimate of the effect of the policy on whether the individual works for the government (as opposed to for the private sector or in a state-owned enterprise). We observe a 1.9 percentage

[23]When estimating the effect of the policy on income, we drop those 340 observations from the optimal bandwidth estimation sample (142 in the treatment group, 198 in the control; out of 106,549 observations) who are working but report zero monthly income.

Table 3: Effects of the policy on labor market outcomes

	(1) Currently employed	(2) Works for government	(3) Log monthly income
Treated (β_1)	0.004	0.019***	0.026**
(no credential fixed effects)	(0.006)	(0.007)	(0.011)
Treated (β_1)	-0.000	0.014***	0.020*
(with credential fixed effects)	(0.006)	(0.006)	(0.010)
Number of observations	106,549	82,607	82,267

Data: census. All samples include only urban residents and non-migrants. Robust standard errors are given below the coefficient estimate in parentheses and are clustered at prefecture level. Estimates are generated using Equation 2. Columns 2 and 3 use only those observations in our sample who are currently employed.

point increase, from a baseline of 28 percent. This decreases to 1.4 percentage points when we control for credential fixed effects.

Next, we estimate the effect of the policy on the natural log of monthly income. These data come from the response to the census question: "in the last month (or calculating monthly income from last year's annual income), what was your overall income?" We estimate an increase of 2.63% in monthly income, statistically significant at 95% confidence[24]. We plot the event study graph with a locally estimated polynomial smoother (estimated separately for the treated and untreated groups) in Figure A.6. This shows a similar pattern - a discrete jump at the threshold of approximately 2.5 log points. When we control for credential fixed effects, the estimate drops to 2.0%, suggesting that about one fifth of the gain in wages came from the change in credentials, with the other four fifths coming from the returns to an additional year of schooling, controlling for highest credential. The IV estimate for these results is a 3.39 percent increase in income, or a 2.58 percent gain after controlling for credential fixed effects.

We next explore heterogeneity in treatment by subgroups, shown in Panels A-C of Table 4 and Figure 5. The coefficients in a given panel are estimated jointly, using an estimating equation similar to Equation 2 but replacing the single treatment variable with interactions between the treat-

[24] For rural residents, our estimate is statistically insignificant and the point estimate less than 0.1%.

ment variable and a dummy for membership in the mutually exclusive and exhaustive subgroups of interest (e.g., highest credential held; men and women; government and non-government workers), adding the relevant, un-interacted group dummies, and excluding the un-interacted treatment variable from the equation. For the estimates in Panels B and C, we also include credential fixed effects for reasons described above. In each panel, we report all relevant group x treatment coefficients.

First we investigate effect heterogeneity by the highest credential an individual holds. Given previous work on the distributional effects of extra instructional time (Meghir and Palme, 2005; Dobbie and Fryer, 2013), we anticipate larger gains for those with lower credentials. As predicted, the coefficient estimates reported in Panel A of Table 4 show that the income gains from the policy are monotonically decreasing in highest educational credential. For exposition, we plot the relevant coefficient estimates and their confidence intervals in Figure 5. This pattern is consistent with the goals of the policy, and imply that the policy is progressive: all who are affected give up a year of earnings, and those with the least education reap the largest gains in later life income.

Panel B of Table 4 shows that the policy slightly increases the proportion of women who are working and increases the proportion who work for the government. We estimate a four percent income gain for women in column 3. While government work is better paid than private sector work in our data, the income gain for women is not driven by a change in employment type. After controlling for type of work (government or not), the income gain women reap from the policy is still 3.8 percent.

Panel C shows that the income of private sector workers affected by the policy is five percent greater than those who are unaffected. Employees of state-owned enterprises gain little from the policy, however, and employees of the government appear to pay an income penalty for having spent an extra year in primary school as a result of the policy. This difference is unsurprising, as pay is almost certainly more closely linked to the relative productivity of labor in the private sector than in the government (Li et al., 2012a). Independently run Wald tests reject equality of the private sector coefficient with that of either other group.

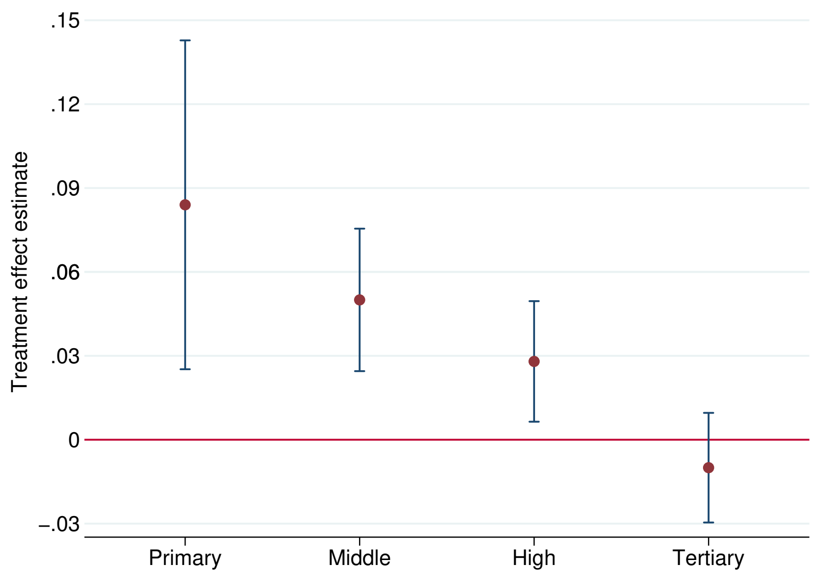
Behind both the positive effects for women's income and the null overall effects for men's in-

Table 4: Subgroup effects of the policy on labor market outcomes

	(1) Currently employed	(2) Works for government	(3) Log monthly income
Panel A: by highest credential			
Primary school	0.018	0.021*	0.085***
	(0.016)	(0.013)	(0.030)
Middle schoool	-0.001	0.012**	0.051***
	(0.007)	(0.006)	(0.013)
High school	0.009	0.030***	0.028***
	(0.007)	(0.007)	(0.011)
Post-secondary	-0.010	0.000	-0.014
	(0.007)	(0.008)	(0.010)
R-squared	0.124	0.275	0.394
Number of observations	106,114	82,488	82,149
Panel B: by gender			
Female	0.012*	0.022***	0.040***
	(0.007)	(0.007)	(0.011)
Male	-0.013**	0.006	0.002
	(0.006)	(0.006)	(0.011)
R-squared	0.124	0.275	0.394
Number of observations	106,114	82,488	82,149
Panel C: by employer			
Private sector	-	-	0.049***
			(0.012)
State-owned enterprise	-	-	0.010
			(0.012)
Government	-	-	-0.021**
			(0.011)
R-squared	-	-	0.399
Number of observations			82,149

Data: census. In each panel, we present coefficients from a single regression generated using a modified version of Equation 2 to generate estimates of heterogeneity across groups as defined in the panel title. Presented coefficients are for dummy variables for membership in the group given in left column (e.g., those whose highest credential is primary school) interacted with the treatment dummy. Robust standard errors are given below the coefficient estimate in parentheses, clustered at prefecture level. Panel A coefficients, when weighted by proportion of the sample corresponding to the proportion of each group among urban residents in the census, sum to 0.0209.

Figure 5: Effect of the policy on log monthly income, by highest credential



Note: This table plots the treatment effect estimates and corresponding 95% confidence intervals for the effect of the policy on log monthly income, conditional on highest credential. Treatment effect estimates and standard errors are presented in Panel A of Table 4, along with relevant details on the regression specification. We cannot reject equality of the estimated effect of the policy on years of schooling, propensity to drop out of school, and receipt of credential for the four different credential groups (primary, middle, high, tertiary).

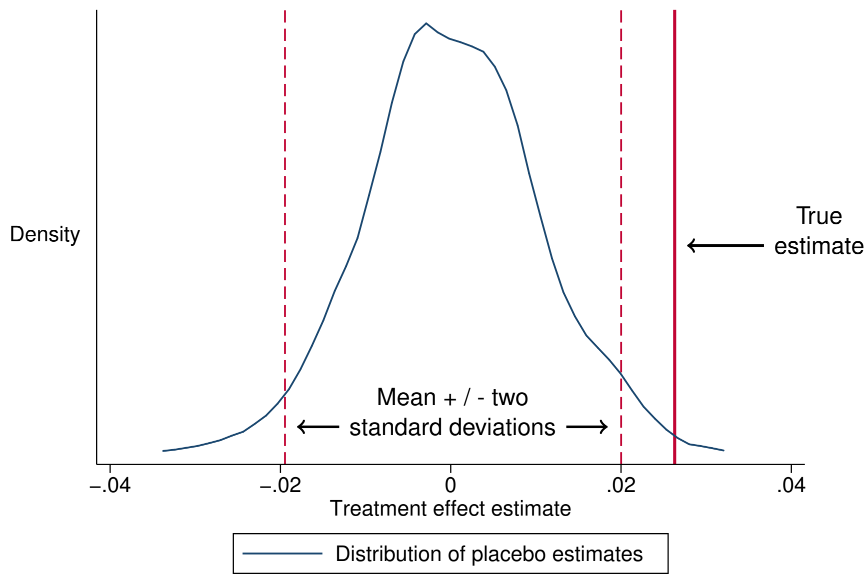
come is the same variation by subgroup we see in Panel A - returns to the policy estimated separately for each gender monotonically decrease by highest credential, with large returns for primary and middle school credential holders and smaller or negative returns for those with high school or tertiary credentials. We also see larger gains among those working in the private sector for both men and women. These two sets of analyses, however, generate wide confidence intervals which include zero, speaking to the limitations of this research design. Comparing the treated and untreated, within subgroups of subgroups in each locality, limited to a narrow bandwidth around the treatment year, leaves us with too few observations per locality to generate precise estimates using the RD design as specified.

