Opportunities swept away: How do floods affect children’s educational outcomes in rural China?

Kevin Chen[a b], Pei Zeng[a], Ziqian Song[b], Xinjie Shi[a*]

[a] *School of Public Affairs, Zhejiang University, Hangzhou, China*

[b] *International Food Policy Research Institute, East and Central Asia Office, Beijing, China*

Kevin Chen: Qiushi Chair Professor, China Academy for Rural Development, School of Public Affairs, Zhejiang University, Hangzhou, China Kzchen@zju.edu.cn.

Pei Zeng: China Academy for Rural Development, School of Public Affairs, Zhejiang University, Hangzhou, China zengpei1998@163.com.

Ziqian Song: International Food Policy Research Institute, East and Central Asia Office, Beijing, China z.song@cgiar.org.

Xinjie Shi (Corresponding author): Assistant Professor, China Academy for Rural Development, School of Public Affairs, Zhejiang University, Hangzhou, China xjshi1990@zju.edu.cn.

Abstract

The frequency of natural disasters has surged globally in recent years. However, the existing literature paid limited attention to the role of government and individual disaster response strategies on shaping human capital outcomes in disaster-affected populations. This paper uses the 1998 nationwide floods in China as a quasi-natural experiment to estimate the loss of educational opportunities for rural children exposed to natural disasters. By matching county-level precipitation anomaly data with the 2010 census data, the difference-in-difference analysis showed that individuals in early life stages (ages 7-15) in rural areas exposed to stronger floods experienced a significant reduction in educational opportunities. This negative impact of the 1998 China floods on educational opportunities was more pronounced among females and impeded intergenerational mobility in rural areas. The mechanism behind this finding is twofold. First, the impact of flooding increased the burden on rural households and reduced their investment in their children’s education, with this deprivation effect being particularly pronounced in areas more dependent on the primary sector. Second, Due to early experiences of flooding, there is an increased likelihood of parents migrating to cities for work. While the potential increase in income may have positive effects, the negative impact of leaving children behind tends to dominate. However, post-disaster investment in education from the government helped mitigate the long-term negative impacts of the floods to some extent. This paper minimizes estimation bias caused by population migration and re-emphasizes the negative impacts of natural disasters on human capital accumulation, elucidating the different response behaviors of the government and farm households, contributing to understanding the impact of different actors’ responses to natural disasters on the long-term human capital accumulation of the offspring, providing a reference for post-disaster relief efforts in countries or regions in transition.

1. Introduction

As climate change intensifies, the frequency of natural disasters has surged globally. Over the past 20 years, there have been 7,348 hazardous events, leading to the loss of 1.23 million lives globally[1]. Notably, Asia has been particularly affected, accounting for nearly 40% of global natural disasters and almost 90% of all victims between 2001 and 2010 (Paudel and Ryu, 2018). Among these disasters, flooding stands out as the most common. In fact, the incidence of flooding has risen by 134% compared to the previous two decades, impacting more than 4.03 billion people worldwide [1]. A significant example is the 1998 China floods, which swept across the country, affecting 223 million individuals and resulting in over $27.98 billion in economic losses.

Natural disasters affect the accumulation of human capital, particularly in less developed regions that struggle to cope with their devastating impacts (Caruso, 2017). In China, vast rural areas are especially vulnerable due to inadequate infrastructure and weak disaster relief capacity. The damage to these rural communities is a major concern highlighted in the disaster indicators published by the Ministry of Emergency Management[2]. During the unprecedented flood disaster of 1998, China also experienced dramatic economic and social reforms. One notable change was the relaxation of the household registration system, which facilitated rural-to-urban migration and provided rural residents with more options for coping with natural disasters. However, much of the existing literature focuses on the direct effects of natural disasters on individuals and the consequences of government interventions, with limited attention to how individual disaster coping strategies impact human capital (Lu et al., 2023; Wang et al., 2017; Liu and Xu, 2021). This study aims to fill this gap by examining how the 1998 floods affected access to education in rural China, taking into account both individual coping strategies and external interventions.

