I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16763

# Hukou Status and Children's Education in China

Yue Sun
Liqiu Zhao
Zhong Zhao

JANUARY 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16763

# Hukou Status and Children's Education in China

**Yue Sun**
*Renmin University of China*

**Liqiu Zhao**
*Renmin University of China*

**Zhong Zhao**
*Renmin University of China and IZA*

JANUARY 2024



ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Hukou Status and Children's Education in China*

Under China's household registration (hukou) system, children with rural hukou do not have equal rights to access education in urban areas. This paper investigates the causal effect of hukou status on children's education by exploiting an exogenous change in hukou status induced by the hukou reform in 1998. Before the reform, children could only inherit their mother's hukou status. After 1998, newborns and preschoolers could inherit either their father's or mother's hukou status, which provided a unique exogenous opportunity for children with urban fathers and rural mothers to obtain urban hukou. Using China's 2010 population census data, we employ a difference-in-differences strategy to examine the impact of hukou status on children's education. We show that the younger cohorts exposed to the reform are 15.1 percentage points more likely to have urban hukou and are 18.9 percentage points more likely to be at the grade level appropriate for their age. Moreover, the effect is more pronounced amongst girls, and children from educated families or from large cities.

**JEL Classification:** I24, I28, O15, R28

**Keywords:** Hukou reform, grade-for-age, education equality, rural-urban disparity

**Corresponding author:**
Zhong Zhao
School of Labor and Human Resources
Renmin University of China
Beijing 100872
China

E-mail: mr.zhong.zhao@gmail.com

* This research is supported by the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (Approval No. 21XNLG03) under the "Thematic Research Project on China's Income Distribution". We would like to express appreciation for comments from two anonymous referees, and the editor Marcel Fafchamps.

## 1 Introduction

Since the government's establishment of the household registration (hukou) system in 1958, all Chinese citizens have been mandatorily divided into agricultural (rural) and non-agricultural (urban) hukou (Song 2014; Pan 2017).[1] Hukou status determines an individual's access to government-provided public services and welfare programs such as public housing (Hui, Yu, and Ye 2014), social security (Huo and Lin 2022), and education (Zhang 2017). Particularly, children with rural hukou do not have equal rights to attend urban schools, which negatively impacts their human capital accumulation (Heckman 2005). The gap in the enrollment rate between rural and urban populations widens as the level of education increases, see panel A of figure 1. Moreover, the gap in years of schooling between the urban and rural working-age population (aged 16-64) have not narrowed over time, see panel B of figure 1. The persistent urban-rural disparities in educational attainment partly reflect unequal access to education among urban and rural hukou holders. This paper investigates the effect of the hukou status of children on their educational outcome by exploiting the hukou reform in 1998, which provides a natural experiment to identify the causality between hukou status and education.

(Insert figure 1 here)

Since the 1950s, children were tied to their parent's birthplace. Newborns were granted the same hukou status as their mothers but not their fathers until the hukou reform in 1998 (Han, Li, and Zhao 2015). Since China's reform and opening-up in 1978, urban areas have seen an increasing demand for inexpensive labor. Simultaneously, rural residents aspired to find

[1] The promulgation of the *Household Registration Regulations of the People's Republic of China* in 1958 marked the establishment of the dualistic rural-urban hukou system.

opportunities in economically developed cities. In order to meet the demands of economic development, population migration, and parents' desire to have more control over their children's hukou status, the government issued a document on July 22, 1998 to reform the hukou system (Pan 2017). The new policy permits parents to decide whether their children's hukou status is inherited from their mother or father.[2] Specifically, children born after the reform can freely choose whether to inherit their father's or mother's hukou; minors born before the reform can also apply to change their hukou status, but preschool-age children are given priority by the government. Therefore, before the reform preschool-age children with a rural mother and an urban father were only able to have rural hukou. After the reform, these children were given the option to have urban hukou. Access to an urban hukou may improve children's education by enhancing the quality of education, reducing the cost of education, and raising educational aspiration. This reform provides a unique opportunity to identify the causality between children's hukou status and their educational outcomes in China.

Using China's 2010 population census data, we employ a difference-in-differences approach, and find that the younger cohorts exposed to the reform are 15.1 percentage points more likely to have urban hukou and 18.9 percentage points more likely to be at the grade level appropriate for their age. Specifically, the reform increases children's probability of attending senior high school at the expected age by 26.2 percentage points. However, the reform shows no significant effect on educational attainment at the primary and junior high school levels. In

[2] Some studies have investigated the impacts of the 1998 hukou reform. For example, Han et al. (2015) show that the change to the hukou inheritance law raises the rate of interprovincial marriages between local urban men and migrant women. Pan (2017) finds that directly granting urban hukou reduces rural youth's high school attendance.

addition, girls benefit more from the reform than boys, and the educational impact is pronounced amongst children from educated families or from large cities.

This paper contributes to existing literature in three ways. Firstly, this paper identifies the causal impact of hukou status on educational attainment by exploiting the 1998 hukou reform. Most existing studies ignore the endogeneity of hukou status (Wu 2010; Ye 2015; Huang, Xie, and Xu 2015). However, hukou identity can be both a cause and a consequence of educational achievement, which may cause reverse causality issue.[3] Furthermore, unobservable factors, such as parental willingness to pay for education, affect both hukou status and educational achievement.[4] To the best of our knowledge, Wu (2012) attempts to identify the causal impact of hukou on education. Nevertheless, the instrument, i.e., mother's occupation when the child is aged 14, might be correlated with unobserved ability or preference.[5]

Secondly, rather than focusing on the end results of education (Wu 2012), we construct the *Grade-for-age* (*GFA*) variable to investigate the overall process of education at different stages. We investigate the stage at which hukou transformation plays a more critical role in promoting educational equity through exploring the effects of hukou on the grade-for-age status at primary, junior high, and senior high schools, respectively.

Thirdly, this study is the first to evaluate the hukou inheritance policy reform in 1998 from the perspective of educational equity. Unlike previous studies that have focused on the impact

[3] Urban hukou facilitates individuals to achieve higher levels of education, and access to higher education is a means of converting from rural to urban hukou (Wu and Treiman 2004).

[4] For example, some parents assist their children to obtain an urban hukou by purchasing a house in the city, and such families usually have more resources and are more willing to make other educational investments such as extracurricular tutoring. With this in mind, a child's hukou and education may be determined simultaneously.

[5] Specifically, Wu (2012) uses endogenous switching regression models. The results show that if actual rural hukou holders were randomly assigned to urban hukou, they would have completed an average of 2.79 more years of schooling.

of this hukou reform on marital matches or educational decisions in the short term (Han et al. 2015; Pan 2017), this study examines the long-term effects of the reform on the educational attainment of children from target families and explores how the educational promotion effects of the reform change over time.

The remainder of this paper is organized as follows: Section 2 introduces the institutional background; Section 3 describes the data and key variables; Section 4 outlines the empirical strategy; Section 5 presents the main results and examines the validity of the DID strategy; Section 6 carries out several robustness checks and the heterogeneous effect analyses, and discusses mechanisms; Section 7 concludes.

## 2 The reform of the hukou system in China

### *2.1 The hukou system and urban-rural education divide*

The household registration (hukou) system was introduced by the Chinese government during the 1950s to control migration from rural to urban areas. Overall, the hukou system aimed to facilitate the industrialization process and support industry-oriented policies by helping to ensure sufficient food supply to urban areas and the accumulation of capital.

Prior to the reform and opening-up in 1978, the hukou system was strictly implemented, population mobility from rural to urban areas was extremely limited and a change to hukou status was virtually impossible.[6] Those with rural hukou could only engage in agricultural production, were forbidden to work in urban areas, and hukou status transmitted from

[6] For example, rural people could only obtain urban hukou if employed by state-owned enterprises (SOEs), in pursuit of a college education, or joining the military. They then enjoyed the benefit of receiving public services in the urban areas (Song 2014).

generation to generation. Children with rural hukou could not access quality educational resources in urban areas, which was the root of the urban-rural education gap.

After the reform and opening-up, the influx of foreign capital and the development of labor-intensive industries generated a tremendous demand for low-skilled labor (Meng and Xue 2020). Subsequently, the hukou system was gradually relaxed to enable rural workers to migrate to cities. However, they remained with rural hukou and were considered "guest workers" and so lacked equal access to key public services, including education (Meng 2012).

Along with the socio-economic development, China has gradually reformed its hukou system. Rural hukou and urban hukou are in the process of unification.[7] However, even today, migrants and urban hukou holders continue to have unequal access to essential public services. There remains a vast difference between school and educational quality in rural and urban areas, which amplifies the rural-urban divide caused by the hukou system. Existing literature explores the following three channels behind such educational disparity.

Firstly, high quality education in urban schools decreases grade repetition and dropout rates. Since the fiscal decentralization in the 1980s, local governments have been responsible for administering and financing basic education, which causes a substantial variation in educational expenditure across urban and rural areas (Hannum 2003). Whilst the ratio of rural to urban education expenditure per student may have risen dramatically since 2008, it remains far below 100%, see figure 2. A poor quality of educational provision in rural schools has meant increased dropout rates or grade repetition (Brown and Park 2002). Access to urban

[7] By 2012, this policy had been implemented in 12 out of 31 provinces in China, such as Hebei, Zhejiang, Jiangsu, and Guangdong.

public education means children are much less likely to fall behind once they have received urban hukou.

(Insert figure 2 here)

Secondly, having an urban hukou can reduce the cost of education for children. Whilst rural children can move to urban areas with their parents, the barrier to attending high-quality public schools remains because they do not have local hukou. For example, in the Rural Urban Migration in China (RUMIC) survey conducted in 2008, 61.5% of migrant children report extra school fees due to their non-local rural hukou; over 20% of temporary migrant children pay more than 2,000 CNY, accounting for more than 11.82% of the average annual income; and 6% pay more than 5,000 CNY, accounting for more than 29.54% of the average annual income (Zhang 2017).[8] The hukou reform in 1998 enabled more rural women to settle in urban areas (Han et al. 2015), and their children could have the same hukou status as their urban fathers; hence their children no longer needed to pay extra schooling fees.

Thirdly, hukou transformation may change an individual's education decisions by raising their educational expectations. On the one hand, the financial returns to education for urban residents are higher than for migrant workers with rural hukou, hence hukou status transformation can improve young people's willingness to pay for education and stimulate their educational aspirations (Qu and Zhao 2017).[9] On the other hand, the reforms of the hukou system provide extensive policy support for migrant children to receive education in urban

[8] The RUMIC is a large-scale survey on Chinese rural-urban migrants conducted by the Australian National University, Beijing Normal University, and the Institute for the Study of Labor (IZA). The data are publicly available on http://idsc.iza.org/?page=27&id=58. In 2008, the average annual income of migrant workers was about 16923.42 CNY or 2438.53 USD (the average exchange rate was 6.94 CNY per 1 USD in 2008).
[9] There was a sharp increase in returns to education from about 4% in the early 1990s to nearly 11% in 2000, which roughly equaled the same level as in western economies (Heckman 2003).

areas. The expectation of fair access to local senior high schools improves the cognitive and non-cognitive outcomes of rural hukou students, which helps improve the educational attainment of the next generation (Guo and Zhao 2019).

### *2.2 The hukou reform in 1998*

On July 22, 1998, the State Council approved the *Opinions on Solving Several Outstanding Problems in the Current Management of Hukou*. The first task was to revise the rules on hukou inheritance for children. Prior to the reform, children could only inherit their mother's hukou, which meant that irrespective of whether their father was an urban male, they were only entitled to a rural hukou if their mother was a rural female and could not access educational resources in urban areas. Since the reform, parents are able to decide whether their children's hukou status is inherited from their mother or their father and permits those children born before the reform to change their hukou status. It is noteworthy that the reform does not alter the hukou policy on children over 18 years old.

Table 1 summarizes the specific rules set out by the hukou reform. There are different policies for children born in different cohorts. Those born on or after July 22, 1998 could voluntarily inherit the hukou status of their father or mother if they had not already registered. Even if they had already registered under their mother's hukou status, they could apply for a transfer to their father's hukou status with no conditions attached. By comparison, those children born before July 22, 1998 who had yet to register, had to register under their mother's hukou status. If they had already inherited their mother's hukou status, they could apply for transfer to the father's hukou status, but there were certain age restrictions with preschool children given priority. For example, the document issued by the government of Guangdong

province clearly stated that applications to inherit the father's urban hukou would be approved according to the following priorities: preschool children and then minor children who were already in school. However, adult children were not allowed to transfer to their father's hukou status. Hukou status is tied to access to urban educational public goods, so the reform has only increased the accessibility to urban quality educational resources for preschoolers in 1998 and those born after the reform.

(Insert table 1 here)

The central government ordered the new policy to be rolled out nationwide and stipulated that all provinces would implement the new policy since July 22, 1998. Every province must formulate related policies, taking into consideration local economic and social development levels, and the local government's financial capability and affordability.[10] More importantly, no local government was allowed to add additional restrictions when implementing the policy. We report the dates of document publication as well as the first affected cohort at the provincial level in table A1. Specifically, the implementation of the hukou reform is characterized by three key features. (1) Despite the differences in the timing of the policy announcement, all provincial governments, except Heilongjiang, Tianjin, and Shanghai, applied July 22, 1998, as the effective date (see column 2 of table A1). As individuals' birth dates are reported in months, the effective dates of 22 provinces and Heilongjiang are considered to be the same, i.e., July 1998. (2) Considering the high population pressure in megacities (e.g., Beijing and Shanghai), these municipalities can implement the reforms later. Specifically, four municipalities directly under the central government had not enacted the new regulation until July 2003. (3) The

[10] Province here includes province, autonomous region, and provincial-level municipality.

information on the effective dates of the policy implementation in six provinces – Tibet, Xinjiang, Hainan, Yunnan, Jiangsu, and Qinghai – is missing.[11]

## 3 Data and key variables

### 3.1 Data

Our primary source of individual-level data is the 0.095 percent sample of China's population census conducted in 2010. The information on the date of the first marriage of adults allows us to identify couples. We then use the variable "relationship to the household head" to match children with their parents and siblings.[12] A total of 406,213 parent-child relationships are matched, and the whole sample is divided into four family types according to the parents' hukou status, see table A2. Those children whose fathers have urban hukou will benefit from the reform if their mothers have rural hukou (type 1); the children whose mothers also have urban hukou will not be affected by the reform (type 2). In addition, households with rural fathers and urban mothers (type 3) may not be very common, as the popular marital practice in China is characterized as "marry up." In our sample, only 1.16% of children have rural fathers and urban mothers. Furthermore, children with rural fathers and rural mothers (type 4) may not be an appropriate control group, as they are differently affected by various reforms. For example, the government implemented a series of educational reforms targeting rural students after 1998, such as mergers and closures of rural schools and free compulsory education reform.