5.2 Robustness

In this subsection, we conduct a few statistical exercises to examine the robustness of our estimates and refine their interpretation. We first conduct a permutation test to examine whether our research design would mechanically generate a difference between the treated and control groups unrelated to the effect of the policy. In this test, we generated 1,000 draws of randomly selected years for each prefecture (sampled from the full support of the estimation sample's potential years, 1981-1997). Then, using the treatment status assigned by these placebo years, we estimate the placebo treatment effect on wages for each draw. Figure 6 gives the probability density function for these estimates. The placebo treatment effect estimates are normally distributed, with a mean of 0.00027 and a standard error of 0.0099, putting the true estimate of 0.0263 well beyond two standard deviations from the mean. We conclude that the sign and significance of our estimates are not merely a mechanical result of our research design.

We next test for the possibility of heterogeneity across time in the treatment effect. This captures two possibilities. The first is general equilibrium effects exerting a downward bias on our results; as the proportion of laborers in a given market treated with the extra year of primary school increases, we might expect to see the labor market returns to the extra year decrease. The second is time-specific effects of labor market entry. As China modernized over this period, the composition of the labor market changed substantially. Those entering the labor market later in life

Figure 6: Permutation test results



Data source: census. This figure plots the distribution of effect estimates of placebo treatment on log monthly income from 1,000 draws of placebo years, using Equation 2. We also show a confidence interval of two standard deviations from the mean. For reference, the overall effect from Table 3 is shown as a thick red vertical line.

may be more likely to enter jobs where the wage is closer to the marginal product of labor. Both of these effects push in the same direction, with larger coefficients predicted for those in prefectures which implement the policy at a later date.

To run this test, we divide our sample into three groups based on when the policy was implemented: 1981-85, 1986-1990, and 1991-1995. Recall that the census was collected in 2005 and so, as described in Section 3, with the census data we only study the first 15 years of implementation of the policy. This sample restriction ensures that those leaving primary school in the early 1990's will have entered the labor market by the time we observe them. Though there is a monotonic relationship between lateness of implementation and the size of the effect estimate (β_1 = 0.016, 0.024, and 0.086 for early, middle, and late implementers), consistent with the general equilibrium effect being smaller than the partial equilibrium effect, a Wald test fails to reject the equality of these three coefficients (p-value of the f-test: 0.2025) or that of the late implementer with either of the two others individually. Note, however, that only 1.2% of our sample of more than 82,000 observations is in the late implementer group, so the large estimate for this group may simply be an artifact of measurement error.

As in all exercises attempting to estimate the labor market returns to a year of schooling, we face the problem that an individual with one more year of schooling has one less year of experience in the labor market, and so our estimate may capture the returns to a year of schooling minus the returns to a year of experience. We calculate upper and lower bounds on this potential contribution, as in Manski (2013). Our lower bound is 0 if we assume that the labor market returns to a year of age are entirely due to maturity and not to work experience specifically. To calculate the upper bound, we estimate the return to a year of experience in the labor market at a given age, δ_a, and weight this by the age composition of our estimation sample ω_a, e.g., $\sum_{age=a} \omega_a \delta_a$. This generates an upper bound of 0.0168, meaning that the Manski bounds on our overall estimate for the effect of the policy on income are (0.0263,0.0431).

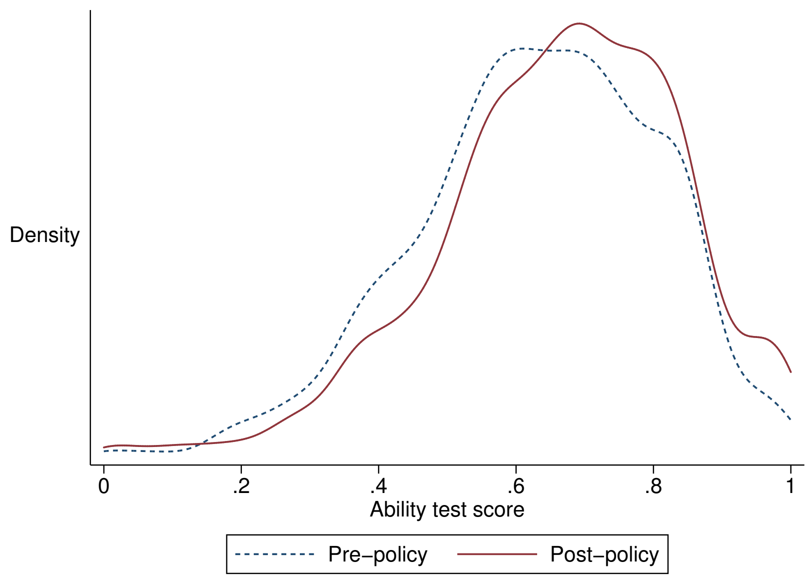
5.3 Mechanisms

In this section we discuss the potential mechanisms for the income results that we find. We estimate the effect of the policy on cognitive skills test scores and then discuss the potential impact of the policy on non-cognitive/socio-emotional skills. We focus on trying to explain the larger effects of the policy among those with lower credentials.

First, we test for a difference between the treated and untreated groups in their cognitive skills test scores administered to adult respondents in the CFPS survey. Figure 7 plots the kernel density functions for treated and untreated individuals using the optimal bandwidth sample. The two distributions track each other quite closely, but there is a visible rightward shift in the treated distribution. A Kolmogorov-Smirnov test rejects the equality of the two distributions with a p-value of 0.003. This comparison uses only the raw data, however, and risks conflating secular time trends with the effect of the policy. The estimated treatment effect of the program generated with our estimating equation and the optimal bandwidth sample is 0.010 points, or 0.055 standard deviations in terms of the untreated group's test scores. This estimate is statistically insignificant ($\sigma = 0.011$ points), however, and quantile regressions generate similar results, with positive coefficients but associated standard errors unable to reject a zero effect. This pattern - a significant positive effect on monthly income, but no statistically significant difference in cognitive ability - is also seen in Heckman (2006) and Chetty et al. (2011), who show that childhood interventions which may initially generate increases in test scores or cognitive skills often bring labor market returns decades after measurable cognitive gains wane.

While we cannot measure non-cognitive skill acquisition in our data, it is another potential source for the income effects that we measure (Kautz et al., 2014). In addition to the curricular knowledge students gain, students also practice social skills while in school. For example, being made to come to class on time and sit still for the duration may teach discipline, punctuality, and attentiveness. Interacting with teachers and administrators may teach the social skill of how to speak to superiors and figures of authority. Completing homework assignments and preparing for tests, in turn, may teach how to complete complex tasks with varying levels of instruction. Such skills carry substantial returns in most modern labor markets (Heckman and Kautz, 2012), and the

Figure 7: Kernel density of cognitive skills test score, by treatment status



This plot shows the kernel density of cognitive skills test scores for observations from the CFPS data, plotted separately for those in affected and unaffected cohorts, in the optimal bandwidth sample.

extra year of primary school is likely to have developed these further, particularly for those who finish their educational careers with only a primary or middle school credential.

Independent of the relative contributions of cognitive and non-cognitive/socio-emotional skill acquisition to the labor market gains we measure, the progressive nature of the policy remains. We argue that an important contributor to this pattern of results is that the marginal contribution of an extra year of schooling may convey the largest benefit to those with the fewest total years of schooling. We base this both on economic intuition and on the information we gleaned from the qualitative work we conducted. To understand affected individuals' perception of the policy's effects, benefits, and costs, we carried out structured interviews with a group of parents, teachers, and students who were affected by the policy. In response to a prompt about the potential benefits of the policy, our respondents were equivocal about the overall effect. Among those who thought the policy was beneficial, several volunteered that it brought the largest benefits to struggling students. Several others were unsure whether the policy helped anyone, but added that if it did, the beneficiaries were likely to be the least able.