This paper uses data from a large-scale nationwide census conducted by the Chinese government to assess the impact of exposure to the 1998 floods on educational opportunities for China's rural residents. Our difference-in-differences analysis results reveals that the floods significantly hindered educational access. Specifically, for every anomalous 1 mm increase in annual average daily precipitation in a region, the likelihood of a typical rural resident continuing their high school education decreased by 1.5 percentage points. In contrast, we found no significant impact of the floods on educational access for urban residents. Furthermore, our findings indicate that females were more likely to experience educational deprivation following the flood, while higher levels of parental education could mitigate this deprivation. The adverse effects of increased household burdens from the floods, which reduced investments in children's education, along with the negative externalities associated with left-behind children due to rural migrant workers coping with the flood's aftermath, are the primary mechanisms through which the floods negatively impacted rural education. The study also highlights the crucial role of external interventions and individual post-disaster coping strategies, confirming the detrimental effects of natural disasters on human capital accumulation (Opper et al., 2023; Tian et al., 2022; Deuchert and Felfe, 2015).

The paper contributes to four strands of literature. First, it relates to the long-term impact

[1] See official document in https://www.un-ilibrary.org/content/books/9789210054478
[2] See official document in https://www.gov.cn/lianbo/bumen/202401/content_6927328.htm

of natural disasters on education. Previous studies have documented the negative effects of such calamities on educational attainment. For instance, Wang et al. (2017) found that individuals affected by the Tangshan Earthquake in 1976 averaged 14% to 21% fewer years of schooling. Some discussions on floods have revealed similar conclusions. Zhang and Zhang (2023) found that survivors of the 1975 Zhumadian flood experienced a reduction of 0.2 years in schooling. Similarly, Yu and Hu (2024) documented that human capital accumulation in the Yellow River floodplain was 12.1% lower than in other regions. Huang and Dong (2025) estimated the impact of floods on individual human capital in China from 1985 to 2015. They found that early-life flood exposure led to an average decrease of 0.786 years in education and a 5.1 percentage point drop in high school enrollment rates. Although extensive literature has examined the impact of natural disasters, including floods, on human capital, two issues remain unresolved. First, there is the estimation bias caused by population migration. Since the 1980s, China's *hukou* system reform and economic liberalization have triggered widespread migration. This makes it challenging to accurately identify an individual's disaster experience, inevitably introducing estimation bias. Most existing studies address this issue by simply excluding migrant samples (Yu and Hu, 2024; Huang and Dong, 2025). However, given the scale and frequency of migration in China, more satisfactory solutions are needed. Second, there is the issue of disaster representativeness. Most studies focus on specific localized disasters without considering a nationwide perspective. Huang and Dong (2025) contributed significantly by estimating the relationship between all major floods in China from 1985 to 2015 and human capital accumulation. However, their causal inference remains challenging. They attempted to identify causal effects using the 1998 flood as a quasi-natural experiment, but they also face the challenge of estimation bias caused by migration. Furthermore, their flood dataset only recorded the locations of major flood events in 1998, despite over 15,000 dykes suffering damage that year. This could lead to an overestimation of the flood's impact. To address these gaps, this study employs the 1998 flood — a nationally representative disaster — as a quasi-natural experiment. We use precipitation anomalies as a proxy for flood exposure to reassess the causal relationship between the 1998 flood and human capital accumulation. Additionally, we leverage China's compulsory education system and rural schooling patterns to define appropriate age groups, thereby mitigating the 33.3% upward estimation bias caused by migration.