[11] According to the central government, all provinces are required to implement the reform. Therefore, the fact that the document did not remain on the webpage does not mean there was no reform in these provinces, only that the exact date of the reform could not be determined. We also include individuals from these six provinces in the robustness checks, and the results are consistent with the baseline regression.

[12] We only retain those households where the head of the household is the father or mother. Otherwise, we cannot identify the number of siblings with certainty.

Moreover, the 9-year compulsory education reform in 1986 had a stronger impact on the education of rural residents as they had lower educational levels prior to the compulsory education reform. Therefore, the treatment group comprises the children with urban fathers and rural mothers (accounts for 2.48%) and the control group comprises children with urban parents (accounts for 26.33%); we obtain an initial sample of 117,009 children.

*Regression sample.* The sample entering regression needs to meet the following five conditions (see table A3). (1) The marital sorting pattern has changed since the hukou reform in 1998 (Han et al. 2015).[13] Therefore, in order to close the potential channel through which the hukou reform may affect children's education by influencing marital matches, we retain the sample with parents whose first marriage occurred in or before 1998 (94,567 samples remained). (2) In order to ensure that the children had reached primary school age when the survey was conducted in November 2010, we set August 2004 as the lower limit of birth date of children. (3) We exclude the sample aged 7-17 years when the reform starts because their treatment status is unclear. Instead, we select individuals aged 18-30 years in August 1998 as the control group because their hukou status is not subject to policy intervention (51,750 samples remained). (4) We drop the samples with missing key variables, such as child's age, gender, ethnicity, residence, number of siblings, parents' age and education, and government's public spending on education (51,197 samples remained). (5) We delete the sample from municipalities and provinces with unclear policies, i.e., Beijing, Tianjin, Shanghai, Chongqing, Tibet, Xinjiang, Hainan, Yunnan, Jiangsu, and Qinghai (41,034 samples remained). Eventually,

[13] After the hukou reform in 1998, more urban men married rural women. Inter-hukou marriage households had less floor space and fewer rooms in their houses, and wives had fewer years of education than their husbands (see table R1, this table is for review only, does not intend for publication).

the sample sizes of the treatment and control groups that remain after screening are 3,604 and 37,430, respectively.

### 3.2 Key variables

*Educational outcome*. As the first cohort fully exposed to the policy was 12 years old when the census was conducted in 2010, we could not obtain information on the individuals' completed years of schooling. In addition, the survey only asked individuals about their "highest level of education" and information on the child's precise grade is not available. Therefore, data limitations mean we cannot measure how far behind they are and can only determine whether they are entering an age-appropriate stage of education.

According to human capital theory, the opportunity cost of education is small in the early stages of the life cycle. Therefore, it is optimal to enroll at the appropriate age during childhood. Moreover, earlier enrollment means that individuals have a potentially longer time frame in which to reap financial gains in the labor market and therefore enjoy higher lifetime earnings (Becker 1993). In the United States, grade retention means it is more difficult for a child to achieve long-term success, and they are less likely to graduate from high school and enroll in college (Andrew 2014). In China, a one-year delay in primary school enrollment reduces the probability of high school enrollment by 6 percentage points for rural children (Chen 2015). As delayed enrollment is widespread in rural areas, grade-for-age status is an essential and important educational indicator in the Chinese context.

Therefore, grade-for-age is an important measurement of educational achievement, and is widely used in the literature. Kearney and Levine (2019) use grade-for-age to measure "whether a child is enrolled in school - or graduated high school for those aged over 18 - at a

grade appropriate for his or her age". This educational outcome has also been used by Oreopoulos, Page, and Stevens (2006) and Silles (2017). In our study, we use similar measurement. Specifically, we use the birth date and education level to determine whether children are in their expected grade, assuming they started school the first year they were eligible and progressed normally (Qureshi and Gangopadhyaya 2021).[14] For example, if 6-year-old children enter or have entered primary school, or 12-year-old teenagers enter or have entered junior high school, they have achieved the desired level of education, and therefore the grade-for-age value is equal to 1; otherwise, it is 0, which infers delayed enrollment, grade retention or drops out.[15]

Figure 3 illustrates that between 1968 and 1992, children in the treatment group were about 30 percent less likely than urban children to be in the anticipated grade.[16] This gap shows a convergence trend from 1992, with little difference for children born after 1998.[17] The previous section has shown that children born in 1992 are the first cohort affected by the hukou reform, so this figure provides direct evidence on the effect of hukou status on education.[18]

[14] According to the "Compulsory Education Law" of China, children are required to enroll in primary school following their sixth birthday, so they are expected to be enrolled in junior high school and senior high school at ages 12 and 15, respectively.

[15] The age is adjusted based on the school academic year. For example, children born between September 2003 and August 2004 are supposed to enroll in primary school in November 2010 and are defined as the same age.

[16] The *x*-axis indicates an individual's birth year, and the *y*-axis indicates the mean grade-for-age value of all individuals born in that year.

[17] There are two possible reasons why the GFA of children from urban families born after 1992 has also been improved. First, as of the fall of 2008, students in urban areas were eligible for the free compulsory education reform. Individuals born after September 1993 (under age 15) in urban areas were affected by the reform. Second, the government set a limit on the number of repeaters nationwide in 1994. Thus, the GFA of both urban and rural children has been improved due to the change in the retention regulations.

[18] Similarly, we calculate the difference in the probability of grade-for-age between treated children and children whose parents both with rural hukou. However, the educational gap between these two groups does not satisfy the parallel trend before the reform. Since 2000, the government has developed a series of education policies for rural areas, such as the "Two Exemption and One Subsidy" reform, which waives the cost of tuition and textbooks and subsidizes the living expenses of students living in school dormitories, was initially introduced in 2003 for students from extremely poor families. In the spring of 2005, the reform was extended to all primary and secondary school students in poor counties nationwide (Tang, Zhao and Zhao

(Insert figure 3 here)

*Reform exposure*. According to the documents collected from provincial governments, children born on or after July 22, 1998 could inherit their father's hukou without conditions, and preschool children would be given priority for hukou conversion. However, adults were not eligible to apply for hukou conversion. Therefore, children who have a rural hukou mother and an urban hukou father and those born in or after August 1998 are fully exposed to this reform; those born between September 1991 and July 1998 are partially exposed, whilst those born in or before August 1980 are nonexposed. We define individuals born between September 1991 and August 2004 as the *Younger Cohort* and those born between September 1967 and August 1980 as the *Older Cohort*. In addition, the reform would not affect children whose parents both had urban hukou, they are considered nonexposed, and we regard these children as the control group.

*Other control variables*. The census data also collect other individual-level demographic information such as gender, ethnicity, date of birth, and education level. Table 2 reports summary statistics for the treatment and control group samples. First, children in the treatment group are 18.4 percentage points less likely to be in the anticipated grade than their urban counterparts, with a mean of 0.479 and 0.663 for the two groups, respectively. The mean difference of grade-for-age between groups for the older cohort is 0.256, and this value decreases to 0.070 for the younger cohort exposed to the reform (see panel A of table 2). Second, in the treatment group, the mean values of the older cohort and younger cohort with urban hukou are 0.277 and 0.392, respectively, suggesting the reform does provide more

2020). In contrast, the educational policies applicable to children from urban households have changed less. Therefore, children from rural families are not a good control group.

opportunities for the younger cohort to obtain urban hukou (see panel B of table 2). Third, we can observe that compared with urban children, those from the treatment group are older, comprise a higher percentage of males, and a lower proportion living in the urban area (see panel C of table 2).[19] Fourth, the treated children have more siblings, and their parents are generally older and less educated (see panel D of table 2).[20]

Although China has basically achieved its goal of universal compulsory education by 2000, the urban-rural educational gap has remained substantial. To address the urban-rural educational disparity, the government has adopted preferential educational policies toward students in rural areas since 2000. Specifically, the government has introduced a series of education policies targeting rural and poor areas, such as mergers and closures of rural schools, two exemption and one subsidy, and free compulsory education (Hannum, Liu, and Wang 2022; Tang et al. 2020; Xiao, Li, and Zhao 2017). The rising public education spending per student in rural areas at the province level after 2000 may improve the educational outcomes of rural children more than that of urban students. Therefore, we control the government's public spending on education, i.e., *per student education expenditure* (see panel E of table 2).[21]

[19] We construct individuals' residence types based on their residence in the 2010 Chinese Population Census. Individuals were required to report their residence type at the time of the survey in 2010. The binary variable urban *residence* takes the value of 1 if the individual lives in a "neighborhood committee" and 0 if he/she lives in a "village committee."

[20] We also compared the pre-reform differences between children in the treatment group and children in the other two types of families, see table A4. Compared to children in the treatment group, children with rural fathers and rural mothers are less likely to be in the right GFA, have a higher probability of holding rural hukou or living in rural areas, and have more siblings; in contrast, children with rural fathers and urban mothers are more likely to be in the right GFA and have a higher probability of holding urban hukou or living in urban areas.

[21] The data sources are the Educational Statistics Yearbook of China and the China Educational Finance Statistical Yearbook. First, we calculate the average government expenditure on education per student in urban and rural areas separately. Second, we match government spending on education when the individual was in grades 1-9, according to individuals' hukou status, i.e., individuals with urban hukou receive urban per student education spending, and individuals with rural hukou receive rural per student education spending. Finally, the average value of government education expenditure during nine years of compulsory education is calculated for each individual.

(Insert table 2 here)

## 4 Empirical strategies

### 4.1 Difference-in-differences model

In this study, we apply a difference-in-differences (DID) strategy to investigate the causal effect of access to urban hukou on educational attainment. The specific DID model takes the following form:

$$EDU_{ijpym} = \alpha_0 + \alpha_1 Young_i * Treat_j + \alpha_2 Young_i + \alpha_3 Treat_j$$

$$+X_{ijpym}\zeta + \lambda_m + \mu_{pym} + \varepsilon_{ijpym} \quad (1)$$

whereby *i*, *j*, *p*, *y*, and *m* denote individual, family type, province, birth year, and birth month, respectively. $EDU_{ijpym}$ indicates the grade-for-age status of individual *i* from family type *j* born in province *p* in year *y* in month *m*. $Young_i$ is a dummy variable indicating whether the child was aged 6 and below at the time of the hukou reform.[22] $Treat_j$ is a dummy variable indicating a child's household type, which is determined by the hukou status of parents. $Treat_j$ = 1 indicates the treatment group in which the child's father holds an urban hukou and the mother holds a rural hukou; $Treat_j$ = 0 indicates the control group in which a child has parents who both hold urban hukou. $X_{ijpym}$ is a vector of covariates, which include individual and family background characteristics (the child's gender, ethnicity, residence, number of siblings, and the parents' age and education) and contemporaneous government's public spending on education.[23] $\alpha_1$ is the parameter of interest and captures the causal effect of hukou on education. Our primary measure of interest is exposure to the hukou reform, not the

[22] The hukou reform stipulated that children aged six and below had the priority to inherit the hukou status of their father.

[23] Both the father's and mother's years of schooling are included in the model as continuous variables.

actual transmission of the hukou status. In this sense, our approach identifies an effect of intention-to-treat (ITT) and not an effect of treatment-on-the-treated (TOT).

Even if children are born in the same year, whether or not they meet the school-age requirement is related to their birth month. Children born between September and December cannot enter primary school in the year of their sixth birthday (Tang et al. 2020). Therefore, we include $\lambda_m$ as children's birth month fixed effects to capture the systematic difference between those born earlier and those born later in the same birth cohort. In addition, province-birth cohort-birth month fixed effects $\mu_{pym}$ are included to allow different provinces to have different cohort trends as well as different birth month trends in the same birth year.[24] Since provincial governments were responsible for formulating and enacting the hukou reform, and the timing of enacting documents varied across provinces, standard errors are adjusted for clustering at the province-year level (Bertrand, Duflo, and Mullainathan 2004).[25]

### 4.2 Event study

The key to the validity of the DID identification strategy is that the treatment and control groups meet the parallel trend assumption, i.e., individuals in the two groups would have the same cohort trends in the educational outcomes if the reform were not present. In addition, the baseline model estimates the aggregate effect of the reform on children's education, which does

[24] It should be noted that month is a dummy variable indicating whether the individual's birth month is in September or later; if so, it is equal to 1. Because individuals born after September 1st are eligible to enter primary school in the year of their sixth birthday, and because the fall semester generally begins in September, we use September as the cutoff point.

[25] We also report the *p*-values for clustering at the province level in the baseline regression. To avoid the asymptotic inconsistency generated by using a small number of clusters, standard errors are corrected for a small number of clusters via wild bootstrapping with 500 replications (Cameron, Gelbach, and Miller 2008). The wild-bootstrap *p*-values are reported in square brackets.

not reflect the variability of the policy effects across birth cohorts. Therefore, we exploit an event study approach to test the parallel trend hypothesis and the dynamic effects of the policy (Duflo 2001). The following equation gives the generalization of the baseline model:

$$EDU_{ijpym} = \beta_0 + \sum_{l=1967}^{2004} (d_{il} * Young_i * Treat_j)\beta_{1,l} + \beta_2 Young_i$$

$$+\beta_3 Treat_j + X_{ijpym}\gamma + \lambda_m + \mu_{pym} + \varepsilon_{ijpym} \quad (2)$$

whereby $d_{il}$ is a dummy that indicates whether individual $i$ is born in year $l$. $\beta_{1,l}$ is an estimate of the impact of the reform on a given cohort $l$. The definitions of remaining variables are the same as equation (1). We apply individuals born between 1967-2004 to conduct the event study analysis. To this end, we generate a series of dummies based on two-year intervals of birth year.[26] The reference group is the cohorts aged 18-19 when the reform was implemented (i.e., born in 1979-1980), who are completely ineligible for the reform.