Intuitively, the fact that this policy adds an extra year to primary school but does not add material to the curriculum supports these claims. Seen through the lens of the literature on human capital accumulation, this intervention targets the intensive margin of skill acquisition. The policy makes individuals spend more time mastering a fixed set of skills set in China's primary school curriculum, as opposed to a year spent learning new skills by advancing in school, e.g., from the 10th to 11th grade, as in most previous work on the returns to schooling (Card, 1999). As mentioned earlier, our results complement those of Meghir and Palme (2005) and Dobbie and Fryer (2013), who find that education policy which increases the amount of time spent mastering a given set of skills can be redistributive toward the disadvantaged.

5.4 Cost-benefit analysis

We next use our results to generate an estimate of the costs and benefits of the program. We borrow our framework directly from Duflo (2001), focusing on the direct private gains and losses and ignoring the other potential benefits of increased income (e.g. decreases in fertility and child morbidity). Though we are aware that this type of exercise involves a precarious amount of uncertainty (Manski, 2013), we feel it important to gain some insight into the net effect of such a large reallocation of resources.

As in Duflo's analysis, we choose our time frame to span from the first cohort in which some students leaving primary school are affected by the policy, 1981, to the end of 2050, and assume an annual discount rate of 5%. Our cost estimate has as its sole argument the lost year of wages[25], w_{it}, that affected students i forgo during the year t they spend in school instead of in the labor market. We multiply this cost by θ, the proportion of individuals induced to spend an extra year in school because of the policy. We add this scale factor to account for the fact that, in our optimal bandwidth sample, about 25% of students spend a sixth year in primary school under the five year system and only 5% spend a seventh year in primary school under the six year system. Given these figures, we set $\theta = 0.8$.

$$Cost = \theta \sum_{i} \frac{w_{it}}{(1+r)^{i-1}} \tag{3}$$

For each cohort, we determine what proportion of individuals leave school with a primary, middle, high school, or tertiary credential. We calculate the value of the year of wages the individual forgoes as the wages they would have earned with zero years of experience and the highest credential they ultimately obtain in the year they earn that credential. For example, an individual born in 1975 would start school in 1981. If her highest educational credential is a middle school degree and she was affected by the policy (and so she spent nine years in school), her "lost year" would be worth the average wages of those obtaining a middle school degree in 1990 in their first year of work. We then calculate the total value of the years lost for all students in each cohort from 1981 to the last cohort entering the labor force in 2050, using the same formula for the value of wages used in the benefit calculation below. Unlike Duflo, we do not incorporate a deadweight loss of taxation, as we see the policy as a transfer and assume there is no productive activity displaced by the policy other than the reallocation of affected individuals' time.

Our structure for estimating the benefit of this policy is also taken directly from Duflo's analysis.

[25]We assume that the extra year of primary school does not induce individuals to remain in the workforce for longer at the end of life. In Appendix 5 we discuss other potential costs and benefits and our decision not to include them in this calculation.

Specifically, we estimate the sum of income gains for all affected cohorts over the time frame we have chosen:

$$Benefit = \sum_{t}\sum_{c} \frac{\alpha GDP(t)S(c,t)P(c)\beta}{(1+r)^{t-1}} \tag{4}$$

Here α is the share of labor in GDP[26], $S(c,t)$ is the size of cohort c divided by the total working population in year t. P(c) is the proportion of cohort c affected by the policy[27], and β is our estimated effect of the policy on income. We sum the benefits earned by each cohort in the labor force in each year, assuming people work from when they leave school until age 65[28].

In Table 5, we present four cost-benefit estimates for the period 1981-2050, varying two key assumptions about the nature of β. In the left column, we present estimates using the average treatment effect estimate of β, while the figures in the right column use credential-specific estimates (i.e., different β for primary, middle, high school, and university degree holders, respectively, as seen in Panel A of Table 4) weighted by the proportion of individuals in a given cohort holding each credential. The right column thus accounts for the large, positive changes in educational attainment over this time, using credential-specific treatment effects and allowing the distribution of highest educational credential to vary by cohort. In the top row, we present estimates using our urban sample, and in the bottom row, our results are the weighted sum of estimates from rural and urban China generated in separate regressions, again weighted by the cohort-specific ratio of rural to urban residents. This accounts for the fact that returns are lower in rural China and, until 2011, more than half of China's population was rural. Estimates for β_{rural} are given in Table A.5.

The sign of the estimate depends on how we treat rural areas. Using the coefficient estimate from urban areas to generate the rural figures, we find the policy generates a net gain. If we use the rural estimates to generate the rural figures, we estimate a net loss. Given the concerns about

[26]This labor share data come from Karabarbounis and Neiman (2014). The rest of the data used in this section was downloaded from stats.gov.cn and the World Bank's World Development Indicators, projected forward using multi-year moving averages.

[27]This is a slight over-estimate of the benefits accruing to cohorts in the first few years of implementation, as we derive this proportion affected measure from national statistics and not local gazetteers. Given the patterns of implementation we observe in the gazetteers, however, gradual implementation applies to a minority of localities. Among those cases, the vast majority reach full implementation in one to three years.

[28]This is a simplifying assumption. For those working in factories, the official retirement age is 60 for men and 50-55 women, but individuals often work well beyond these ages. In addition, the official age is slated to be changed in the next five years. Using the official retirement age would reduce the amount of years during which benefits accrue and thus reduce our estimate of the benefits of the policy.

Table 5: Cost-benefit calculation: 1981-2050

	Assumption about heterogeneity in β_1 by highest credential held	
β_1 estimate used for rural areas	Using average treatment effect for all	Using credential-specific treatment effects
Using estimates from urban areas	33,498	62,825
Using rural effect estimates from our data	-15,181	-144,688
Costs	117,047	-
Cost-benefit calculation for cohort leaving primary school in 2015 (urban, per-credential β_1)	-155	

*Estimates in millions of 2015 US Dollars.

the rural estimates outlined in Section 5.1, we are inclined toward the two positive estimates in the top row. Note that we use our reduced form estimates of β for this analysis. We calculated an analog to this table using IV estimates. While it increased the magnitude of the estimates in each cell, the signs were unchanged. This stems from the fact that the sign of the estimate is driven largely by whether we use the negative estimated returns to those in rural areas with either a high school or tertiary credential.

To provide an estimate of the current per-year cost of the policy, we compare the value of a lost year of productive work to the lifetime productivity benefits of the extra year of primary schooling for the cohort leaving primary school in 2015. We estimate this to be a net loss of approximately 155 million dollars, or less than $15 per person in the cohort.

We offer two further observations. First, China is a victim of its own success in increasing educational attainment. The largest benefits accrue to those with the lowest credentials, but the proportion of individuals in a cohort with these lower credentials is steadily decreasing over time. We interpret this as evidence of the following claim: while the policy was a success based on its original aims (to increase skills, particularly among the least educated), it may have outlived its purpose given the dramatic increase in average educational attainment over time.

The second observation is that the per-person cost of the policy is very small and we are likely understating the benefits. Prior research suggests there are several possible positive social impacts of this policy change, most of which we either lack the data to analyze or are not powered to test for, that could well add to the benefits side of our calculation. These include a beneficial effect of the policy on crime, health, and mortality (Lochner and Moretti, 2004; Lleras-Muney, 2005). It is likely too early to measure the mortality effects, but we anticipate various other salutary effects of the policy. We leave these to future research.

6 Conclusion

We find that the Chinese government's expansion of compulsory primary education from five years to six changed the primary school experience for hundreds of millions of individuals, increased income of affected individuals by 2.6-3.4 percent on average, and had the largest income effects for a key disadvantaged group in China: those who leave schooling with only a primary or middle school credential. A rough estimate suggests that the policy has reallocated over 450 million years of productive time from the labor market to the pursuit of schooling to date and its benefits likely exceed its costs. Over time, however, these benefits dwindle with the dramatic increase in educational attainment in China and thus the decrease in the size of the group likely to benefit most.

More broadly, our analysis highlights two features of policy setting the duration of primary school. One, we highlight the importance of this policy decision in developing countries. Every government must decide upon the number of years children spend in primary school, and our analysis suggests this policy decision affects how many years the vast majority of individuals will ultimately spend in school as well as their subsequent earning power in adulthood. Two, we show that extending the duration of primary schooling has large, progressive effects on later life income, a potentially important lesson for the many developing countries with a sizable proportion of individuals who never go beyond primary school.

References

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

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak, "The elite illusion: Achievement effects at Boston and New York exam schools," *Econometrica*, 2014, *82* (1), 137–196.

Almond, Douglas, Hongbin Li, and Shuang Zhang, "Land reform and sex selection in China," *Journal of Political Economy*, Forthcoming.

Angrist, Joshua and Alan Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 1991, *106* (4), 979–1014.

Angrist, Joshua D and Alan B Krueger, "Empirical strategies in labor economics," *Handbook of labor economics*, 1999, *3*, 1277–1366.

Arteaga, Carolina, "The effect of human capital on earnings: Evidence from a reform at Colombia's top university," *Journal of Public Economics*, 2018, *157*, 212–225.

Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying education: Evidence from two randomized experiments in India," *Quarterly Journal of Economics*, 2007, *122* (3), 1235–1264.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust differences-in-differences estimates?," *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 2005, *95* (1), 437–449.