Second, this paper further deepens the understanding of the government's role in post-disaster relief. Most existing literature focuses on the negative impact of natural disasters on human capital accumulation. However, there are contrasting views regarding the educational impact of natural disasters, particularly regarding the potential for post-disaster government and international aid to foster local development. For example, one study found that following the devastating Indian Ocean tsunami, economic growth in Aceh was bolstered by external aid (Philipp and Neumayer, 2019). Liu and Xu (2021) noted that increased investment in disaster-affected regions often leads to educational improvements. These mixed results prompt us to further explore whether government assistance can mitigate the negative shocks caused by natural disasters. We provide evidence of the negative effects of natural disasters on education and discuss how government relief efforts can alleviate these impacts, bridging the gap in the discourse surrounding disasters and education.

Third, this article enriches the discussion on disaster coping strategies among rural households. Due to poverty, rural households often lack investment in disaster prevention. In response to natural disasters, these households typically reduce expenditures, sell productive assets (Duflo, 2003; Jensen, 2000), and optimize agricultural operations through land transfers (Eskander and Barbier, 2023). Additionally, farmers often seek non-farm employment opportunities post-disaster (Mueller and Quisumbing, 2011). Before the 1980s in China, compensating for losses from natural disasters through non-farm work was challenging due to strict household registration controls. However, after the 1980s, the Chinese government relaxed these controls, making migrant labor more common (Chan and Zhang, 1999). Existing discussions have primarily focused on how migrant labor mitigates economic losses from disasters, often overlooking the negative impact on the education of children left behind. This article addresses the relationship between natural disasters and the coping responses of farming households, emphasizing the educational losses faced by left-behind children due to family migration.

Finally, this paper contributes to the research on natural disasters and inequality. Previous studies have examined inequalities related to gender (Paudel and Ryu, 2018), income (Bui et al., 2014; Pathak and Schündeln, 2022), and regional disparities (Warr and Aung, 2018) resulting from natural disasters. These inequalities, stemming from differences in mortality rates (Caruso and Miller, 2015) and varying levels of resilience (Van, 2010), can persist across generations (Caruso and Miller, 2015; Nakamura et al., 2022). We expand the discussion on natural disasters and inequality from an educational perspective, highlighting that, in the context of natural disasters, parental education levels significantly influence their children's educational outcomes, thereby exacerbating intergroup inequalities and hindering intergenerational mobility in education.

The remainder of this paper is structured as follows. Section 2 provides background on the 1998 floods, outlining the characteristics and context. Sections 3 and 4 detail the data and econometric specifications used in the study, respectively. Section 5 presents the main results while Section 6 explores the channels through which the floods affected rural human capital. Section 7 discusses the bias caused by migration and data. Finally, Section 8 concludes with a discussion.

2. Background

2.1 Unusual climatic phenomena

Three major climatic phenomena contributed to the 1998 China floods. First, the strongest El Niño of the 20th century occurred in 1997 and lasted until June 1998. This event led to significantly heavy precipitation in southern China during the spring of 1998, resulting in floods in various locations and a decrease in the storage capacity of rivers during the flood season. Second, an unusually strong western Pacific subtropical high-pressure system persisted from June to August 1998, a rare occurrence in the previous 40 years, and was a decisive factor influencing precipitation across China. Additionally, a high snowpack on the Qinghai-Tibetan Plateau during the winter of 1997 contributed to the substantial rainfall in the Yangtze River Basin. Collectively, these factors triggered the once-in-a-century floods that occurred from mid-June to September 1998 (Ministry of Water Resources of China, 2002). This information underscores that the flooding in China was largely due to anomalous precipitation patterns in 1998.