In addition, concurrent educational reforms may have differential impacts on the educational outcomes of individuals in different birth cohorts, for example, the Compulsory Education Law (CEL) and China's Higher Education Expansion (CHEE). Both of these educational reforms were implemented at the province level. The effective date of the CELs varies across provinces, from 1984 to 1994 (Cui, Liu, and Zhao 2019). Although the CHEE started in 1999, the expansion rate varies across provinces (Liu and Wan 2019; Wang et al. 2022). Moreover, individuals under the age of 15 or 18 when the CEL or CHEE was effective were affected by these reforms. However, the province-birth cohort-birth month fixed effects $\mu_{pym}$ could absorb the effects of these educational reforms.

[26] Due to the limited number of observations for each birth year, we aggregate birth cohorts into two-year intervals to increase the power of the estimation (Kearney and Levine, 2019).

Another concern about the DID strategy is that other educational policies or socioeconomic environmental changes that occur during the same period may influence educational decisions or educational outcomes. Such events include the one child policy from 1979, the closures and mergers of rural schools which was started in 2001, the free compulsory education reform in rural China that was started in 2006, and the Four Thousand Billion Stimulus Plan from 2008-2009 in response to the financial crisis. Section 6.1 details the robustness checks used to address these concerns.

## 5 Results

### 5.1 Baseline results

Before presenting our formal results from estimating equation (1), we first present the impact of the reform on the hukou status of children in the treatment group. Figure 4 depicts the estimates of timing dummies with a 95% confidence interval in the regression of children's hukou status (urban=1) on time dummy variables (28 years pre-reform and 12 years post-reform), personal and family background characteristics, month dummies, and province dummies.[27] The reference group comprises the cohorts born 30 years before 1998. The estimated coefficients (0.223) are significantly different from zero, starting with the 6-year-old cohort in 1998, followed by a significant jump for those born two years after 1998. Therefore, newborns and preschoolers at the time of the reform were more likely to be policy compliers and generally inherited urban hukou from their fathers.

(Insert figure 4 here)

[27] On the x-axis, positive numbers indicate the age in 1998, and negative numbers denote how many years after the reform the person was born.

Then, we investigate how the reform affects the hukou status using a similar framework as equation (1), and using hukou status (urban=1) as the dependent variable. Results are reported in table 3. In column 1, we only control for individual characteristics of the child, including gender, ethnicity and residence. In column 2, we add family background characteristics such as the number of siblings, age of parents, and educational level of parents. In column 3, we further control the *per student education expenditure*. The birth month and province-birth cohort-birth month fixed effects are included in all regressions. The reform resulted in a 15.1 percentage point higher probability of obtaining an urban household for the younger cohort relative to the older cohort in the treatment group than in the control group.

(Insert table 3 here)

Table 4 reports the baseline results of the DID estimation. The coefficient on *Young* × *Treat* can be interpreted as the causal effect of exposure to the reform on children's hukou and grade-for-age status. Column 3 demonstrates that children exposed to the reform are 18.9 percentage points more likely to be at the grade level appropriate for their age. To interpret the magnitude of the coefficient, we need to be aware that 39.2 percent of the younger cohort in the treatment group obtained an urban hukou and 74.4 percent of them achieved the anticipated grade.[28] Therefore, the conversion of hukou type from 0 to 1 would increase the grade-for-age status by 48.2 percentage points (0.189 / 0.392 = 0.482), which reflects a 64.8 percent (0.482 / 0.744 = 0.648) improvement in the probability of age-appropriate enrollment.[29]

[28] Not all eligible families choose urban hukou for their children. Descriptive statistics for different birth cohorts in the treatment group with urban or rural hukou status show that parents of children with urban hukou are usually younger, have more years of schooling, and have fewer siblings (see table R2, this table is for review only, does not intend for publication).

[29] We follow Kearney and Levine (2019) to construct this measure, who studied the impact of early childhood television programs on educational outcomes, including grade-for-age status.

Note that the treatment and control groups are defined based on parents' hukou status in 2010 rather than the hukou status when the reform was implemented. If a rural woman who married an urban man obtained an urban hukou in cities, her children would be considered as the control group. In this case, our estimates can be regarded as the lower bound of the effect of hukou reform on children's grade-for-age. However, migration for marriage is not very common in China. For instance, the 2010 Chinese Population Census indicates that only 4.42% of rural women left their place of household registration for marriage. Moreover, the 2010 Chinese General Social Survey (CGSS) suggests that only 2.29% of rural females obtained an urban hukou by marriage in adulthood. Similarly, the China Health and Retirement Longitudinal Study (CHARLS) in 2014 shows that approximately 1.6% of rural women obtained an urban hukou following their spouses' hukou type. Therefore, the estimation bias due to changes in the mother's hukou status tends to be small in our study.

(Insert table 4 here)

To explore the educational stage in which the hukou plays the most significant role, we construct three dummy variables indicating whether individuals enter or have entered primary, junior, and senior high school at the expected age. Results are shown in columns 4-6 of table 4.[30] It is noteworthy that the impact of the reform on the educational outcome is most significant in senior high school, where grade-for-age status is estimated to increase by 26.2 percentage points. In contrast, the reform has a much smaller impact on the junior high school level with a value of just 1.6 percentage points. The estimated coefficients are not statistically significant at the primary and junior high school levels. Since 1986, the universalization of

[30] The expected ages of entry into primary, junior high, and senior high school are 6, 12, and 15 years old, respectively.

compulsory education in China has alleviated the disparity of educational opportunities at the primary and junior high school levels. However, most high-quality senior high schools are located in urban areas, and many rural children must move to urban areas to pursue post-compulsory education. Therefore, it is reasonable that the reform of the hukou system has mainly promoted educational equity at the high school level.[31]

If it holds that obtaining an urban hukou improves educational outcomes, we can expect the positive impact of the reform on education to be greater in regions with significant urban-rural educational disparities. To explore the effects of the reform in areas with different urban-rural education gaps, we calculate the ratio of years of schooling for urban and rural adults over 25 years old, this age group is expected to have completed their education, at the prefecture level. Further, we define prefectures equal to and above the median as having a *Large* gap and those below the median as having a *Small* gap. As shown in columns 1 and 2 of table 5, the reform increases the probability of a child being grade-for-age by 20.2 and 16.0 percentage points in prefectures with large and small education gaps, respectively. In particular, the effect that achieving hukou status conversion has on junior high school enrollment is only significant in the group with a sizeable urban-rural gap (see column 5 of table 5). It is reasonable that the more significant the education gap between urban and rural areas within the same prefecture, the higher the added value of education within the local urban hukou.

(Insert table 5 here)

[31] We also combine matching and DID methods to investigate the effect (Heckman, Ichimura, and Todd 1997). As shown in table A5 (the matching process for review only, do not intend for publication), the matched DID estimates are similar to the main estimation results.

## 5.2 Validity of DID strategy and dynamic effect

### *5.2.1 Parallel cohort trends and dynamic effect*

The key hypothesis of the identification strategy is that the educational outcomes of the two groups of children would have met parallel trends if there had been no hukou reform. Figure 5 shows no significant trend before the event except a drop in 1984, which provides supporting evidence for the parallel-trend assumption.[32] Figure 5 also illustrates a dynamic effect of the hukou reform, i.e., the later the child is born, the greater the policy exposure and so the more prominent the improvements in education. For one thing, the rate of grade-for-age shows a significant increase in 1992, and an upward trend over time, because children born in 1992-1997 are partially exposed, and the later the birth, the higher the probability of obtaining an urban hukou. For another, the significant increase in the estimates after 1998 suggests a more remarkable educational improvement for those birth cohorts fully exposed to the hukou reform.

(Insert figure 5 here)

### *5.2.2 Placebo tests*

In order to ascertain that progress in educational outcomes for treated children is indeed due to the hukou reform rather than other structural changes, we perform two placebo tests.

Firstly, we assume that the hukou reform was implemented 8-13 years earlier than the actual reform year.[33] Based on the pseudo-policy timing, we construct a fake birth cohort

[32] We do not have a robust explanation for the drop in 1984, which is an open question for future investigation.

[33] We have chosen this setting as delayed enrollment is quite common in rural China. For example, children who were already 7 years old when the reform started may not have entered elementary school, and thus they were still preschool-age children affected by the policy.

exposed to the policy. For example, when we assume that the reform occurred in 1990, those exposed to the policy were born from September 1983 to August 1996 (*Placebo Young* = 1), and the corresponding nonexposed cohorts were born from September 1959 to August 1972 (*Placebo Young* = 0). We then re-estimate the DID model based on the interaction of *Placebo Young* and *Treat* as in equation (1), and all other model specifications remain unchanged. Figure 6 demonstrates that the estimated coefficients of all interaction terms are not statistically different from 0.

(Insert figure 6 here)

Secondly, the placebo treatment group is defined as children from the same cohorts whose fathers have rural hukou and mothers have urban hukou. When the father has rural hukou, irrespective of whether the 1998 hukou reform grants the child the freedom to settle in the father's place of hukou registration, the parents generally have no desire for their children to obtain rural hukou. In short, the reform does not have an exogenous shock on the hukou status of children with rural fathers. Therefore, we construct the "*Placebo treat*" dummy using children with rural fathers and urban mothers as the treated samples and re-estimate equation (1). Figure 7 shows the estimation results are almost distributed on both sides of the line at 0, as consistent with our expectations.

(Insert figure 7 here)

## 6 Robustness checks, heterogenous effects, and mechanisms

### 6.1 Controlling for confounding factors

To isolate the policy effect and control for other co-founding factors and shocks, we re-estimate the effect of the hukou reform, considering other contemporaneous policies and socioeconomic factors.

#### *6.1.1 One child policy*

Since the 1980s, the one child policy has been strictly enforced, especially in urban China. For example, the penalty rate increased rapidly from 0.8% to 3.0% of annual household income on average between 1989 and 1992 (Huang, Lei, and Sun 2021). As suggested by the quantity-quality trade-off theory, the one child policy may improve the educational attainment of individuals, and exposure to stricter fertility restrictions is associated with higher education levels (Qin, Zhuang, and Yang 2017; Huang et al. 2021). If the enforcement coincides with the timing of the hukou reform, the omission of the one child policy may bias our results. As the enforcement of the one child policy varies across counties and years, we additionally control for county-birth cohort fixed effects to capture the impacts of the family planning policy. As shown in panel A of table 6, after controlling for county-birth cohort fixed effects, the coefficients of interest are similar to the baseline results.

#### *6.1.2 Mergers and closures of rural schools*

Due to the decline in rural school-age children, the State Council issued the *Decisions on the Reform and Development of Primary Education* in 2001 to adjust the layout of schools in

rural areas.[34] From 2000 to 2010, the number of rural primary schools saw a 59.38% decrease from 416,198 to 169,045; the number of regular junior high schools decreased from 35,023 to 15,112, a reduction of 56.85%, which led to an increase in dropout rates (Educational Statistics Yearbook of China 2001, 2011). As the school closure process almost coincided with the hukou reform in 1998, the validity of identification could be challenged. We use the change in the number of schools per student from 2000 to 2010 to measure the intensity of primary and junior high school closures, and include the interactions between the intensity of school closure at the prefecture level and birth cohort dummies in the model.[35] After controlling for the accessibility of schools, the results reported in panel B of table 6 are consistent with the baseline.

### *6.1.3 Free compulsory education reform*

In order to promote educational equity, the Chinese government began to implement Free Compulsory Education (FCE) reform in rural areas in 2006. The policy specified that tuition and miscellaneous fees for primary and secondary school students should be exempted. The reform was to be implemented in western rural areas in 2006, central and eastern rural areas in 2007, and urban areas in 2008 (Xiao et al. 2017).[36] Individuals born after September 1991 were the main beneficiaries of the 1998 hukou reform. They were under 15 years old in 2006,

[34] With the relaxation of immigration controls in the country, some farmers would bring their children with them when they moved to the cities to seek work. Simultaneously, the one child policy reduced the number of children in rural families.

[35] The number of primary and junior high schools and the number of students at the city level in 2000 and 2010 are obtained from the China City Statistical Yearbook in 2001 and 2011. As a robustness test, we calculate the proportion of school reductions from 2000 to 2010 to measure the policy strength, and results remain unchanged.

[36] The specific timing of introducing the FCE reform in each province is as follows. In rural areas, Tianjin, Inner Mongolia, Shanghai, Fujian, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang began implementation in March 2006; Beijing, Jiangsu, Zhejiang, and Guangdong in September 2006; Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Anhui, Jiangxi, Shandong, Henan, Hubei, and Hunan in March 2007. Urban areas in all provinces began implementation in September 2008.

attending primary or junior high school, and were exposed to the FCE reform. Therefore, the previous estimation results may be confounded by the effect of the FCE reform. To allay this concern, we control the number of semesters that an individual is supposed to be exposed to the FCE reform and its interaction term with place of residence. Our baseline results are robust to considering the school subsidy program, as shown in panel C of table 6.

### *6.1.4 Four Thousand Billion Stimulus Plan*

After the global financial crisis in 2008, the Chinese government launched the Four Thousand Billion Stimulus Plan in November of that year to boost economic growth and maintain stable employment. The funds were mainly spent on housing, transportation, and rural infrastructure construction, which provided many jobs for migrant workers in related industries. Education is a rational decision considering budget constraint, costs, expected benefits, and the perceived probabilities of successful outcomes. The stimulus package increases the opportunity cost of education for children and thus may increase the dropout rate in rural areas; relaxes the household budget constraints and may improve the probability of enrollment. To control for the impact of socioeconomic factors on individual education decisions, we exclude individuals in the treatment group who were aged 16 years or above in 2008 and re-estimate equation (1).[37] The regression results based on the subsample suggest that the effect of the hukou reform is not confounded by the economic crisis in 2008 (see panel D of table 6).