Cai, Fang, Albert Park, and Yaohui Zhao, "The Chinese labor market in the reform era," in Loren Brandt and Thomas G. Rawski, eds., *China's great economic transformation*, Cambridge University Press, 2008.

Card, David, “The causal effect of education on earnings,” *Handbook of Labor Economics*, 1999, *3*, 1801–1863.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star,” *Quarterly Journal of Economics*, 2011, *126* (4), 1593–1660.

Clark, Damon and Heather Royer, “The effect of education on adult mortality and health: Evidence from Britain,” *American Economic Review*, 2013, *103* (6), 2087–2120.

Clay, Karen, Jeff Lingwall, and Melvin Stephens Jr, “Laws, educational outcomes, and returns to schooling: Evidence from the full count 1940 census,” *National Bureau of Economic Research Working Paper 22855*, 2016.

Devereux, Paul J and Robert A Hart, “Forced to be rich? Returns to compulsory schooling in Britain,” *Economic Journal*, 2010, *120* (549), 1345–1364.

Dobbie, Will and Roland G Fryer, “Getting beneath the veil of effective schools: Evidence from New York City,” *American Economic Journal: Applied Economics*, 2013, *5* (4), 28–60.

Duflo, Esther, “Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment,” *American Economic Review*, 2001, *91* (4), 795–813.

Erten, Bilge and Pinar Keskin, “For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey,” *American Economic Journal: Applied Economics*, 2018, *10* (1), 64–105.

Fang, Hai, Karen N. Eggleston, John A. Rizzo, Scott Rozelle, and Richard J. Zeckhauser, “The returns to education in China: Evidence from the 1986 compulsory education law,” *NBER Working Paper*, 2012, (18189).

Fukunaga, Keinosuke and Larry Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, 1975, *21* (1), 32–40.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs," *NBER Working Paper 20405*, 2014.

Heckman, James J, "Skill formation and the economics of investing in disadvantaged children," *Science*, 2006, *312* (5782), 1900–1902.

_ **and Tim Kautz**, "Hard evidence on soft skills," *Labour economics*, 2012, *19* (4), 451–464.

Imbens, Guido W and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," *Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans, "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," *National Bureau of Economic Research Working Paper 20749*, 2014.

Lee, David S and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

_ **and Thomas Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

Li, Haizheng, "Economic transition and returns to education in China," *Economics of Education Review*, 2003, *22* (3), 317–328.

Li, Hongbin, Lei Li, Binzhen Wu, and Yanyan Xiong, "The end of cheap Chinese labor," *Journal of Economic Perspectives*, 2012, *26* (4), 57–74.

— , **Pak Wai Liu, and Junsen Zhang**, "Estimating returns to education using twins in urban China," *Journal of Development Economics*, 2012, *97* (2), 494–504.

Liu, Elaine and Shu Zhang, "A Meta-Analysis Of The Estimates Of Returns To Schooling In China," *Working Paper*, 2013.

Liu, Yingjie, *Book of Major Educational Events in China 1949-1990 (in Chinese)*, Zhejiang Education Publishing House, 1993.

Lleras-Muney, Adriana, "The relationship between education and adult mortality in the United States," *The Review of Economic Studies*, 2005, *72* (1), 189–221.

Lochner, Lance and Enrico Moretti, "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports," *American Economic Review*, 2004, *94* (1), 155–189.

Lucas, Adrienne M and Isaac M Mbiti, "Access, sorting, and achievement: The short-run effects of free primary education in Kenya," *American Economic Journal: Applied Economics*, 2012, *4* (4), 226–253.

Lv, Ping and Yu Xie, "Sampling Design of the Chinese Family Panel Studies," Working Paper, Institute of Social Sciences Surveys, Peking University May 2012.

Manski, Charles F, *Public policy in an uncertain world: Analysis and decisions*, Harvard University Press, 2013.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Meghir, Costas and Mårten Palme, "Educational reform, ability, and family background," *American Economic Review*, 2005, *95* (1), 414–424.

Mo, Di, Linxiu Zhang, Hongmei Yi, Renfu Luo, Scott Rozelle, and Carl Brinton, "School dropouts and conditional cash transfers: Evidence from a randomised controlled trial in rural China's junior high schools," *Journal of Development Studies*, 2013, *49* (2), 190–207.

Morin, Louis-Philippe, "Estimating the benefit of high school for university-bound students: Evidence of subject-specific human capital accumulation," *Canadian Journal of Economics*, 2013, *46* (2), 441–468.

Munshi, Kaivan and Mark Rosenzweig, "Networks, commitment, and competence: Caste in Indian local politics," *NBER Working Paper 19197*, 2013.

National Institute of Education Sciences, Beijing, *Chronicle of Education Events in China (in Chinese)*, Educational Science Publishing House, 1984.

Oreopoulos, Philip, "Estimating average and local average treatment effects of education when compulsory schooling laws really matter," *American Economic Review*, 2006, *96* (1), 152–175.

_ , "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling," *Journal of Public Economics*, 2007, *91* (11), 2213–2229.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a better school: Effects and behavioral responses," *American Economic Review*, 2013, *103* (4), 1289–1324.

Stephens, Melvin and Dou-Yan Yang, "Compulsory education and the benefits of schooling," *American Economic Review*, 2014, *104* (6), 1777–1792.

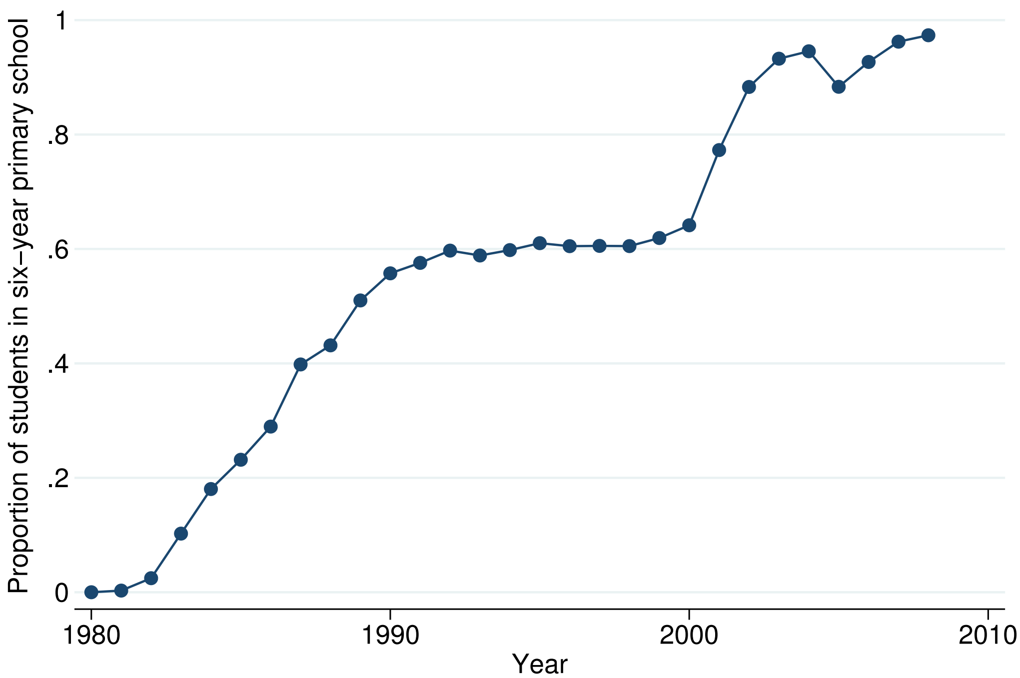
Vogel, Ezra F, *Deng Xiaoping and the transformation of China*, Belknap Press of Harvard University Press, 2011.

Zhang, Junsen, Yaohui Zhao, Albert Park, and Xiaoqing Song, "Economic returns to schooling in urban China, 1988 to 2001," *Journal of Comparative Economics*, 2005, *33* (4), 730–752.