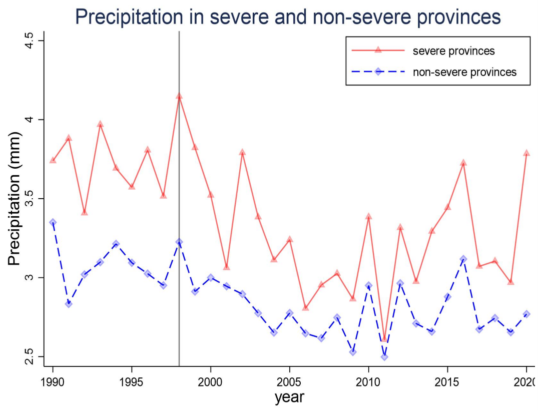
To illustrate the precipitation trends in 1998, we plotted the changes in average daily precipitation in China from 1990 to 2020, categorizing provinces into flood-severe and non-severe regions[3]. As shown in Figure 1, Panel (a) presents the average daily precipitation for the entire year, while Panel (b) shows the average daily precipitation for the summer months. The data clearly indicate that average daily precipitation in flood-severe regions was significantly higher than in non-severe regions in 1998. Notably, the average daily precipitation in the worst-affected areas was the highest recorded over the past 30 years. The non-flood severe regions also exhibited a wave peak in average daily precipitation, reflecting the widespread precipitation anomalies across China in 1998. Figure A.1 provides additional insights into precipitation changes in each province, corroborating that those along the Yangtze, Songhua, and Nenjiang rivers experienced significant peaks in average daily precipitation that year.

2.2 Post-disaster government response

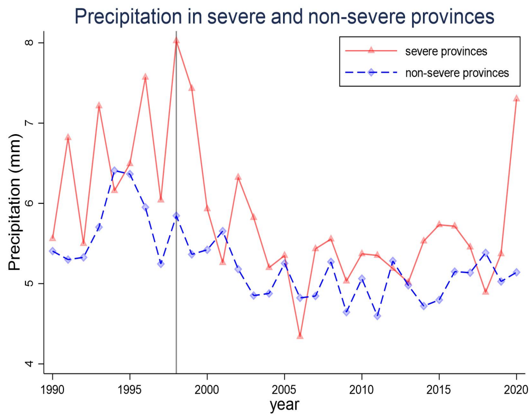
In response to the 1998 floods, the Chinese government mobilized 362,400 People's Liberation Army (PLA) troops to assist in rescue and relief efforts. More than 8 million government officials and citizens participated in the relief operations, deploying a total of 566,700 vehicles, 32,300 ships, and 2,241 aircraft and helicopter flights, with various rescue and relief materials valued at approximately $1.826 billion. Despite these extensive measures, the floods inflicted severe damage. During the floods, over 9,000 dykes in the Yangtze River basin were at risk, along with more than 6,000 dykes in the Songhua River basin. The Jiujiang dyke in the Yangtze River basin breached on August 7, 1998, threatening the safety of hundreds of thousands of people. Overall, 29 provinces experienced flooding of varying severity, which destroyed 55 million acres of farmland, resulting in over $21 billion in direct economic losses[4].

In the aftermath, the Chinese government increased financial support for the reconstruction of affected areas. By the end of August 1998, the Standing Committee of the Ninth National People's Congress approved the issuance of an additional $14 billion in treasury bonds by the Ministry of Finance to state-owned commercial banks. The central government's fiscal deficit expanded from $6.4 billion at the beginning of the year to $13.4 billion to provide special subsidies to local governments in disaster-stricken areas (He and Liu, 2024). Furthermore, to prevent future floods, the central government launched a large-scale "returning farmland to forest" project in 1999, completely banning logging of natural forests in the middle and upper reaches of the Yangtze and Yellow Rivers to conserve soil and water resources.