(Insert table 6 here)

[37] We use 16 as a cut-off age since according to the law, state organs, social bodies, enterprises, institutions, non-governmental not-for-profit organizations, and private businesses are prohibited from employing children under 16 years of age (Tang et al. 2020).

### 6.2 Other robustness checks

This subsection carries out five robustness tests. Firstly, we adjust the age of entry into the primary school to 7 years old because the delayed entry may underestimate the probability of grade-for-age status (Cui et al. 2019). The results in panel A of table 7 show that the magnitude and significance of the coefficients are generally consistent with the main results, suggesting that an age-specific setting does not drive our findings.

Secondly, individuals who benefit from the hukou reform may have spillover effects on their siblings. Therefore, we create a new variable, *Siblings' exposure*, i.e., the total number of siblings who are supposed to be exposed to the reform, and then control it in the regression. After considering the spillover effect, our key results remain robust, as shown in panel B of table 7.

Thirdly, China's hukou system has been gradually deregulated since the 1980s, meaning that migrant workers could bring their children to urban areas and enroll them in local schools (Zhang 2017). To exclude this interference of migration, we remove individuals that leave their original place of hukou registration; our findings do not change after re-estimation (see panel C of table 7).

Fourthly, we change the range of the older cohort from 18-30 to 7-17 years old. As panel D of table 7 shows, the reform's effects reduce at junior and senior high school levels because a proportion of the cohort also achieves a hukou conversion. However, the findings are broadly consistent with the baseline regression.

Finally, the reform policies in some large cities were implemented late, and some local governments refused to address the issue of urban hukou for those who violated the one child

policy. Therefore, we remove samples from 15 sub-provincial cities and individuals with siblings. The coefficient estimation results are consistent with the baseline regression after making stricter restrictions on the sample (see panel E of table 7).

(Insert table 7 here)

### 6.3 Heterogeneous effects

This section explores the heterogeneous effects of the 1998 hukou reform on children's educational outcomes. Specifically, we analyze from three perspectives: gender of children, parental education level, and city type.

#### *6.3.1 Gender of children*

There is a strong son preference in rural China and resource allocation within a family is often biased in favor of boys. Moreover, girls' education is more sensitive to direct costs, opportunity costs, and household economic constraints (Hannum 2003). After obtaining an urban hukou, rural children do not need to pay extra fees to attend urban schools, thus reducing education costs. As a consequence, there may be gender differences in term of hukou reform on promoting educational equity. The results of separate regressions using subsamples of boys and girls are listed in panel A of table 8. The effect on educational outcome is particularly pronounced amongst girls, with an 80.9 percent reduction in the probability of falling behind in school, compared with the 57.4 percent for boys.[38] In addition, girls are 9.5 percentage points higher than boys in terms of senior high school education attainment. These results suggest that the hukou reform can help reduce the gender gap in education.

[38] The economic meaning of the coefficients is calculated in the same way as in the baseline regression, except that the mean of the grade-for-age and hukou variables are grouped by gender when they are counted.

### *6.3.2 Parental education level*

Parents with higher levels of education have a stronger financial capacity and are more efficient at investing in their child's human capital because they have the knowledge to transfer education to their children (Becker et al. 2018). Furthermore, knowledgeable adults have higher educational expectations of their children (Spagat 2006). Therefore, the effect of the hukou reform may differ depending on parents' cultural capital. A family is defined as *Educated* if either parent has attended senior high school or above, and *Less-educated* otherwise. As shown in panel B of table 8, the reform increases the grade-for-age status by 32 percentage points for children with educated parents; however, the figure is only 6.5 percentage points amongst children of whom neither parent have a high school education. The hukou reform effect is complementary to the parental effect, which means educated parents can better seize the educational opportunities that the reform brings to their children.

### *6.3.3 City type*

Inequality in educational opportunities in China exists between urban and rural areas; differences also exist between cities of different hierarchies because large cities are better funded and have more qualified teachers (Tani, Xu, and Zhu 2021).[39] Since there are significant differences in educational opportunities and quality between different city types, we conjecture that the reform's effect is heterogeneous across cities. We define the *Capital* dummy variable that distinguishes provincial capitals from smaller cities and towns. As shown in panel C of table 8, children in capital cities show a 5.3 percentage points higher improvement level

[39] Migrant students who move to larger cities have 10 percentage points higher cognitive gains than their peers who move to smaller cities.

in grade-for-age status than their counterparts in smaller cities. The increase in the likelihood of being at the suitable grade level is 70.1 and 56.8 percentage points for children in capital cities and non-capital cities, respectively.[40] In short, obtaining an urban hukou can improve educational outcomes in either large or small cities, with larger impacts for an urban hukou in capital cities.

(Insert table 8 here)

### 6.4 Mechanisms

As noted in Section 2.1, holding the urban hukou typically implies a better quality of education, lower cost of education, and higher educational expectations. We estimate the impacts of the hukou reform on three channel variables in a DID framework:

$$Y_{ijpyw} = \gamma_0 + \gamma_1 Young_i * Treat_j + \gamma_2 Young_i + \gamma_3 Treat_j$$
$$+X_{ijpyw}\zeta + \delta_p + \eta_y + \tau_w + \varepsilon_{ijpyw} \quad (3)$$

whereby $i$, $j$, $p$, $y$, and $w$ denote individual, family type, province, birth year, and survey wave, respectively. $Y_{ijpyw}$ indicates potential channel variables, including quality of education, cost of education and educational aspirations. The definition of $Young_i$ and $Treat_j$ are the same as those used in equation (1). $X_{ijpyw}$ refers to the individual's gender, ethnicity, and the parents' age and years of schooling. $\delta_p$, $\eta_y$, and $\tau_w$ denote a full set of the province, birth year, and the survey wave fixed effects, respectively.

To investigate the potential channels through which hukou status affects children's education, we apply the data from the China Family Panel Studies (CFPS) and the Chinese

[40] The economic meaning of the coefficients is calculated in the same way as in the baseline regression, except that the mean of the grade-for-age and hukou variables are grouped by city type when they are counted.

General Social Survey (CGSS), both of which are nationally representative. However, we could not rely on a single dataset for the following two reasons. First, the CFPS collects the current quality of education and education cost of teenagers; however, the CFPS does not contain the quality of education and education cost during compulsory education of an adult. Second, although the CGSS collects information on the quality of education and education cost at each stage of education of respondents, all respondents are adults. As a result, we pooled the older cohorts (aged 30-43 years in 2010) derived from the CGSS and the younger cohorts (aged 6-19 years in 2010) derived from the CFPS.[41]

### 6.4.1 Quality of education

First, urban children have increased access to better schools in cities and thus have better educational outcomes. Given that the 9-year compulsory education is primarily financed by the local government in China, the quality of education depends on the fiscal capacity of the local government.[42] Hence, we apply the "school location" as a proxy for the quality of education. The *school location* equals 1 for cities and 0 for counties and villages. We pool the young cohorts from the CFPS 2010-2020, which investigates the location of schools for students under age 16, and the old cohorts from the 2003 and 2008 waves of the CGSS, which contain the detailed retrospective educational history of respondents (e.g., the location of schools at each stage of education), and estimate equation (3). The result in column 1 of table 9 indicates

[41] Before examining the mechanisms, we first test whether the samples from the two surveys are nationally representative. To this end, we compare the demographic characteristics of samples from the two surveys with that of the samples from Census 2010 data. The results in table A6 demonstrate that the demographic characteristics, hukou status, and educational attainment of samples from CFPS 2010 and CGSS 2010 are comparable with those of the Census 2010.

[42] In cities, primary schools are administered and funded by the district government, and secondary schools are administered and funded by the city governments. In rural areas, primary schools are financed primarily from community contributions, and secondary schools are administered and funded by the county governments.

that the hukou reform significantly increased the probability of enrolling in urban schools for the young cohorts in the treatment group by 20.8 percentage points. The hukou reform increases access to better schools.

### 6.4.2 Cost of education

Second, urban hukou allows students to attend local urban schools without paying sponsorship fees and thus reduces education costs, which tends to decrease school dropout. To investigate this channel, we use the "extra-paid educational expenses" to measure the cost of education. We pool the young cohorts from the CFPS 2010-2020 and the old cohorts from the CGSS 2003.[43] Since the dependent variable has many zeros (only 7.27% of the sample had ever paid sponsorship fees to enroll in a school), we apply a Tobit model here. As shown in column 2 of table 9, the reform reduces the extra educational expenses for the young cohorts in the treatment group by 3,587 CNY, although not statistically significant. Furthermore, we use a decomposition method developed by McDonald and Moffitt (1980) and decompose the Tobit coefficient into two parts: (1) the effect of the hukou reform on the amount of sponsorship fees for those who paid sponsorship fees (intensive margin); and (2) the effect of the hukou reform on the probability of paying positive sponsorship fees (extensive margin). In our study, 84.77% is attributable to the extensive margin, i.e., the waiver of sponsorship fees, and only 15.33% is attributable to the reduction in educational costs.[44]

[43] The CGSS contains the information on extra-paid educational expenses at each educational stage only in 2003.

[44] In our study, 7.27% of the observations paid non-zero sponsorship fees on choosing schools, i.e., the fraction of the sample above zero $F(z)=0.0727$. Thus, the fraction of mean total response resulting from marginal changes in the value of (positive) expenditures $1-zf(z)/F(z)-f(z)^2/F(z)^2 = 0.1533$, whereas 0.8467 would be generated by changes in the probability of spending anything at all.

### 6.4.3 Educational aspirations

Third, since the returns to education are higher in urban than rural areas and individuals in urban areas have more information on returns to education, the hukou reform can promote educational attainment by raising educational aspirations. We measure educational aspiration by whether the highest educational attainment a student wants to achieve is a three-year college and above. The *educational aspiration* takes the value of 1 if an individual desires college or above and 0 otherwise. Both the young and old cohorts are from the CFPS 2010-2020, as the survey has detailed records of educational aspirations for all age groups of respondents. The estimate in column 3 of table 9 demonstrates that the hukou reforms raised the probability that an individual expects to obtain college and above education by 13.4 percentage points.

Overall, the empirical evidence supports for these three mechanisms. Specifically, the hukou reform has led to a significantly higher probability of attending urban schools and a significantly higher educational aspiration for the young cohorts in the treatment group. Moreover, the hukou reform has reduced sponsorship fees for the treated cohorts, although the effect is not precisely estimated.

(Insert table 9 here)

## 7 Conclusions

This study identifies and estimates the causal link between hukou status and educational attainment using the change in the hukou inheritance law in 1998, which permits children with urban fathers and rural mothers to change from rural to urban hukou. We find that the reform increases the probability of children's grade-for-age status throughout their educational career, especially at the senior high school level; this also suggests a decrease in ending schooling after

the primary or junior high school level, as the course of education is continuous. Furthermore, we explore three possible mechanisms and find that educational quality and educational aspirations are essential in improving educational attainment.

In addition, we find heterogeneity in the policy effects of hukou reform on promoting educational equity. First, urban hukou facilitates educational attainment for girls more than boys, and thus hukou reform is beneficial in reducing the gender gap in education. Second, children with educated parents benefit more from the reform, reflecting the lack of intergenerational mobility in education and the trend toward class solidification. Third, the educational value-added of urban hukou in capital cities is larger than that in small cities, which provides a new perspective on educational inequality in China.

Our findings provide important policy implications for promoting rural-urban educational equity and enhancing human capital accumulation in China. First and foremost, it is essential to remove the linkage between children's access to education and their hukou identities, with a focus on increasing government funding for education in rural areas, which will help to eliminate the pre-market discrimination caused by the hukou system (Song 2014). Secondly, children from disadvantaged families encounter severe selection barriers, so it is important for policymakers provide them with more information and educational assistance (Mare 1980); otherwise, the reform will only benefit more for the higher socioeconomic status group, which will reduce the intergenerational mobility of education. Finally, the educational inequities between cities of different tiers must be acknowledged. As rapid urbanization in China continues, unless the concentration of high-human capital extends beyond the capital cities (Spagat 2006), it will become a new source of social stratification in the future.

## References


Andrew, Megan. 2014. "The Scarring Effects of Primary-Grade Retention? A Study of Cumulative Advantage in the Educational Career." *Social Forces* 93, no. 2 (December): 653-85.

Becker, Gary S. 1993. *Human capital: a theoretical and empirical analysis with special reference to education.* Third ed. Chicago: The University of Chicago Press.

Becker, Gary S., Kevin M. Murphy, Scott Duke Kominers, and Jörg L. Spenkuch. 2018. "A Theory of Intergenerational Mobility." *Journal of Political Economy* 126: 7-25.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119, no. 1: 249-75.

Brown, Philip H., and Albert Park. 2002. "Education and Poverty in Rural China." *Economics of Education Review* 21, no. 6: 523-41.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90, no. 3 (August): 414-27.

Chen, Qihui. 2015. "Ready for School? Impacts of Delayed Primary School Enrollment on Children's Educational Outcomes in Rural China." *International Journal of Educational Development* 45 (November): 112-28.

Cui, Ying, Hong Liu, and Liqiu Zhao. 2019. "Mother's Education and Child Development: Evidence from the Compulsory School Reform in China." *Journal of Comparative Economics* 47, no. 3 (September): 669-92.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *The American Economic Review* 91, no. 4 (September): 795-813.

Guo, Yu, and Liqiu Zhao. 2019. "The Impact of Chinese Hukou Reforms on Migrant Students' Cognitive and Non-Cognitive Outcomes." *Children and Youth Services Review* 101 (June): 341-51.

Han, Li, Tao Li, Yaohui Zhao. 2015. "How status inheritance rules affect marital sorting: theory and evidence from urban China." *Economic Journal* 125, no. 589: 1850-87.