Appendix (for online publication only)

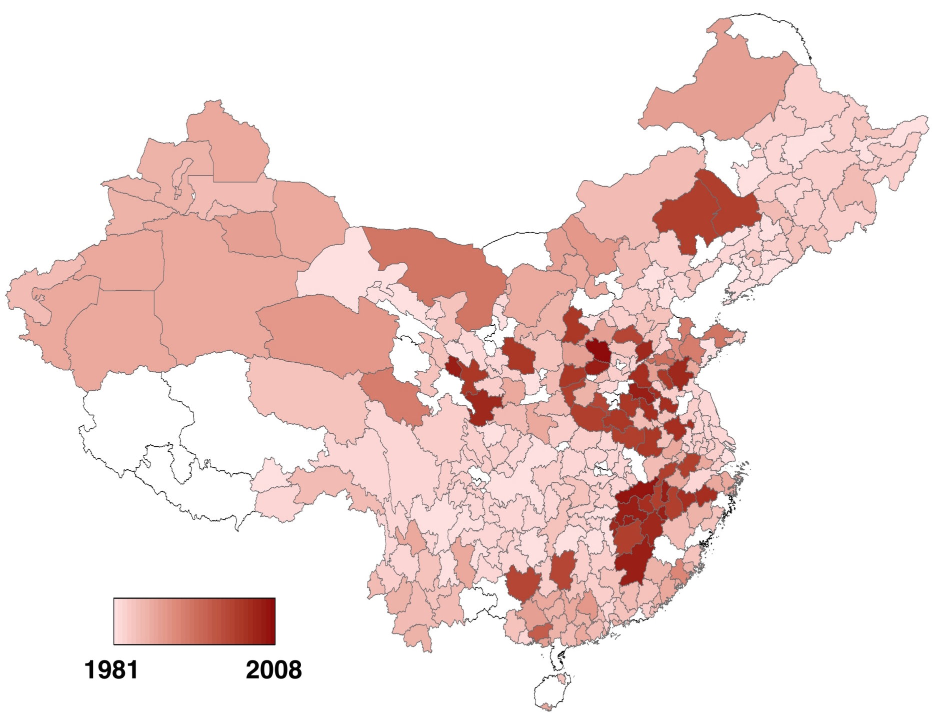
Appendix 1 - Appendix figures

Figure A.1: National data on proportion of students in six year primary education



Data source: Chinese Ministry of Educational Statistics

Figure A.2: Timing of implementation by prefecture in Mainland China



Prefectures are classified by our coding of gazetteer data. In some cases, prefecture boundaries have changed since the archive was published. In these cases, we take the prefecture level-average of the treatment year in all previous prefecture capitals falling within the new prefecture. In prefectures that are colored white, we are unable to assign a unique prefecture-level treatment year for reasons discussed in Section 3. Note that we exclude from this map Hong Kong, Macau, Taiwan, and various islands.

Figure A.3: Example of a page from a county gazetteer

续表

1979	440	85 846	3 721
1980	355	84 274	3 565
1981	362	83 952	3 604
1982	386	81 925	3 509
1983	353	78 492	3 328
1984	361	80 059	3 373

2．学制、课程设置及教法

1976年，全县小学实行五年制。1984年，按照省教育厅要求，将小学五年制改为六年制，并在两年内完成。按照教育部门的要求，境内小学采取农村以乡，城镇以校为单位自行命题进行考试、择优的办法，使五年级应届毕业生，三分之二毕业升入初中。1985年应届毕业生三分之一升入初中，留校三分之二升入六年。

1976年秋，学校所开课程有语文、数学、自然、地理、历史、音乐、美术、体育、劳动，使用全国统编教材。1982年增设思想品德课。1984年，仍延用原有的课程设置及教材。

1976年后，各小学校逐步恢复正常教学秩序。1979年，县革委会发出了55号文件，摘掉了小学附设的初中班。恢复了中心校管理体制。1981年，学校工作的立足点转为以教学为中心，注重提高教学质量，加强教学研究和改革，强调“双基”（ 基础知识、基本技能）教学。要求教师立足课堂，提高课堂教学效果，着眼点放在培养学生能力、开发学生智力上。

1984年，各校进行教学改革试验，坚持以学生为主体，以教师为主导，全面进行教学工作整体改革。按照教育局要求，各校成立了课外活动小组，开展了“第二课堂”活动。学校在加强基础知识教学同时，还开展了小学语文、数学知识竞赛。是年，全县小学毕业生双科合格率达90%。

第二节　教育体制改革后的小学教育

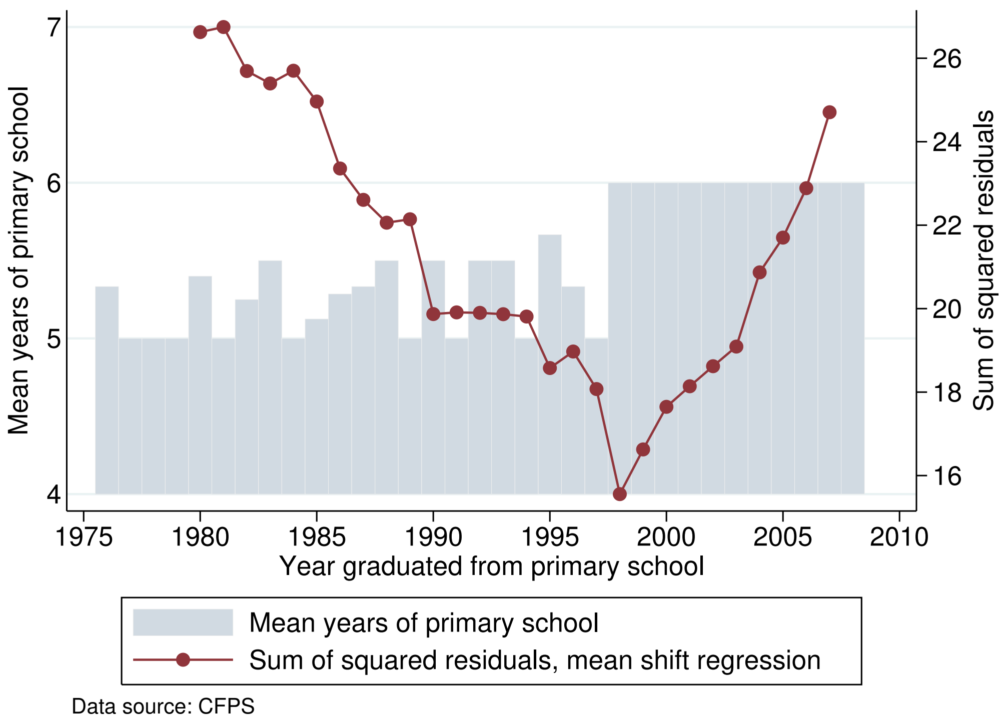
一、学校发展概况

1985年，全县有小学校335所，其中教育部门和集体办320所，其他部门办15所，设立下伸点65个（教育部门和集体办58个，其他部门办7个）。在校学生80 464人（女学生38 583人），其中教育部门和集体办76 135人，其他部门办4 329人。教学班2 431个，其中教育部门办和集体办2 294个（复式班105个），其他部门办137个（复式班2个）。二部制学校7所，有教学班42个。教职工3 653人，其中教育部门办1 848人，集体办1 605人，其他部门办200人。

1987年，林海、三合乡划归白城市（现洮北区）后，境内有学校318所（小学中心校28所），教学点58个。在校学生74 992人，教学班2 334个（复式班100个，二部制班56个），教职工3 517人。

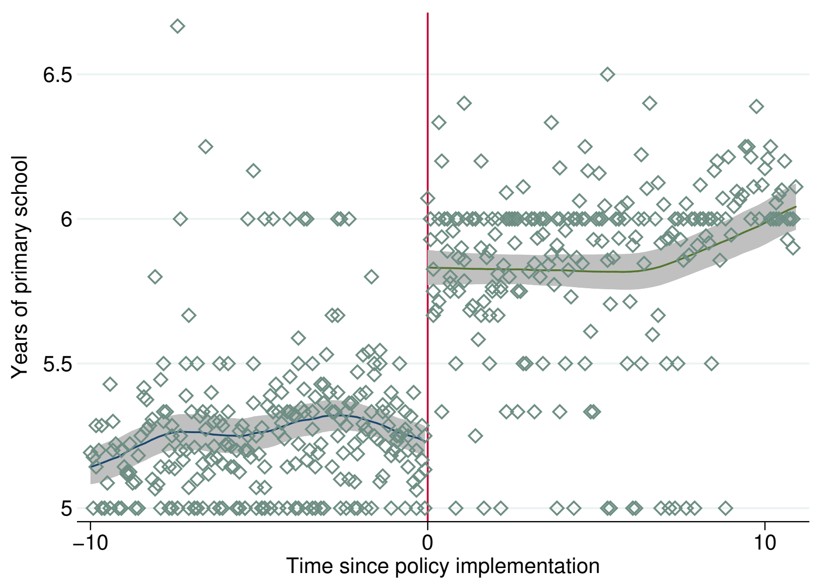
1989年，城内小学建成2座教学楼，建筑面积6 460平方米。其中，实验小学教学楼建筑面积3 260平方米；第四小学教学楼建筑面积3 200平方米。

Figure A.4: Example of mean shift algorithm identifying year of policy implementation