[3] Flood severe provinces include the Yangtze River Basin and the Songhua and Nenjiang River Basin provinces (including Sichuan, Chongqing, Hubei, Hunan, Jiangxi, Anhui, Jiangsu, Heilongjiang, and Jilin).
[4] See official document in https://www.gov.cn/xinwen/2015-06/25/content_2884116.htm



(a)Average daily precipitation in 1998



(b)Average daily precipitation in 1998 summer

Figure 1. Average daily precipitation in different provinces in China 1990 - 2020

Data source: the European Union and the European Organization of Medium-Range Weather Forecast Centers

3. Data

The primary data source for this paper is the 2010 census conducted by the Chinese government. This census includes individual-level information such as age, gender, hukou (a unique population registration system in China that distinguishes between agricultural and non-agricultural populations and records place of residence), and education status. To measure the impact of the 1998 floods, which were directly linked to anomalous regional precipitation caused by the El Niño phenomenon, we utilize the ERA5-Land dataset published by the European Union and the European Organization for Medium-Range Weather Forecasts. This dataset provides month-by-month daily mean precipitation raster data ($0.1^{\circ} \times 0.1^{\circ}$). We average the raster values within each county to obtain the final daily average precipitation for that area. The two databases are matched at the county level using county codes to construct cross-sectional data. To facilitate our discussion on government actions following the floods and farmers' disaster coping strategies, we also refer the 1995-2000 China Education Statistical Yearbook, rural fixed observation point data, and the 2000 census, matching these with the 1998 precipitation data in a similar manner.

Our dependent variable is whether the surveyed individual had a high school education by the end of 2010.. In China, high school education is not compulsory, and the costs associated with continuing high school are borne by the individual. In our econometric analysis, we restrict the sample to cohorts aged 7 to 24 (born between 1991 - 1974) in 1998, designating individuals aged 7 to 15 as the treatment group and those aged 16 to 24 as the control group.

There are several reasons for defining the treatment group as ages 7 to 15. First, children in China typically begin school at ages 6 or 7. Given that rural children may start school later, we set age 7 as the benchmark for schooling, ensuring that all individuals in the sample were at least 19 years old by 2010 and had completed high school. Age 15 represents the final year of compulsory education, after which individuals must decide whether to continue to high school. The control group consists of rural individuals aged 16 to 24 in 1998, as this cohort had already made decisions regarding high school attendance at the time of the floods. A potential concern is that individuals in the control cohort who were enrolled in high school might drop out due to the flood's impact, potentially biasing our estimates. However, this concern is mitigated by the census data, which records an individual's educational attainment as high school as long as they had chosen to attend, regardless of whether they completed it[5].

Table 1 presents the sample means for the outcome and control variables for both the control and treatment cohorts. These variables include high school continuation rates and demographic information such as ethnicity, gender, and age. We find that the control cohort had a high school continuation rate of 17.5%, while the treatment cohort's rate was 38%, indicating that the treatment cohort was significantly more likely to continue high school by 20%. Regarding demographic information, the differences between the control and treatment cohorts are relatively small, except for age.

Table 1. Summary statistics (2010 census)

	Control Cohort （16 - 24 years old in 1998）	Treatment Cohort （7 - 15 years old in 1998）	Difference
Variables	(1)	(2)	(3)
Whether to continue high school study = 1	0.175	0.38	0.204***
	(0.380)	(0.485)	
Han ethnic = 1	0.971	0.972	-0.001
	(0.168)	(0.165)	
Male = 1	0.567	0.541	-0.026***
	(0.495)	(0.498)	
Age	31.821	22.583	-9.237***
	(2.672)	(2.477)	
Observations	79025	97107	

Notes: Standard deviations are in parentheses. ***, **, * indicates significance at the 1%, 5%, 10% levels.

4. Empirical strategy

We employ a difference-in-differences identification strategy based on two sources of variation. First, different counties experienced varying intensities of flooding in 1998. Second, individuals from different cohorts within the same county were exposed to different levels of flood impact. Following Chen et al. (2020), we estimate the following equation:

$$Y_{i,g,c,p} = \beta_0 + \beta_1 Flood_{c,p} \times I(7 \leq g \leq 15) + \beta_2 X_{i,g,c,p} + \lambda_c + \gamma_{g,p} + \phi_c \times \gamma_g + \varepsilon_{i,g,c,p} \quad (1)$$

where $Y_{i,g,c,p}$ indicates whether individual i in cohort g residing in county c of p province had received a high school education. $Flood_{c,p}$ represents the flood intensity, quantified by the level of precipitation anomaly for each county in 1998, as the primary cause of the floods was abnormal precipitation started in 1998 spring. We use the absolute precipitation anomaly to capture precipitation extremes across China in 1998, a metric