Hannum, Emily. 2003. "Poverty and Basic Education in Rural China: Villages, Households, and Girls' and Boys' Enrollment." *Comparative Education Review* 47, no. 2 (May): 141-59.

Hannum, Emily, Xiaoying Liu, and Fan Wang. 2022. "Estimating the Effects of Educational System Consolidation: The Case of China's Rural School Closure Initiative." *Economic Development and Cultural Change* 70, no. 1 (October): 485-528.

Heckman, James J. 2003. "China's Investment in Human Capital." *Economic Development and Cultural Change* 51, no. 4: 795-804.

Heckman, James J. 2005. "China's Human Capital Investment." *China Economic Review* 16, no. 1 (January): 50-70.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* 64, no. 4 (October): 605-54.

Huang, Guoying, Yu Xie, and Hongwei Xu. 2015. "Cognitive Ability: Social Correlates and Consequences in Contemporary China." *Chinese Sociological Review* 47, no. 4: 287-313.

Huang, Wei, Xiaoyan Lei, and Ang Sun. 2021. "Fertility Restrictions and Life Cycle Outcomes: Evidence from the One-Child Policy in China." *Review of Economics and Statistics* 103, no. 4: 694-710.

Hui, Eddie C., Ka Hung Yu, and Yinchuan Ye. 2014. "Housing Preferences of Temporary Migrants in Urban China in the Wake of Gradual Hukou Reform: A Case Study of Shenzhen." *International Journal of Urban and Regional Research* 38, no. 4: 1384-98.

Huo, Xuan, and Mingang Lin. 2022. "Evolution of the Rural Social Security System in a Large Country over 35 Years: Institutional Transformation and the Chinese Experience." *China Agricultural Economic Review* 14, no. 1 (January): 1-16.

Kearney, Melissa S., and Phillip B. Levine. 2019. "Early Childhood Education by Television: Lessons from Sesame Street." *American Economic Journal: Applied Economics* 11, no. 1 (January): 318-50.

Liu, Ling, and Qian Wan. 2019. "The Effect of Education Expansion on Intergenerational Transmission of Education: Evidence from China." *China Economic Review* 57 (October).

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75, no. 370 (June): 295-305.

McDonald, John F., and Robert A. Moffitt. 1980. "The Uses of Tobit Analysis." *Review of Economics and Statistics* 62, no. 2: 318-21.

Meng, Xin, and Sen Xue. 2020. "Social Networks and Mental Health Outcomes: Chinese Rural-Urban Migrant Experience." *Journal of Population Economics* 33, no. 1: 155-95.

Meng, Xin. 2012. "Labor Market Outcomes and Reforms in China." *Journal of Economic Perspectives* 26, no. 4: 75-102.

Ministry of Education of the People's Republic of China. 1999-2019. *Educational Statistics Yearbook of China*. Beijing: People's Republic of China People's Education Press.

Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens. 2006. "The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics* 24, no. 4: 729-60.

Pan, Yao. 2017. "The Impact of Removing Selective Migration Restrictions on Education: Evidence from China." *Journal of Human Resources* 52, no. 3 (June): 859-85.

Qin, Xuezheng, Castiel Chen Zhuang, and Rudai Yang. 2017. "Does the One-Child Policy Improve Children's Human Capital in Urban China? A Regression Discontinuity Design." *Journal of Comparative Economics* 45, no. 2 (May): 287-303.

Qu, Zhaopeng, and Zhong Zhao. 2017. "Glass Ceiling Effect in Urban China: Wage Inequality of Rural-Urban Migrants During 2002-2007." *China Economic Review* 42 (February): 118-44.

Qureshi, Javaeria A., and Anuj Gangopadhyaya. 2021. "Childhood Medicaid Eligibility and Human Capital." *Economics of Education Review* 82 (June): 1-12.

Silles, Mary A. 2017. "The Intergenerational Transmission of Education: New Evidence from Adoptions in the USA." *Economica* 84, no. 336: 748-78.

Song, Yang. 2014. "What Should Economists Know About the Current Chinese Hukou System?" *China Economic Review* 29 (June): 200-12.

Spagat, Michael. 2006. "Human Capital and the Future of Transition Economies." *Journal of Comparative Economics* 34, no. 1 (January): 44-56.

Tang, Can, Liqiu Zhao, and Zhong Zhao. 2020. “Does Free Education Help Combat Child Labor? The Effect of a Free Compulsory Education Reform in Rural China.” *Journal of Population Economics* 33, no. 2: 601-31.

Tani, Massimiliano, Lei Xu, and Yu Zhu. 2021. “The Impact of an Un(Der)Funded Inclusive Education Policy: Evidence from the 2013 China Education Panel Survey.” *Journal of Economic Behavior and Organization* 190 (October): 768-84.

Wang, Chuhong, Xingfei Liu, Zizhong Yan, and Yi Zhao. 2022. “Higher Education Expansion and Crime: New Evidence from China.” China Economic Review 74 (August).

Wu, Xiaogang, and Donald J. Treiman. 2004. “The Household Registration System and Social Stratification in China: 1955-1996.” *Demography* 41, no. 2 (May): 363-84.

Wu, Xiaogang. 2010. “Economic Transition, School Expansion and Educational Inequality in China, 1990-2000.” *Research in Social Stratification and Mobility* 28, no. 1 (January): 91-108.

——. 2012. “The Household Registration System and Rural-Urban Educational Inequality in Contemporary China.” *Chinese Sociological Review* 44 (February): 31-51.

Xiao, Yun, Li Li, and Liqiu Zhao. 2017. “Education on the Cheap: The Long-Run Effects of a Free Compulsory Education Reform in Rural China.” *Journal of Comparative Economics* 45, no. 3 (August): 544-62.

Ye, Hua. 2015. “Key-Point Schools and Entry into Tertiary Education in China.” *Chinese Sociological Review* 47, no. 2 (January): 128-53.

Zhang, Huafeng. 2017. “Opportunity or New Poverty Trap: Rural-Urban Education Disparity and Internal Migration in China.” *China Economic Review* 44 (July): 112-24.

## Figures and tables

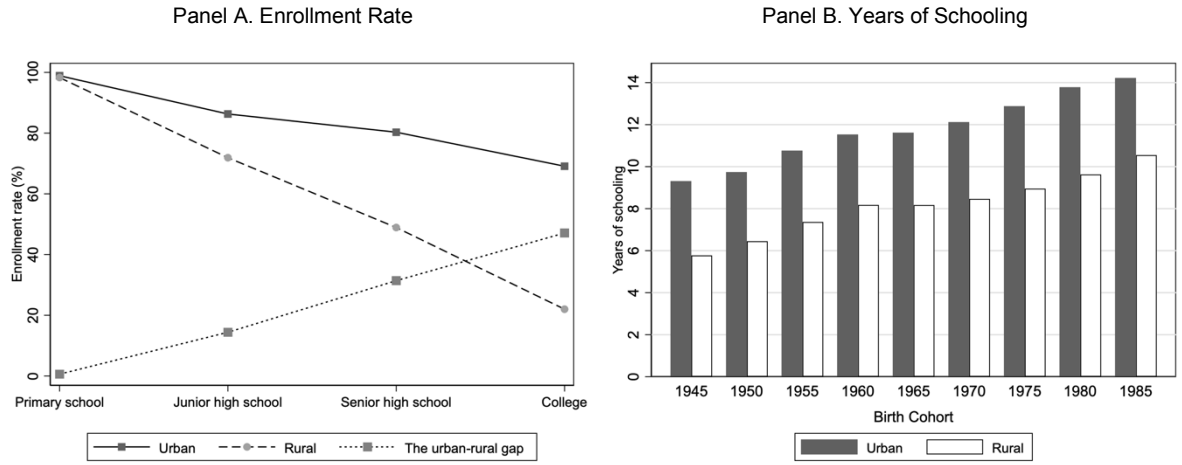


**Figure 1.** The urban-rural education gap in China. Panel A illustrates the urban-rural gap in enrollment rates. The enrollment rate is the proportion of children enrolled in school among the appropriate school-age population. The appropriate school age is 6-11, 12-14, 15-17, and 18-21 for primary, junior high, senior high, and college students, respectively. Panel B presents the urban-rural difference in the average years of schooling for the working-age population (aged 16-64) born in different cohorts. For example, the label 1945 represents people born between 1945 and 1949; the label 1950 represents people born between 1950 and 1954. Source: The 0.095% sample of the 2010 Chinese Population Census.

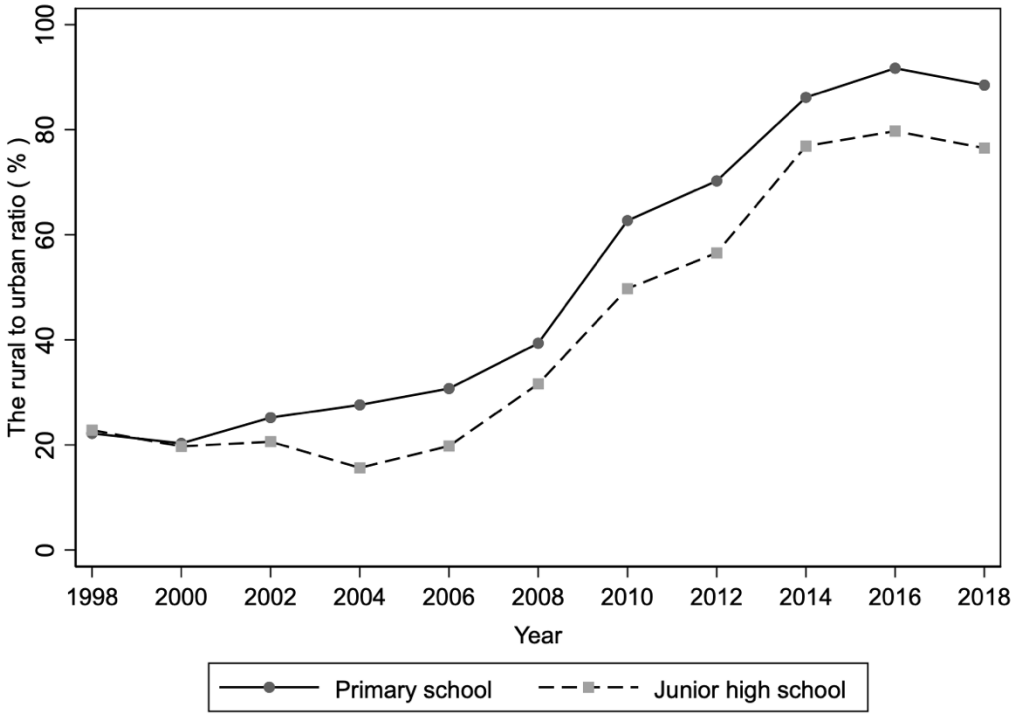


**Figure 2.** The urban-rural gap in education expenditure per student in China, 1998-2018. The rural-urban ratio is the relative value of rural education expenditure per student divided by urban education expenditure per student. Source: Educational Statistics Yearbook of China (1999 - 2019).

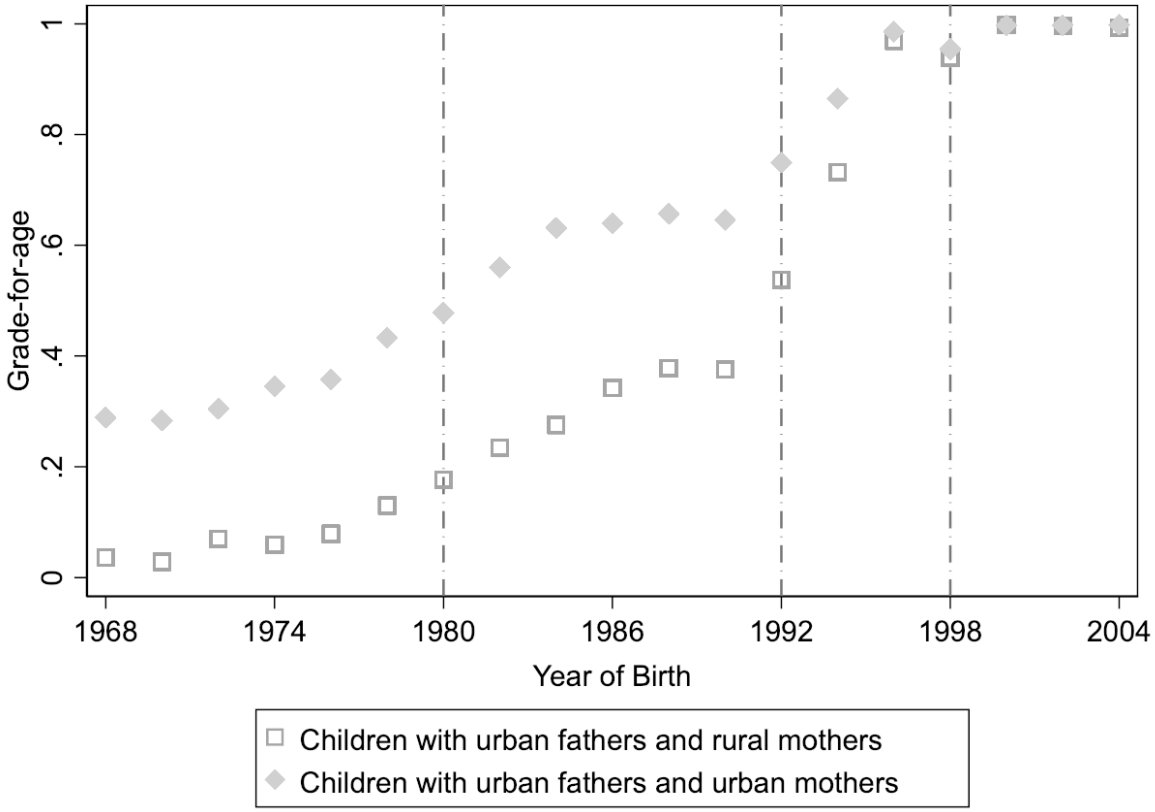


**Figure 3.** The educational outcomes of children from different family types. The x-axis indicates an individual's birth year, and the y-axis indicates the mean grade-for-age status of all individuals born in that year. Source: The 0.095% sample of the 2010 Chinese Population Census.