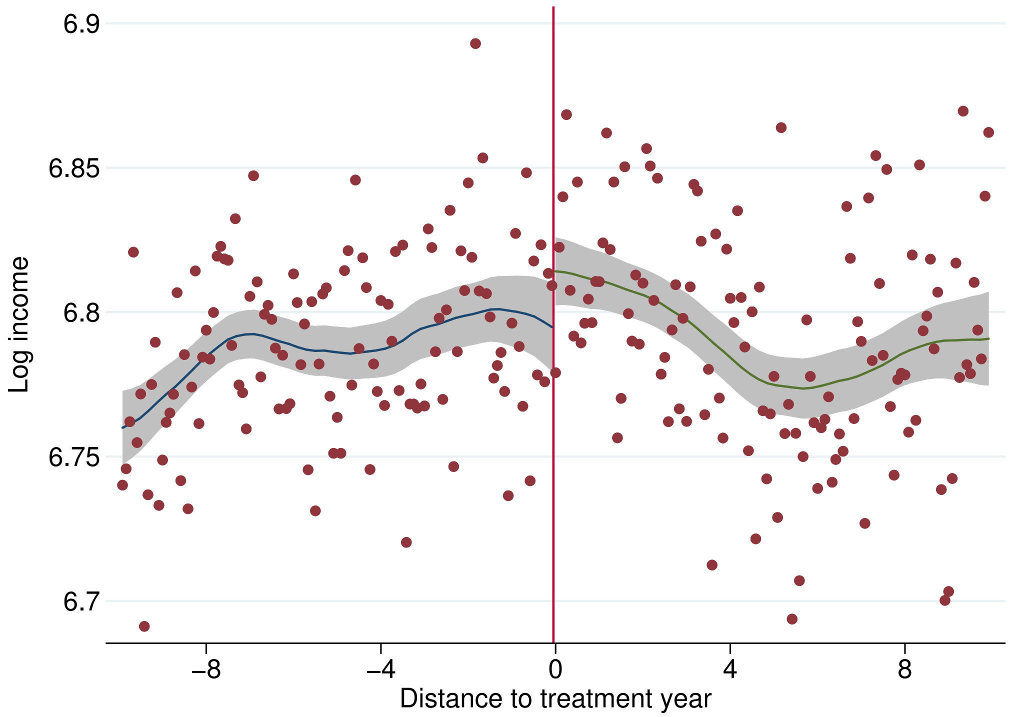
This figure shows the histogram of cohort-mean years of primary school in a county and plots the ssr estimates generated by Equation 1 for each treatment year. The ssr sequence reaches its minimum at 1997, where we also observe a clear shift upwards in mean years of primary school from approximately five to six.

Figure A.5: First stage, month-by-year bins



Data: CFPS. This plot shows the month-by-year-to-treatment bin means of years of primary schooling and a locally smoothed polynomial with its confidence interval, estimated separately for the treated and untreated groups.

Figure A.6: Event study of log income



Data: census. This figure is the more-detailed analog to Figure 3 in the main text. It shows the month-by-year-to-treatment bin mean of log income over the nine years before and after the policy is implemented (stacked across different implementation years) along with a locally smoothed polynomial and the related confidence interval, estimated separately for the treated and untreated group.

Appendix 2 - Appendix tables

Table A.1: Anecdotes on implementation of six year primary education from across China

Locality	Strategy of Changing 5-Year to 6-Year Primary Education	Source
Beijing	In 1969, the duration of primary education was shortened to five years. Starting from September 1st, 1980, the duration of primary education was prolonged to six years, with the fifth-grade students continuing to be in the old system and students of other grades entering into the new system.	Beijing General Education Chronicle (Part I)
Xinle County, Hebei Province	In 1967, the duration of primary education was shortened to five years. Starting from August 1985, first-grade students and one half of second-grade students entered into the new six year system, while the rest of students remained in the old system.	Xinle Educational Chronicle
Nanjing City, Jiangsu Province	In 1969, the duration of primary education was shortened to five years. Starting from 1982, the duration of primary education was prolonged to six years in urban districts and children started primary education at the age of six. By contrast, the duration of primary education remained to be five years until 1999 in five other counties (Jiangning, Jiangpu, Liuhe, Gaochun, and Lishui).	Nanjing Educational Chronicle
Wuyi County, Zhejiang Province	In 1984, first-grade students entered into the new six year system in half of primary schools, with the other half following the next year. In 1987, however, all primary schools were required to resume the five year system. In September 1999, all primary schools were restored to the new system of six years. The final cohort under the old system graduated in June 2004.	Wuyi Educational Chronicle
Dongying District, Shandong Province	In 1997, the compulsory education system changed from the 5-3 (years of primary school-years of middle school) to the 5-4 system. In 2003, first-grade students entered into the new 6-3 system.	Dongying District Chronicle (1998-2005)
Xishui County, Hubei Province	In 1986, the first-grade students entered into the new 6-3 system in the primary schools located in county seats, while other primary schools remained in the old 5-3 system. In 1987, the first-grade students in the primary schools located in the township entered into the new 6-3 system. In 1991, first-grade students in the remaining primary schools entered into the new 6-3 system.	Xishui Educational Chronicle (1986-2006)

Table A.2: Data sources

Data source	Year(s) collected	Sample size	Relevant contents	Policy year identification method	Geographic specificity
County educational gazetteers	Various	345*	Date policy implemented, implementation details	N/A	County**
Mini-census	2005	2.59m	Schooling: highest credential, dropout labor market: income, employment	Archival records	Prefecture**
China Family Panel Studies (CFPS)	2010	33,600	Schooling: years spent in each level of schooling	Algorithm	County
China Labor-Dynamics Survey (CLDS)	2012	16,253	Schooling: years spent in each level of schooling	Algorithm and archival records	Prefecture

*Number of prefectures for which we collect gazetteer data
**Note: There are multiple counties in each prefecture

Table A.3: Summary statistics

	CFPS		*Census*	
	(1)	(2)	(3)	(4)
Outcome	Urban	Rural	Urban	Rural
Demographics				
Female	0.491	0.471	0.516	0.519
Ethnic minority	0.061	0.068	0.071	0.174
Highest credential				
Primary school	0.056	0.310	0.037	0.319
Middle school	0.323	0.516	0.333	0.606
High school	0.282	0.146	0.321	0.067
Post-secondary	0.339	0.028	0.309	0.008
Labor market characteristics				
Currently working	0.689	0.628	0.775	0.905
Works for government	0.278	0.030	0.269	0.009
Number of observations	1,164	2,240	107,422	199,126

This table presents descriptive statistics of demographic, educational, and labor market characteristics for our two main datasets using the optimal bandwidth estimation sample. This is the same sample used to generate the regression estimates presented in the paper's main tables.

Table A.4: Regression tests of continuity of conditional expectation of outcome variable

Outcome	CFPS	2005 Census
Female	-0.027	0.001
	(0.037)	(0.006)
Urban hukou	0.016	0.020
	(0.032)	(0.021)
Relative cohort size	-1.381	-92.909
	(1.647)	(69.246)
Observations	3,404	306,548

This table estimates Equation 2, using the predetermined characteristic as the outcome variable, to test for our main identifying assumption. Coefficients from six separate regressions are presented, estimating the "effect" of the policy on the predetermined characteristic listed in the first column for the census data in the second column and for the CFPS data in the third column.

Table A.5: Effect estimates for rural residents' labor market outcomes

	(1) Currently employed	(2) Works for government	(3) Log monthly income
Panel A: overall estimates			
Treated (β_1)	0.004	-0.000	0.000
	(0.003)	(0.001)	(0.009)
R-squared	0.074	0.011	0.335
Number of observations	194,613	176,098	175,580
Panel B: by highest credential			
Primary school	0.005	-0.000	-0.001
	(0.004)	(0.001)	(0.010)
Middle schoool	0.006*	-0.000	0.002
	(0.004)	(0.001)	(0.009)
High school	-0.018***	0.000	0.008
	(0.007)	(0.004)	(0.016)
Post-secondary	-0.084***	-0.066***	-0.080**
	(0.020)	(0.028)	(0.037)
R-squared	0.081	0.084	0.345
Number of observations	185,907	168,683	168,172
Panel C: by gender			
Female	-0.010***	-0.000	0.022***
	(0.004)	(0.001)	(0.009)
Male	0.017***	-0.001	-0.018**
	(0.003)	(0.001)	(0.009)
R-squared	0.082	0.083	0.345
Number of observations	185,907	168,683	168,172
Panel D: by employer			
Private sector	-	-	-0.001
			(0.009)
State-owned enterprise	-	-	0.006
			(0.015)
Government	-	-	0.046*
			(0.027)
R-squared	-	-	0.349
Number of observations			175,580

This table generates labor market results for rural residents analog to those presented for urban residents in Tables 3 and 4.

Appendix 3 - Details on Our Use of the Mean Shift Model

As we use an algorithm to determine treatment year status for observations in the CFPS, we introduce a few safeguards to ensure that we are not erroneously choosing a year when there is no change in treatment status or not enough evidence to determine the year. Specifically, we exclude counties in which:

1. The cohort-level median number of years of primary school in a given county is never less than 5.5 before the identified treatment year or is never more than 5.5 after;

2. The mean shift model predicts a treatment year after which individuals have less primary schooling, that is, $\gamma_1 < 0$ in Equation 1;

3. There are more than four cohorts for which we have no observations within five years before or after the identified treatment year (that is, data for at least five of the twelve cohorts in this county's restricted bandwidth sample are missing).