[5]See official document in https://www.stats.gov.cn/sj/pcsj/rkpc/6rp/html/fu09.htm

commonly employed in meteorology to reflect changes in precipitation anomalies at specific locations (Dong et al., 2024). The calculation method is detailed in equation (2), which takes the difference between each county's average daily precipitation in 1998 and its average daily precipitation over the preceding 40 years. We visualize the anomaly of different counties in Figure 2.

$$Flood_{c,p} = \overline{precipitation_{c,p,1998}} - \overline{precipitation_{c,p,1958-1997}} \qquad (2)$$

$I(7 \leq g \leq 15)$ represents whether an individual was aged 7 to 15 at the time of the floods, assigning a value of 1 to this cohort and 0 to those aged 16 to 24 during the floods. $X_{i,g,c,p}$ is a series of control variables including an individual's ethnic and gender. λ_c is the county fixed effect and $\gamma_{g,p}$ is the province-by-cohort fixed effect, which accounts for differences at the provincial level across cohorts. To control for the initial level of education in each county, we include $\phi_c \times \gamma_g$, an interaction term between the level of basic education in each county and the birth cohort. The basic education level is calculated using the high school entrance rates of the control group in each county, as derived from the 2010 census.

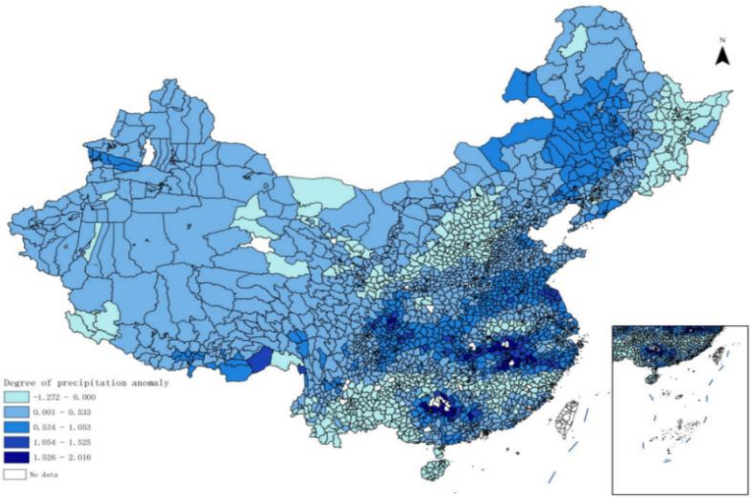


Figure 2. Precipitation anomalies by county in China in 1998

Two key challenges remain for our estimation. The first is policy overlap. In 1999, the Chinese government implemented a significant expansion of college enrollment, which inevitably influenced individuals' decisions regarding high school enrollment. With increased opportunities to attend college, many individuals were more inclined to pursue a high school education. This policy change had a limited effect on the control cohort, as they had already made their decisions about continuing high school. In contrast, the treatment cohort was directly affected, partly explaining why the treatment group had a much higher continuation rate compared to the control group. To mitigate this issue, we include province-by-cohort fixed effects to account for variations in university acceptance rates across different birth cohorts during their provincial university entrance exams, which can heavily influence decisions about high school enrollment (Xing, 2014). However, it is worth noting that these results might still underestimate the true impact of the floods to some extent.