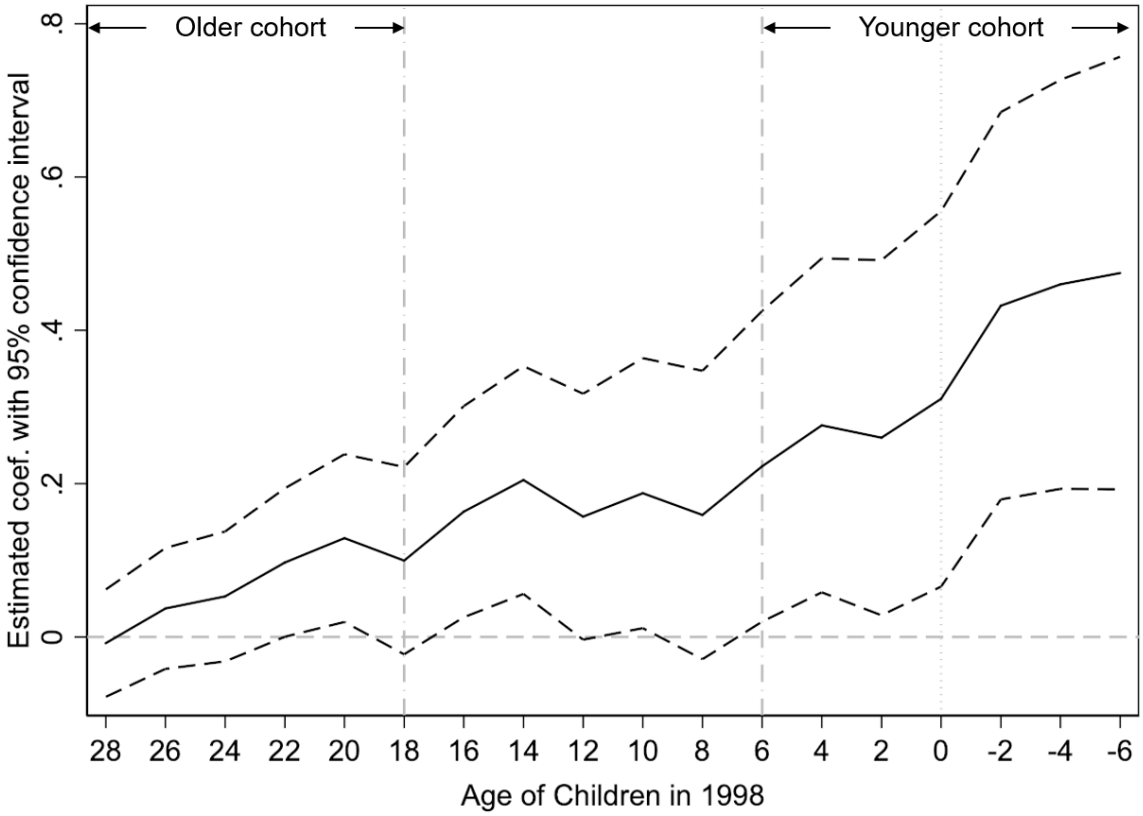


**Figure 4.** The effect of the 1998 reform on children's hukou status. The sample comprises individuals from households with urban fathers and rural mothers. The figure reports the estimates of timing dummies with a 95% confidence interval in the regression of children's hukou status (urban=1) on time dummy variables (28 years pre-reform and 12 years post-reform), personal and family background characteristics, and province dummies. Birth cohorts are aggregated into two-year intervals to increase the power of the analysis. On the x-axis, positive numbers indicate the age in 1998, and negative numbers denote how many years after the reform the person was born. Source: The 0.095% sample of the 2010 Chinese Population Census.

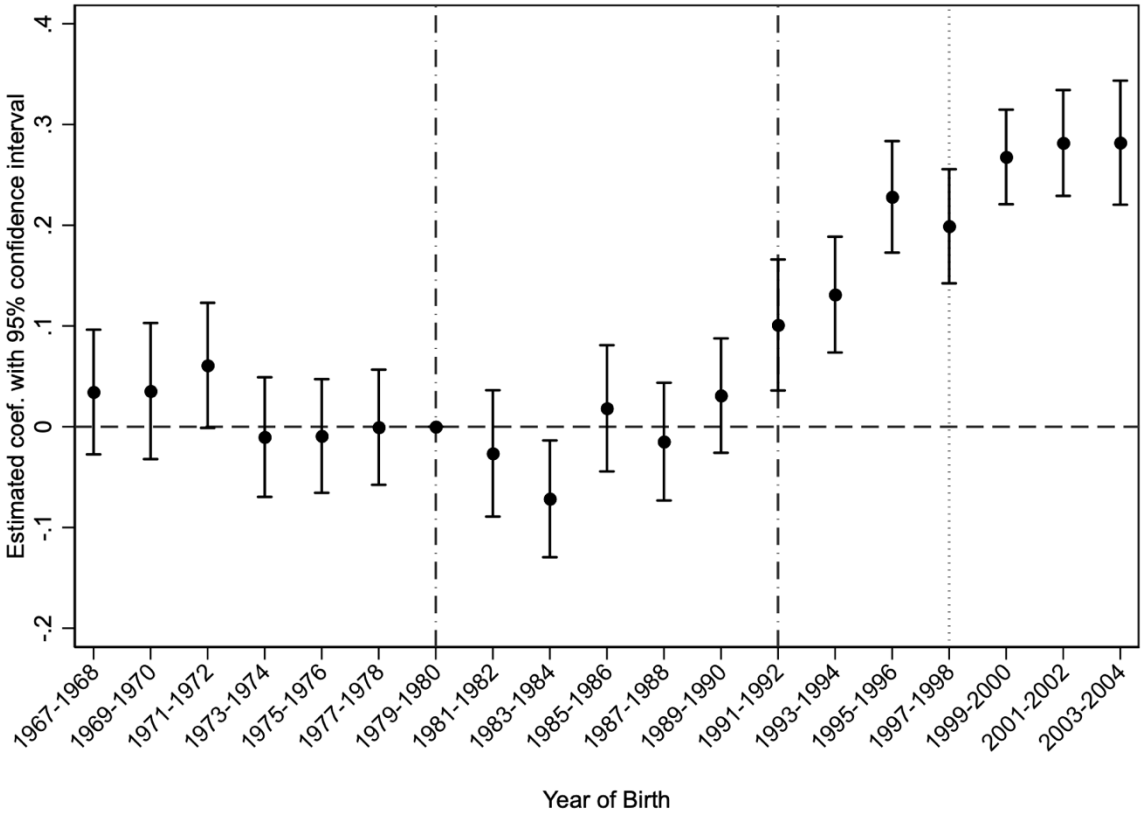


**Figure 5.** The event study of the reform in 1998. The figure reports the estimates of timing dummies with a 95% confidence interval in the regression of children's grade-for-age status on a sequence of dummies from 1968 to 2004, personal and family background characteristics, per student education expenditure, month dummies, and province-birth cohort-birth month dummies. The x-axis represents an individual's birth year. Birth cohorts are aggregated into two-year intervals to increase the power of the analysis. The control group consists of those born in 1979-1980 and is omitted in the regression. Source: The 0.095% sample of the 2010 Chinese Population Census.

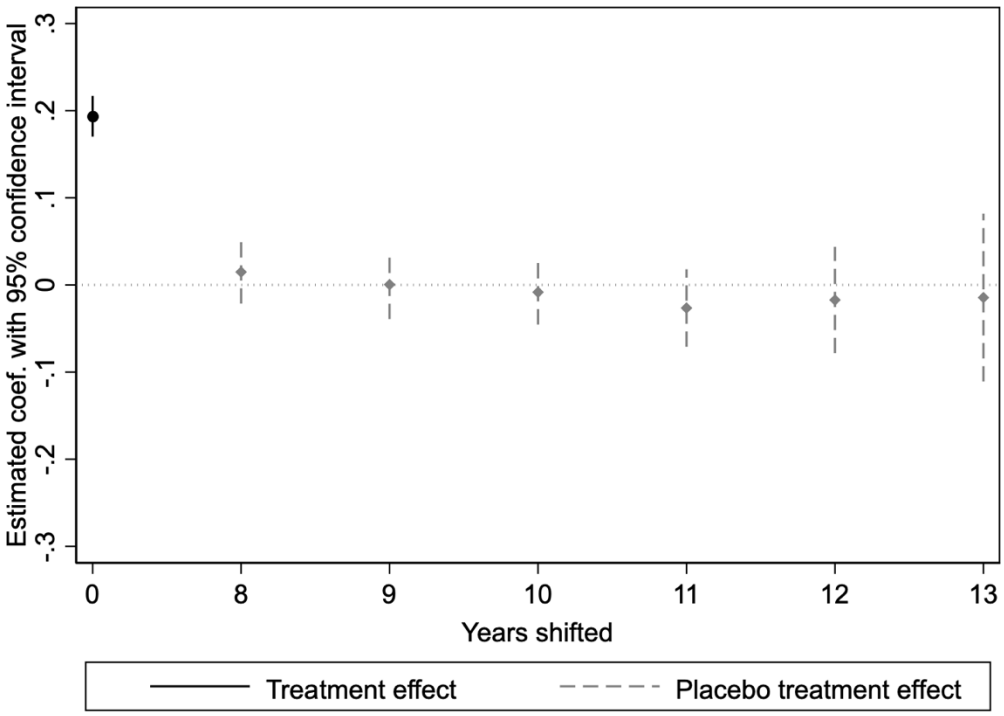


**Figure 6.** Estimated coefficients from the time-shifted placebo tests. The solid lines denote the true treatment effects estimated by equation (1). The dashed lines represent the estimates of the interaction of *Placebo Young* and *Treat* with a 95% confidence interval in the regression of DID models based on the pseudo-policy timing. The hukou reform is assumed to have occurred 8-13 years before the actual time. Source: The 0.095% sample of the 2010 Chinese Population Census.

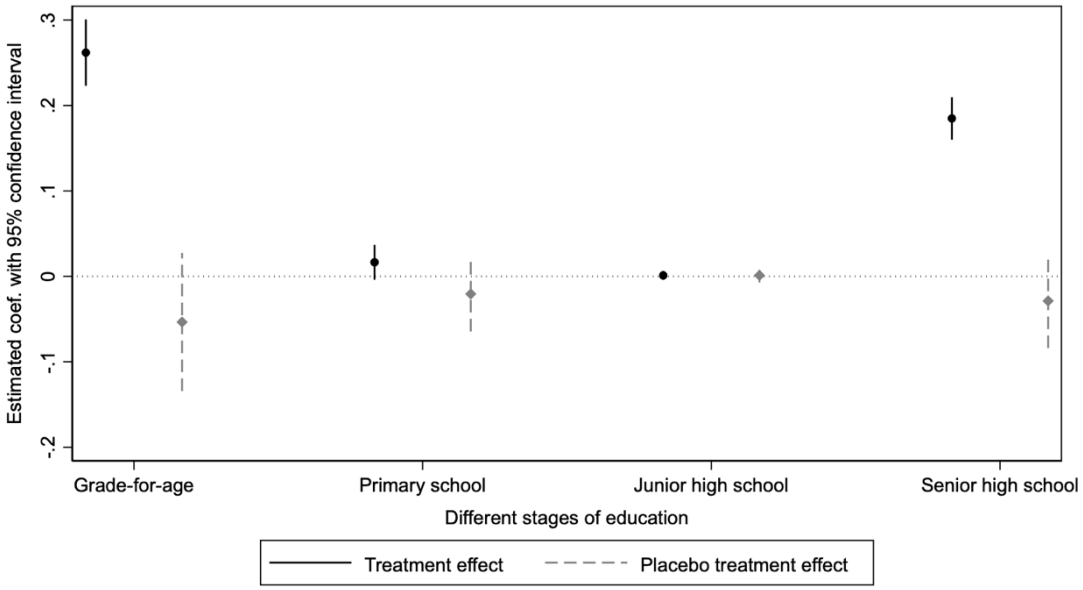


**Figure 7.** Estimated coefficients from the placebo tests using untreated samples. The solid lines denote the true treatment effects estimated by equation (1). The dashed lines represent the estimates of the interaction of *Young* and *Placebo Treat* with 95% confidence interval. *Placebo Treat* = 1 if a child's father has rural hukou and the mother has urban hukou; = 0 if a child's parents both have rural hukou. Source: The 0.095% sample of the 2010 Chinese Population Census.

**TABLE 1**

SPECIFIC RULES FOR THE HUKOU REFORM IN 1998

| Date of birth (1) | Age range (2) | Hukou registration policy (3) |
|---|---|---|
| Born on or after July 22, 1998 | Newborns | Voluntarily inherit the hukou status of their father or mother, with no conditions attached. |
| Born before July 22, 1998 | Preschoolers (Aged 1-6) | Register in the mother's domicile place first. Permission to apply for transfer to the father's hukou status and given priority. |
| | Minor children (Aged 7-17) | Register in the mother's domicile place first, gradually changing to the father's domicile place but not given priority. |
| | Adults | Not eligible for hukou transformation. |

**Note.** The State Council approved the *Opinions on Solving Several Outstanding Problems in the Current Management of Hukou* on July 22, 1998. We summarize the specific rules of the hukou reform based on various local legal and administrative documents.