This exercise excludes 32 of the 144 non-Shanghai counties, or about 22% of the sample. All of the results we show are qualitatively similar with the entire 144 county sample and are available by request.

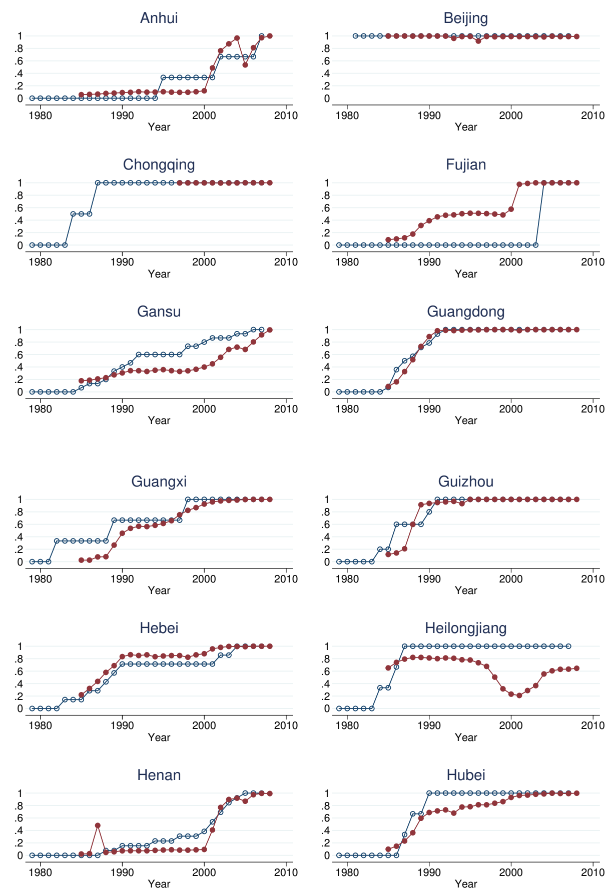
As a final check of accuracy, we visually compare the treatment year generated by the mean shift model to each county's histogram. In four cases, our inspection of the histograms suggested a year more than one year different than the mean shift model's choice and we use this visually identified year in our main analyses, controlling for those counties whose year was chosen visually as opposed to algorithmically.

To check that these algorithmically identified years are credible, we conduct two exercises. First, we compare the mean shift years for all counties within each province to national statistics on provincial-level implementation, shown in Figures A.7 and A.8. The two series track closely and their correlation is 0.7759. Second, we apply both the mean shift and archival match policy year identification methods to the China Labor-force Dynamics Survey data. Figure A.9 shows the distribution of the difference between archival and algorithmically identified treatment years.

These values are normally distributed around 0, and in 67% of prefectures the years identified by the two methods are within 3 years of each other. This amount of discrepancy is not surprising, as in the gazetteers we see that counties occasionally implemented at different times within a prefecture. Additionally, the number of observations in cohort-by-prefecture bins in the Labor-force Dynamics dataset is small, which implies imprecision (relative to using the CFPS) in our ability to algorithmically determine the implementation year.

Figure A.7: National statistics and province-level CDFs of policy implementation

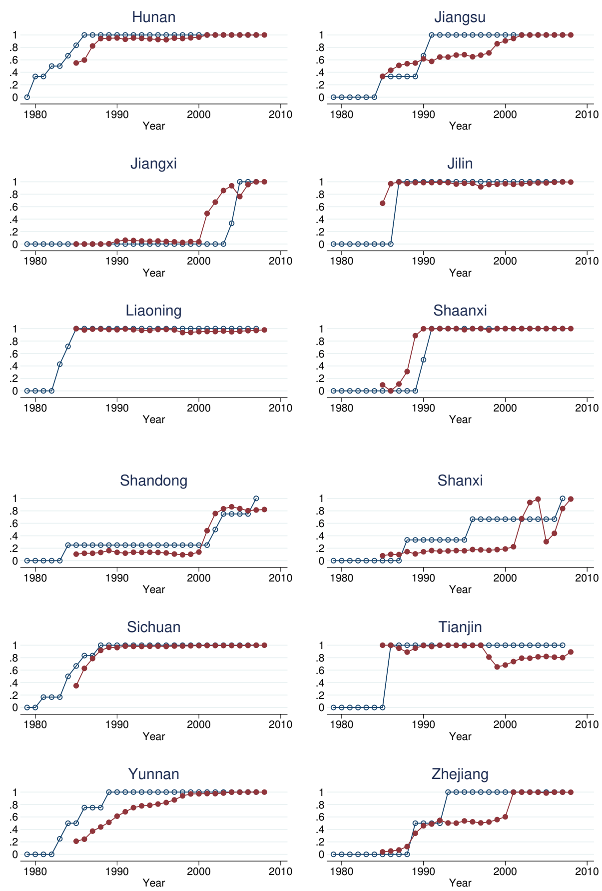
First 12 provinces



These plots show the proportion of individuals enrolled in six-year primary schools in the 24 non-Shanghai CFPS provinces. The sequence of hollow circles represent the proportion of treated counties in a given province by year according to county-specific implementation years identified in the CFPS data. The sequence of solid circles are from publicly available, Chinese Ministry of Education statistics on the number of schoolchildren benefiting from six year education in a given province by year. Measurement error and policy experimentation cause these to briefly trend downward in a few cases.

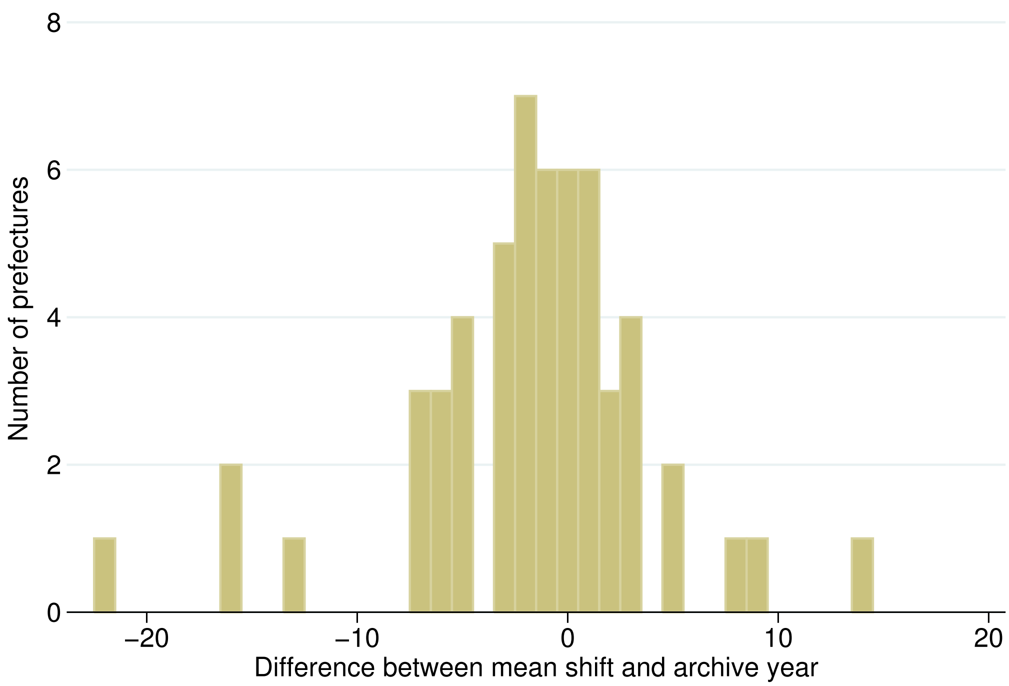
Figure A.8: National statistics and province-level CDFs of policy implementation

Second 12 provinces



These plots show the proportion of individuals enrolled in six-year primary schools in the 24 non-Shanghai CFPS provinces. The sequence of hollow circles represent the proportion of treated counties in a given province by year according to county-specific implementation years identified in the CFPS data. The sequence of solid circles are from publicly available, Chinese Ministry of Education statistics on the number of schoolchildren benefiting from six year education in a given province by year. Measurement error and policy experimentation cause these to briefly trend downward in a few cases.

Figure A.9: Comparing archival and algorithmically identified treatment years



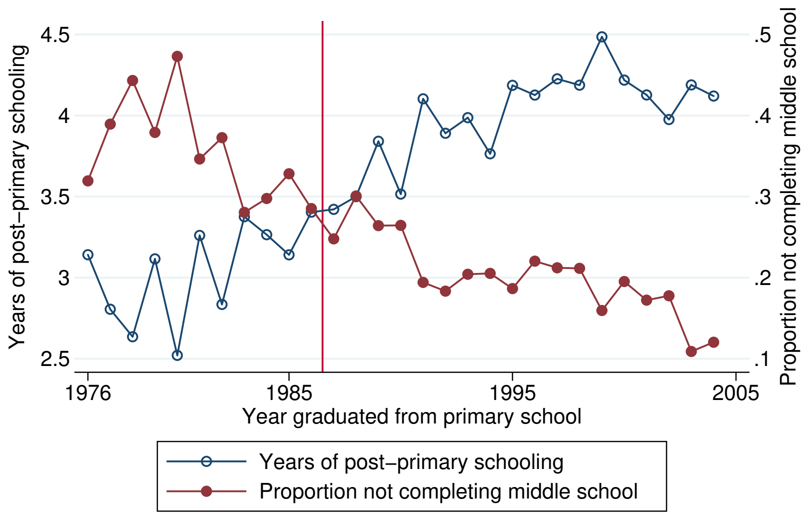
Data source: CLDS

Appendix 4 - China's 1986 compulsory education law

China's 1986 compulsory education law stipulated that all Chinese students had to complete primary and middle school. While the policy could conceivably have generated much of the bunching at the middle school credential attainment year we observe, we estimate that it is unlikely to have done so. Recent empirical work shows the law was gradually implemented and remains porously enforced, with many places not implementing it until the mid-2000s (Fang et al., 2012). Using province-specific policy implementation years from that study, we estimate the policy's rollout has no effect on whether individuals attain a middle school credential or on the number of years they spend in primary school[29]. We believe this null result is driven by porous enforcement of the law in rural areas and the fact that, in urban areas, high educational attainment levels (as seen in our data) make the law non-binding for the vast majority of affected individuals. Further supporting the claim of porous enforcement in rural areas, recent work has shown that as recently as 2010 the proportion of students not completing middle school was over 20% in some areas of rural China (Mo et al., 2013). In Figure A.10, we show that the cohort-specific proportion of students who get less than a full middle school education (three years of post-primary schooling) is stable at around 20% throughout most of the period we study.

[29]Fang et al. (2012) estimate that the policy generated a 0.8 year increase in the average number of years spent in middle school. Though we can replicate their results using their chosen empirical specification, when we apply the treatment years identified in that study to our data using our RD design, we find no significant difference between treated individuals in terms of either years spent in middle school or the likelihood that individuals completed at least a middle school degree.

Figure A.10: Evidence of students' ability to adjust on lower margin



*Vertical line at annoucement of compulsory middle school
Data source: CFPS

Appendix 5 - Difference in differences results

In this appendix, we present difference-in-differences results. The equation we estimate is

$$Y_{iclp} = \beta_0 + \beta_1 * Treated_{cl} + \xi_m + \mu_l + \eta_{cp} + \varepsilon_{iclp} \tag{5}$$

The notation is nearly identical to our estimating equation (Equation 2). The point of departure is that, since we are not stacking the discontinuities, we do not include pre- and post-distance-to-treatment linear trends. In this DD design, all localities are eventually treated, so the $Treated_{cl}$ indicator is equivalent to the $Treated * Post$ indicator commonly emphasized in most work using a DD design (Bertrand et al., 2004).

In Table A.11, we present two sets of difference-in-differences results to mirror our main results in the body of the paper: in column one, we show estimates for the same restricted bandwidth sample that we use in our RD analysis, which also gives us a balanced panel in terms of distance to policy implementation; in column two, we use the entire sample within the cohort range specified in Section 3.3.

We give more credence to the column 1 results, as column 2 has several issues which impede clear interpretation. First, as described in the body of the paper, we are concerned about the likely violation of the parallel trends assumption. The full sample uses an unbalanced panel, which means that early implementers have more treated observations and fewer untreated observations, and later implementers have the reverse. This may exacerbate the risk of violation of the parallel trends assumption if, as documented in descriptions of the policy, timing of implementation is correlated with locality-specific economic conditions that affect schooling decisions or the labor market outcomes we attribute to the effect of the policy. Second, the full sample also includes controls and treated groups far from the implementation of the policy, which increases the time window in which other policies could be implemented, another potential source of bias.

Our results in column 1 largely mirror our main findings using the RD design. We find a similar effect on years of primary and post-primary schooling, small positive effects on credential attainment at the middle school and high school level. We find no effects on the probability of

Figure A.11: Difference-in-differences results

Outcome	(1) Balanced panel / optimal bandwidth	(2) Full sample
CFPS		
Years of primary schooling	0.620***	0.580***
	(0.039)	(0.029)
Years of post-primary schooling	-0.105	-0.010
	(0.143)	(0.115)
Number of observations	3,404	8,021
Census		
Education results		
Highest credential: at least middle school	0.010***	-0.008***
	(0.004)	(0.003)
Highest credential: at least high school	0.018***	0.031***
	(0.004)	(0.003)
Number of observations	306,548	659,537
Labor market results		
Currently employed	0.003	-0.001
	(0.006)	(0.005)
Log income	0.022*	0.012
	(0.011)	(0.008)
Number of observations	82,267	153,125

Notes: each cell in this table which is not preceded by the title "observations" presents a regression estimate of β_1 from Equation 5. Data source and dependent variable given in leftmost column.

employment, and a similar effect (2.2 percent) on log income. In column 2 we find similar effects, though the magnitudes change - larger effects for educational attainment and smaller effects for log income.

Appendix 6 - Details of cost-benefit analysis

This section outlines our data sources and extrapolation assumptions for the cost benefit analysis we perform. The costs of the policy we include consist of one year of forgone wages for all affected individuals. We assume the value of one individual's forgone year of wages is equal to $\frac{\alpha * GDP_t}{N_t}$, where N_t is the size of the active workforce in year t. We calculate when the forgone year falls for each individual born on or after 1969 (as the first affected cohort would be leaving primary school in 1981). To calculate the total number losing a given year in each cohort, we multiply the total number of individuals in each cohort (generated using data from the Chinese National Statistical Bureau and World Bank World Development Indicators data) by the proportion of individuals with a given credential in each cohort (calculated from the CLDS data). We assume the lost year is the year a person would have entered the labor force in the absence of the policy, e.g. for someone who was born in 1980 and finishes school with a middle school credential, we calculate the value of the work they would have done in 1995 when they would have left school in the absence of the policy (starting school at age 7, spending five years in primary school and three years in middle school). We assume that those leaving school after completing primary school do not lose a year of work, as they will leave school at the age of 12 and are not likely to immediately enter gainful employment. For simplicity, we assume that everyone enters school at age 7.

The benefits from the policy consist of the boost in earnings we estimate for affected individuals accruing throughout their lives. This parameter is taken directly from our estimates in Section 5. We assume that the benefits affected individuals enjoy are constant throughout their lifetimes, and calculate the total value of the earnings gain for each affected individual up to the year 2050. Table A.6 provides our data sources for each of these figures and assumptions used to extrapolate into the future and where data is not available.

We exclude a few potential sources of costs and benefits, believing them to be several orders of magnitude smaller than the value of the lost year of wages and of the lifetime income gains. On the costs side, we exclude the cost of hiring new teachers and the cost of building new facilities. We do so because archival records suggest that in the short term, the current staff was used to meet most of the increased personnel needs imposed by the policy and classrooms could be split, thus requiring no sizable infrastructure additions. In the long term, the one child policy reduced the number of children in schools, and so existing staff levels and infrastructure house a smaller and smaller number of children. We assume that this would largely offset the need for long term staff and infrastructure adjustments to accommodate the extra cohort of primary students. On the benefits side, we disregard the possible benefit of the creation of new teacher jobs, as our assumption about the staffing costs of the policy implies very few additional jobs would be created. As in Duflo's approach, we exclude the possible beneficial impact of the extra year of schooling on health, fertility and other non-financial outcomes.

Table A.6: Cost-benefit calculation details

Data point	Source	Range of data	Range for extrapolation	Assumptions used	What used for	Notes
Birth rate per 1,000 women	stats.gov.cn	1966-2013	N/A	N/A	Size of cohort	Assume: cohort size constant 2013-2050
Number of women	stats.gov.cn	1970-2013	N/A	N/A	Size of cohort	-
Total labor force	World Bank WDI (World Dev. Indicators)	1981-2015, five-yearly estimates for 2020-2050	Gaps in estimates for 2020-2050	Linear trend between estimates	Total labor force	Close to stats.gov.cn figures, but with predictions to 2050
Proportion of cohort with each education level	China Labor-force Dynamics Survey (CLDS)	Birth cohorts 1969-1987	Birth cohorts 1987-2035	Linear trends from previous 9 years and rules of probability $(0 \leq P \leq 1)$	CB estimates with cohort-specific effects	Begin extrapolation at 1987 cohort to avoid measurement error
Proportion of cohort affected by policy	China's National Bureau of Statistics	1981-2010	2010-2050	Full coverage after 2010	Proportion affected by policy	Equate 5 year primary + 4 year middle with 6 primary + 3 middle
Alpha (labor's share of GDP)	Karabarbounis and Neiman (2014)	1992-2009	1981-1991 2010-2050	1981-91: average of t+1 to t+3; 2010-50: average of t-1 to t-3	Argument in final calculation	-
GDP (constant 2005 US$)	World Bank WDI	1981-2013	2014-2050	Growth rate declines annually from 2013 rate to 4%, annual decrease of 0.15%	Argument in final calculation	-