**TABLE 2**

SUMMARY STATISTICS

| | Control group (*N*=37,430) | | Treatment group (*N*=3,604) | | Difference |
|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | Col.1 - Col.3 |
| | (1) | (2) | (3) | (4) | (5) |
| | A. Outcome variable | | | | |
| Grade-for-age: | 0.663 | 0.473 | 0.479 | 0.500 | 0.184*** |
| Older cohort | 0.361 | 0.480 | 0.106 | 0.307 | 0.256*** |
| Younger cohort | 0.814 | 0.389 | 0.744 | 0.436 | 0.070*** |
| | B. Treatment variable | | | | |
| Hukou (Urban = 1): | 1.000 | 0.000 | 0.345 | 0.475 | 0.655*** |
| Older cohort | 1.000 | 0.000 | 0.277 | 0.448 | 0.723*** |
| Younger cohort | 1.000 | 0.000 | 0.392 | 0.488 | 0.608*** |
| | C. Individual characteristics | | | | |
| Age | 20.854 | 10.579 | 22.346 | 10.965 | -1.492*** |
| Gender (Male = 1) | 0.566 | 0.496 | 0.625 | 0.484 | -0.059*** |
| Ethnicity (Han = 1) | 0.962 | 0.192 | 0.955 | 0.207 | 0.007** |
| Residence (Urban = 1) | 0.899 | 0.302 | 0.368 | 0.482 | 0.530*** |
| | D. Household characteristics | | | | |
| Number of siblings | 0.405 | 0.640 | 0.694 | 0.773 | -0.289*** |
| Father's age | 49.900 | 11.954 | 51.353 | 12.089 | -1.452*** |
| Father's years of schooling | 11.222 | 3.296 | 9.675 | 2.710 | 1.547*** |
| Mother's age | 47.696 | 11.334 | 49.370 | 11.529 | -1.674*** |
| Mother's years of schooling | 10.373 | 3.432 | 7.767 | 2.920 | 2.606*** |
| | E. Government's public spending on education | | | | |
| Per student education expenditure | 2.098 | 1.301 | 1.673 | 1.313 | 0.424*** |
| Older cohort | 0.762 | 0.263 | 0.556 | 0.197 | 0.206*** |
| Younger cohort | 2.768 | 1.079 | 2.466 | 1.185 | 0.301*** |

**Note.** The data source is the 0.095% sample of the 2010 Chinese Population Census. The control group consists of children with an urban father and an urban mother; the treatment group consists of children with an urban father and a rural mother. Individuals born from September 1991 to August 2004 are defined as the *Younger Cohort*; individuals born from September 1967 to August 1980 are defined as the *Older Cohort*. The paired *t*-test is applied for mean differences in the outcome and treatment variables of the control and treatment groups.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 3**

EFFECTS OF THE REFORM ON HUKOU STATUS

| | Hukou status | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treat | -0.667*** | -0.668*** | -0.592*** |
| | (0.014) | (0.014) | (0.017) |
| Young × Treat | 0.091*** | 0.093*** | 0.151*** |
| | (0.019) | (0.020) | (0.024) |
| Gender (Male=1) | 0.003** | 0.003* | 0.003** |
| | (0.001) | (0.001) | (0.001) |
| Ethnicity (Han=1) | 0.001 | 0.001 | 0.000 |
| | (0.004) | (0.004) | (0.004) |
| Residence (Urban=1) | 0.078*** | 0.079*** | 0.070*** |
| | (0.004) | (0.004) | (0.004) |
| Number of siblings | | -0.004*** | -0.001 |
| | | (0.001) | (0.001) |
| Father's age | | -0.001* | -0.000* |
| | | (0.000) | (0.000) |
| Father's years of schooling | | -0.001** | -0.001** |
| | | (0.000) | (0.000) |
| Mother's age | | -0.000 | -0.000 |
| | | (0.000) | (0.000) |
| Mother's years of schooling | | 0.000 | 0.000 |
| | | (0.000) | (0.000) |
| Per student education expenditure | | | 0.589*** |
| | | | (0.082) |
| Observations | 41,024 | 41,024 | 41,024 |
| $R^2$ | 0.663 | 0.663 | 0.726 |

**Note.** The dependent variable is children's hukou status (urban hukou=1). Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 4**

EFFECTS OF THE REFORM ON EDUCATIONAL OUTCOMES

| | Grade-for-age | | | Primary school | Junior high school | Senior high school |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Young × Treat | 0.151*** | 0.185*** | 0.189*** | 0.001 | 0.016 | 0.262*** |
| | (0.013) | (0.013) | (0.013) | (0.002) | (0.010) | (0.020) |
| | [0.000] | [0.000] | [0.000] | [0.544] | [0.036] | [0.000] |
| Treat | -0.180*** | -0.164*** | -0.158*** | -0.000 | -0.020*** | -0.199*** |
| | (0.011) | (0.011) | (0.011) | (0.002) | (0.008) | (0.017) |
| Gender (Male=1) | -0.031*** | -0.031*** | -0.031*** | 0.000 | -0.003 | -0.035*** |
| | (0.004) | (0.004) | (0.004) | (0.001) | (0.002) | (0.005) |
| Ethnicity (Han=1) | -0.033*** | -0.019* | -0.019* | 0.000 | 0.003 | 0.011 |
| | (0.011) | (0.011) | (0.011) | (0.001) | (0.007) | (0.016) |
| Residence (Urban=1) | 0.111*** | 0.065*** | 0.064*** | 0.001 | 0.016*** | 0.145*** |
| | (0.007) | (0.006) | (0.006) | (0.001) | (0.004) | (0.010) |
| Number of siblings | | -0.026*** | -0.026*** | -0.000 | -0.010*** | -0.037*** |
| | | (0.004) | (0.004) | (0.000) | (0.002) | (0.005) |
| Father's age | | 0.001 | 0.001 | -0.000** | 0.001** | 0.000 |
| | | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Father's years of schooling | | 0.019*** | 0.019*** | 0.000* | 0.003*** | 0.022*** |
| | | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Mother's age | | 0.002*** | 0.002*** | -0.000 | -0.001 | 0.002* |
| | | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Mother's years of schooling | | 0.012*** | 0.012*** | 0.000 | 0.002*** | 0.017*** |
| | | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Per student education expenditure | | | 0.049** | 0.002 | 0.011 | 0.211*** |
| | | | (0.019) | (0.002) | (0.017) | (0.053) |
| Observations | 41,024 | 41,024 | 41,024 | 41,024 | 33,743 | 25,248 |
| $R^2$ | 0.414 | 0.447 | 0.448 | 0.063 | 0.197 | 0.247 |

**Note.** The dependent variables in columns 1-3 are grade-for-age status. The dependent variables in columns 4-6 are dummies indicating whether individuals enter or have entered primary, junior high, and senior high school at the expected age. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level. The *p*-values (in square brackets) are computed using the wild bootstrap with 500 replications

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 5**

EFFECTS OF THE REFORM ON EDUCATIONAL OUTCOMES WITH LARGE OR SMALL URBAN-RURAL EDUCATION GAP

| | Grade-for-age (GFA) | | Primary school | | Junior high school | | Senior high school | |
|---|---|---|---|---|---|---|---|---|
| | Large | Small | Large | Small | Large | Small | Large | Small |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Young × Treat | 0.202*** | 0.160*** | 0.006 | -0.005** | 0.034** | 0.001 | 0.274*** | 0.243*** |
| | (0.019) | (0.018) | (0.004) | (0.002) | (0.016) | (0.012) | (0.029) | (0.026) |
| *p*-value | [0.111] | | [0.030] | | [0.110] | | [0.451] | |
| Mean of dependent variable | 0.743 | 0.745 | 0.995 | 0.998 | 0.915 | 0.911 | 0.642 | 0.692 |
| Mean of hukou | 0.399 | 0.387 | 0.399 | 0.387 | 0.367 | 0.384 | 0.480 | 0.366 |
| Observations | 21,942 | 19,004 | 21,942 | 19,004 | 18,193 | 15,503 | 13,690 | 11,513 |
| $R^2$ | 0.434 | 0.492 | 0.116 | 0.069 | 0.207 | 0.222 | 0.267 | 0.267 |

**Note.** We measure the urban-rural education gap by calculating the ratio of years of schooling for urban and rural adults aged over 25 in 2010. The prefectures equal to and above the median are defined as having a *Large* gap, whilst those below the median are defined as having a *Small* gap. The *p*-values (in square brackets) reflect the difference between the two groups. The mean of dependent variables and hukou (urban=1) are calculated based on the younger cohorts in the treatment group. Regressions control for a set of individual (gender, ethnicity, residence), household (number of siblings, parent's age, and years of schooling), and government's public spending on education (per student education expenditure) variables. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 6**

ROBUSTNESS CHECKS: CONFOUNDING FACTORS

| | Grade-for-age | Primary school | Junior high school | Senior high school |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | A. One child policy | | | |
| Young × Treat | 0.144*** | -0.001 | 0.021 | 0.196*** |
| | (0.022) | (0.003) | (0.013) | (0.033) |
| Observations | 34,985 | 34,985 | 28,980 | 21,081 |
| $R^2$ | 0.580 | 0.320 | 0.431 | 0.460 |
| | B. Mergers and closures of rural schools | | | |
| Young × Treat | 0.132*** | 0.002 | 0.050*** | 0.219*** |
| | (0.022) | (0.005) | (0.018) | (0.044) |
| Observations | 40,959 | 40,959 | 33,692 | 25,226 |
| $R^2$ | 0.451 | 0.065 | 0.201 | 0.249 |
| | C. Free compulsory education reform | | | |
| Young × Treat | 0.135*** | 0.002 | 0.051*** | 0.221*** |
| | (0.021) | (0.006) | (0.018) | (0.043) |
| Observations | 41,024 | 41,024 | 33,743 | 25,248 |
| $R^2$ | 0.450 | 0.063 | 0.198 | 0.248 |
| | D. Four Thousand Billion Stimulus Plan | | | |
| Young × Treat | 0.198*** | 0.002 | 0.013 | 0.261*** |
| | (0.013) | (0.002) | (0.012) | (0.022) |
| Observations | 37,656 | 37,656 | 30,375 | 21,880 |
| $R^2$ | 0.448 | 0.066 | 0.197 | 0.242 |

**Note.** In panel A, we additionally control for county-birth year fixed effects. In Panel B, we further control the interactions between the intensity of school closure at the city level from 2000 to 2010 and birth cohort dummies. In Panel C, we further control the number of semesters that an individual is supposed to be exposed to the FCE reform. As employers are prohibited from employing children under 16, Panel D excludes individuals in the treatment group who were 16 or above in 2008. Regressions control for a set of individual (gender, ethnicity, residence), household (number of siblings, parent's age, and years of schooling), and government's public spending on education (per student education expenditure) variables. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 7**

ROBUSTNESS CHECKS: OTHERS

| | Grade-for-age | Primary school | Junior high school | Senior high school |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | A. Redefining dependent variables | | | |
| Young × Treat | 0.209*** | 0.001 | 0.029*** | 0.265*** |
| | (0.013) | (0.002) | (0.009) | (0.022) |
| Observations | 41,004 | 40,656 | 30,939 | 22,442 |
| $R^2$ | 0.521 | 0.042 | 0.080 | 0.250 |
| | B. Considering spillover effects | | | |
| Young × Treat | 0.185*** | 0.001 | 0.018* | 0.260*** |
| | (0.013) | (0.002) | (0.010) | (0.020) |
| Siblings' exposure | 0.033*** | -0.001 | -0.010* | 0.009 |
| | (0.007) | (0.001) | (0.005) | (0.010) |
| Observations | 41,024 | 41,024 | 33,743 | 25,248 |
| $R^2$ | 0.448 | 0.063 | 0.198 | 0.247 |
| | C. Removing migrant children | | | |
| Young × Treat | 0.171*** | -0.002 | 0.017 | 0.217*** |
| | (0.017) | (0.003) | (0.014) | (0.027) |
| Observations | 25,686 | 25,686 | 20,930 | 15,180 |
| $R^2$ | 0.472 | 0.096 | 0.224 | 0.265 |
| | D. Redefining older cohort to 7-17 years old | | | |
| Young × Treat | 0.177*** | -0.002 | 0.018 | 0.240*** |
| | (0.017) | (0.003) | (0.014) | (0.027) |
| Observations | 25,686 | 25,686 | 20,930 | 15,180 |
| $R^2$ | 0.472 | 0.096 | 0.224 | 0.266 |
| | E. Removing 15 sub-provincial cities and individuals with siblings | | | |
| Young × Treat | 0.158*** | -0.005 | 0.003 | 0.261*** |
| | (0.022) | (0.004) | (0.016) | (0.035) |
| Observations | 14,073 | 14,073 | 11,574 | 8,097 |
| $R^2$ | 0.507 | 0.121 | 0.237 | 0.312 |

**Note.** In panel A, the dependent variables are redefined by adjusting the age of entry into a primary school to 7 years. In panel B, we further control Siblings' exposure to the hukou reform. In panel C, we remove samples that leave the place of hukou registration. In panel D, the age range of the older cohort is changed from 18-30 to 7-17 years old. In panel E, we remove samples from 15 sub-provincial cities and individuals with siblings. Regressions control for a set of individual (gender, ethnicity, residence), household (number of siblings, parent's age, and years of schooling), and government's public spending on education (per student education expenditure) variables. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 8**

HETEROGENEOUS EFFECTS

| | Grade-for-age | | Primary school | | Junior high school | | Senior high school | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | A. Gender of children | | | | | | | |
| | Girl | Boy | Girl | Boy | Girl | Boy | Girl | Boy |
| Young × Treat | 0.232*** | 0.170*** | -0.000 | 0.002 | 0.040** | 0.005 | 0.335*** | 0.240*** |
| | (0.022) | (0.016) | (0.005) | (0.003) | (0.019) | (0.013) | (0.033) | (0.027) |
| *p*-value | [0.019] | | [0.707] | | [0.120] | | [0.029] | |
| Mean of dependent variable | 0.755 | 0.735 | 0.995 | 0.998 | 0.915 | 0.910 | 0.680 | 0.662 |
| Mean of hukou | 0.380 | 0.403 | 0.380 | 0.403 | 0.368 | 0.384 | 0.343 | 0.393 |
| Observations | 17,534 | 23,429 | 17,534 | 23,429 | 14,225 | 19,487 | 10,273 | 14,944 |
| $R^2$ | 0.433 | 0.475 | 0.150 | 0.086 | 0.218 | 0.215 | 0.285 | 0.256 |
| | B. Parental education level | | | | | | | |
| | Educated | Less-educated | Educated | Less-educated | Educated | Less-educated | Educated | Less-educated |
| Young × Treat | 0.320*** | 0.065*** | -0.001 | 0.003 | 0.017 | 0.018 | 0.347*** | 0.178*** |
| | (0.025) | (0.016) | (0.004) | (0.003) | (0.015) | (0.014) | (0.032) | (0.029) |
| *p*-value | [0.000] | | [0.397] | | [0.988] | | [0.000] | |
| Mean of dependent variable | 0.786 | 0.713 | 0.996 | 0.998 | 0.930 | 0.900 | 0.738 | 0.620 |
| Mean of hukou | 0.418 | 0.373 | 0.418 | 0.373 | 0.385 | 0.370 | 0.393 | 0.352 |
| Observations | 23,946 | 17,025 | 23,946 | 17,025 | 18,804 | 14,915 | 12,947 | 12,277 |
| $R^2$ | 0.366 | 0.491 | 0.097 | 0.110 | 0.227 | 0.211 | 0.228 | 0.220 |
| | C. City type | | | | | | | |
| | Capital | Not capital | Capital | Not capital | Capital | Not capital | Capital | Not capital |
| Young × Treat | 0.216*** | 0.163*** | 0.007 | -0.002 | 0.017 | 0.018 | 0.260*** | 0.249*** |
| | (0.025) | (0.015) | (0.006) | (0.003) | (0.019) | (0.011) | (0.039) | (0.023) |
| *p*-value | [0.068] | | [0.252] | | [0.986] | | [0.740] | |
| Mean of dependent variable | 0.743 | 0.745 | 0.996 | 0.997 | 0.913 | 0.912 | 0.678 | 0.668 |
| Mean of hukou | 0.415 | 0.385 | 0.415 | 0.385 | 0.395 | 0.370 | 0.397 | 0.360 |
| Observations | 15,502 | 25,452 | 15,502 | 25,452 | 12,879 | 20,820 | 9,858 | 15,346 |
| $R^2$ | 0.431 | 0.487 | 0.194 | 0.069 | 0.237 | 0.217 | 0.273 | 0.261 |

**Note.** In Panel A, we run separate regressions for the subsamples of boys and girls. Panel B defines a family as *Educated* if either parent has attended senior high school or above and Less-educated otherwise. Panel C defines the *Capital* dummy variable that distinguishes municipalities and provincial capitals from smaller cities. The *p*-values (in square brackets) reflect the difference between the two groups. The mean of dependent variables and hukou (urban=1) are calculated based on the younger cohorts in the treatment group. Regressions control for a set of individual (gender, ethnicity, residence), household (number of siblings, parent's age, and years of schooling), and government's public spending on education (per student education expenditure) variables. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE 9**

MECHANISM ANALYSIS

| | School location (Located in cities = 1) | Extra-paid educational expenses (CNY: thousand) | Educational aspiration (College degree or above = 1) |
|---|---|---|---|
| | (1) | (2) | (3) |
| Young × Treat | 0.208*** | -3.587 | 0.134** |
| | (0.041) | (5.306) | (0.052) |
| Observations | 12,994 | 2,820 | 5,336 |
| $R^2$ | 0.173 | — | 0.162 |

**Note.** In column (1), The dependent variable is "school location," which equals 1 for cities and 0 for counties and rural areas. In column (2), we apply the "extra-paid educational expenses" as a proxy for the cost of education. In column (3), The dependent variable is educational aspiration, which equals 1 if an individual expects to obtain college or above degree and 0 otherwise. All regressions control for a set of individual and household characteristics, including age, gender, ethnicity, parents' age, and years of schooling. Wave, province, and birth year fixed effects are included in all regressions. Standard errors clustered at the province-birth year level are in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

## Appendix

**TABLE A1**

POLICY ISSUE DATE AND THE FIRST AFFECTED COHORT

| Province | Affected cohort born from | The issue date |
|---|---|---|
| (1) | (2) | (3) |
| Heilongjiang | Jul. 01, 1998 | Aug. 19, 1998 |
| Ningxia | Jul. 22, 1998 | Sep. 11, 1998 |
| Guangdong | Jul. 22, 1998 | Oct. 23, 1998 |
| Inner Mongolia | Jul. 22, 1998 | Nov. 4, 1998 |
| Jilin | Jul. 22, 1998 | Nov. 13, 1998 |
| Zhejiang | Jul. 22, 1998 | Dec. 10, 1998 |
| Henan | Jul. 22, 1998 | Dec. 18, 1998 |
| Shaanxi | Jul. 22, 1998 | Dec. 22, 1998 |
| Sichuan | Jul. 22, 1998 | Dec. 22, 1998 |
| Fujian | Jul. 22, 1998 | Dec. 25, 1998 |
| Shanxi | Jul. 22, 1998 | Dec. 25, 1998 |
| Guangxi | Jul. 22, 1998 | Dec. 30, 1998 |
| Hebei | Jul. 22, 1998 | Mar. 22, 1999 |
| Hunan | Jul. 22, 1998 | Mar. 22, 1999 |
| Anhui | Jul. 22, 1998 | Jan, 22, 1999 |
| Hubei | Jul. 22, 1998 | Apr. 26, 1999 |
| Shandong | Jul. 22, 1998 | Jun. 6, 1999 |
| Jiangxi | Jul. 22, 1998 | Sep. 30, 1999 |
| Gansu | Jul. 22, 1998 | Mar. 31, 2000 |
| Liaoning | Jul. 22, 1998 | Mar. 31, 2000 |
| Guizhou | Jul. 22, 1998 | Apr. 30, 2000 |
| Beijing | Jul. 22, 1998 | Feb. 26, 2001 |
| Chongqing | Jul. 22, 1998 | Aug. 27, 2003 |
| Shanghai | Aug. 07, 2003 | Aug. 21, 2003 |
| Tianjin | Aug. 07, 2003 | Aug. 08, 2003 |
| Hainan | — | — |
| Jiangsu | — | — |
| Qinghai | — | — |
| Tibet | — | — |
| Xinjiang | — | — |
| Yunnan | — | — |

**Note.** Authors' construction based on various local legal and administrative documents. Province here includes province, autonomous region, and provincial-level municipality. The issue date refers to the date that provinces/municipalities enacted the hukou reform. The information on the effective dates of the policy implementation in six provinces – Tibet, Xinjiang, Hainan, Yunnan, Jiangsu, and Qinghai – is missing.

**TABLE A2**

COMPOSITION OF FAMILY TYPES

| Family type | Hukou status (Father, Mother) | Percent (%) | Sample size |
|---|---|---|---|
| (1) | (2) | (3) | (4) |
| Type 1 | (Urban, Rural) | 2.48 | 10,067 |
| Type 2 | (Urban, Urban) | 26.33 | 106,942 |
| Type 3 | (Rural, Urban) | 1.16 | 4,706 |
| Type 4 | (Rural, Rural) | 70.04 | 284,498 |
| Total | | 100.00 | 406,213 |

**Note.** The data source is the 0.095% sample of the 2010 Chinese Population Census. Couples are identified based on the date of the first marriage. The variable "relationship to the household head" is then used to match children with their parents and siblings.

**TABLE A3**

SAMPLE SCREENING PROCESS

| Basis of sample selection | Sample size |
|---|---|
| (1) | (2) |
| Retain Type 1 (treatment group) and Type 2 (control group) families | 117,009 |
| Retain the sample with parents whose first marriage occurred in 1998 and before | 94,567 |
| Restrict age range | 51,750 |
| Drop the sample with missing key variables | 51,197 |
| Delete the sample from municipalities and provinces with unclear policies | 41,034 |

**Note.** The data source is the 0.095% sample of the 2010 Chinese Population Census.

**TABLE A4**

Comparison of children born before the reform

| | Treatment group (*N*=3,604) | | Children with rural fathers and urban mothers (*N*=1,632) | | Children with rural fathers and rural mothers (*N*=114,000) | | Difference | Difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. | Col.1-Col.3 | Col.1-Col.5 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Grade-for-age | 0.367 | 0.482 | 0.574 | 0.495 | 0.336 | 0.473 | -0.207*** | 0.031*** |
| Hukou (Urban = 1) | 0.321 | 0.467 | 0.717 | 0.451 | 0.012 | 0.110 | -0.395*** | 0.309*** |
| Age | 25.223 | 10.108 | 19.080 | 8.139 | 21.190 | 9.035 | 6.143*** | 4.033*** |
| Gender (Male = 1) | 0.650 | 0.477 | 0.591 | 0.492 | 0.604 | 0.489 | 0.059*** | 0.046*** |
| Ethnicity (Han = 1) | 0.955 | 0.206 | 0.960 | 0.197 | 0.950 | 0.217 | -0.004 | 0.005 |
| Residence (Urban = 1) | 0.329 | 0.470 | 0.563 | 0.496 | 0.097 | 0.296 | -0.234*** | 0.232*** |
| Number of siblings | 0.676 | 0.785 | 0.690 | 0.777 | 0.975 | 0.918 | -0.014 | -0.299*** |
| Father's age | 53.852 | 11.776 | 46.304 | 9.201 | 48.869 | 10.741 | 7.548*** | 4.982*** |
| Father's years of schooling | 9.497 | 2.703 | 9.108 | 2.542 | 8.031 | 2.228 | 0.388*** | 1.466*** |
| Mother's age | 51.814 | 11.196 | 44.588 | 8.883 | 47.070 | 10.167 | 7.226*** | 4.744*** |
| Mother's years of schooling | 7.458 | 2.956 | 8.999 | 2.896 | 7.000 | 2.728 | -1.541*** | 0.458*** |

*Notes*: Based on individuals born between September 1967 and July 1998 from the 0.095% sample of the 2010 Chinese Population Census. Treatment group refers to children with urban fathers and rural mothers.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE A5**

EFFECTS OF THE REFORM ON EDUCATIONAL OUTCOMES WITH MATCHED SAMPLES

| | Grade-for-age | Primary school | Junior high school | Senior high school |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Young × Treat | 0.094*** | -0.001 | 0.018* | 0.216*** |
| | (0.013) | (0.002) | (0.010) | (0.021) |
| Treat | -0.088*** | 0.001 | -0.026*** | -0.167*** |
| | (0.011) | (0.002) | (0.008) | (0.017) |
| Gender (Male=1) | -0.013*** | 0.000 | -0.000 | -0.014** |
| | (0.004) | (0.001) | (0.003) | (0.006) |
| Ethnicity (Han=1) | -0.017 | 0.000 | 0.004 | 0.003 |
| | (0.012) | (0.002) | (0.009) | (0.021) |
| Residence (Urban=1)) | 0.073*** | 0.001 | 0.012*** | 0.143*** |
| | (0.007) | (0.001) | (0.004) | (0.010) |
| Number of siblings | -0.029*** | 0.001 | -0.006** | -0.032*** |
| | (0.004) | (0.001) | (0.003) | (0.006) |
| Father's age | -0.000 | -0.000 | 0.001* | -0.002 |
| | (0.001) | (0.000) | (0.000) | (0.001) |
| Father's years of schooling | 0.015*** | 0.000 | 0.003*** | 0.023*** |
| | (0.001) | (0.000) | (0.001) | (0.001) |
| Mother's age | 0.002*** | -0.000 | -0.001 | 0.004*** |
| | (0.001) | (0.000) | (0.001) | (0.001) |
| Mother's years of schooling | 0.002** | 0.000 | 0.001** | 0.011*** |
| | (0.001) | (0.000) | (0.001) | (0.002) |
| Per student education expenditure | 0.052*** | 0.001 | 0.012 | 0.201*** |
| | (0.018) | (0.001) | (0.015) | (0.046) |
| Observations | 30,006 | 30,006 | 24,810 | 17,862 |
| $R^2$ | 0.506 | 0.091 | 0.223 | 0.274 |

**Note.** Estimates are obtained using matched samples. The dependent variable in column 1 is grade-for-age status. The dependent variables in columns 2-4 are dummies indicating whether individuals enter or have entered primary, junior high, and senior high school at the expected age, respectively. Birth month and province-birth cohort-birth month fixed effects are included in all regressions. Standard errors (in parentheses) are clustered at the province-year level.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

**TABLE A6**

Demographic characteristics

| | Younger cohort | | | Older cohort | |
|---|---|---|---|---|---|
| | Census | CFPS | | Census | CGSS |
| | (1) | (2) | | (3) | (4) |
| Age group | | | Age group | | |
| Aged 6-8 | 0.217 | 0.233 | Aged 30-32 | 0.227 | 0.193 |
| Aged 9-11 | 0.220 | 0.241 | Aged 33-35 | 0.229 | 0.217 |
| Aged 12-14 | 0.237 | 0.238 | Aged 36-38 | 0.266 | 0.282 |
| Aged 15-17 | 0.327 | 0.288 | Aged 39-41 | 0.279 | 0.308 |
| Level of education | | | Level of education | | |
| Primary school and below | 0.440 | 0.407 | Primary school and below | 0.145 | 0.248 |
| Junior high school | 0.313 | 0.302 | Junior high school | 0.524 | 0.384 |
| Senior high school | 0.207 | 0.282 | Senior high school | 0.173 | 0.183 |
| College degree and above | 0.040 | 0.009 | College degree and above | 0.158 | 0.185 |
| Hukou (Urban = 1) | 0.169 | 0.143 | Hukou (Urban = 1) | 0.202 | 0.222 |
| Gender (Male = 1) | 0.300 | 0.330 | Gender (Male = 1) | 0.279 | 0.277 |
| Ethnicity (Han = 1) | 0.531 | 0.527 | Ethnicity (Han = 1) | 0.520 | 0.501 |

**Note.** The older cohorts (individuals born between September 1967 and August 1980) are derived from the CGSS 2010, and the younger cohorts (individuals born between September 1991 and August 2004) are derived from the CFPS 2010. Values are weighted descriptive statistics. Source: Authors' calculations based on the 0.095% sample of the 2010 Chinese Population Census, CFPS 2010, and CGSS 2010.