The second challenge involves population migration. Following the reform of the household registration system in China in 1984, restrictions on population mobility were relaxed. Since the 2010 census data is significantly removed from the year of the floods, many individuals may have relocated, making it crucial to accurately identify those who experienced the floods for reliable estimation results. To address this, our study matches the intensity of the 1998 floods to the individuals' hukou locations. This approach is justified for two reasons. First, our sample consists of individuals with rural hukou as recorded in the 2010 census, which confirms that they were rural hukou holders in 1998. In China, hukou migration generally occurs from rural to urban areas, with relocations between rural areas or from urban to rural areas facing stricter conditions. Although policies vary by county, most allow only three scenarios for hukou relocations between rural areas or from urban to rural areas: between spouses, from parents to adult children, or from children to parents. Second, compulsory education in China typically must be completed at the hukou location, making migration during this stage unlikely. Consequently, since all individuals in our treatment group were in compulsory education in 1998, the likelihood of relocation during that time is minimal.

Another potential concern is that rural individuals, despite retaining their hukou, may have moved with their parents for schooling in 1998, potentially leading them to not experience the floods in their hukou location. However, the Interim Measures on Schooling for Migrant Children and Adolescents issued by the Chinese government in March 1998 imposed strict regulations on the schooling of migrant workers' children at their parents' workplaces, which alleviates this concern.

5 Results

5.1 Baseline results

We begin by estimating the impact of the 1998 floods on children or rural residents' access to high school education. As illustrated in Table 2, all columns reveal negative and statistically significant effects of the floods on educational attainment. In columns (1) to (3), we gradually incorporate control variables related to individual characteristics and the basic education status of the county. In column (3), the coefficient on the interaction term indicates that a 1mm increase in precipitation anomaly is associated with a 1.5 percentage point decrease in the likelihood of rural individuals attending high school. To account for the potential influence of individuals who relocated prior to the floods, we exclude samples from our analysis for those who were away from their hukou locations for more than six years (the 2010 census data only identify individuals who were away for this duration). The results presented in Table A1 continue to reflect a significant negative impact, with the coefficients remaining largely consistent.

Table 2. Impact of the 1998 floods on rural individual education attainment

	Whether to continue high school studies		
Variables	(1)	(2)	(3)
Affected cohort (7 - 15 in 1998) × Flood intensity	-0.016*	-0.017*	-0.015*
	(0.009)	(0.009)	(0.009)
Observations	176,293	176,293	176,132
R-squared	0.168	0.169	0.169
Controls	NO	YES	YES
Base education × Cohort trend	NO	NO	YES

County FE	YES	YES	YES
Province by cohort FE	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE and Province by cohort FE. Columns (2)-Column (3) sequentially incorporate individual control variables (including gender, and ethnicity) and Base education × Cohort trend. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

5.2 Test for parallel trends

A key assumption for the validity of difference-in-differences estimation is that the control and treatment groups exhibit similar trends in the outcome variable prior to the floods. We test for parallel trends by estimating the following equation using the event study method:

$$Y_{i,g,c,p} = \beta_0 + \sum_{t=7}^{23} \beta_{1,\tau} Flood_{c,p} \times I(g = \mathrm{t}) + \beta_2 X_{i,g,c,p} + \lambda_c + \gamma_{g,p} + \varepsilon_{i,g,c,p} \quad (3)$$

where $\beta_{1,\tau}$ represents the coefficients for each cohort aged 7 to 23 in 1998, with the cohort that turned 24 in 1998 serving as the reference cohort. Figure 3 illustrates the estimation results for each cohort along with the 95% confidence interval. The figure indicates that the cohort older than 15 years old in 1998 shows no significant differences from the control cohort, while the cohort younger than 15 years old demonstrates a significant negative effect. These results support the parallel trend hypothesis. The lack of significance for the 15-year-old cohort may be attributed to some individuals having completed their mandatory education and having made decisions regarding higher education. The impact of the floods appears to dissipate for those aged 10 and older, suggesting that the floods had a short-term effect on the educational outcomes of rural individuals. This also implies that government relief efforts may have mitigated the floods' impact. Consequently, we will further investigate the government's post-flood relief actions in the mechanism section.

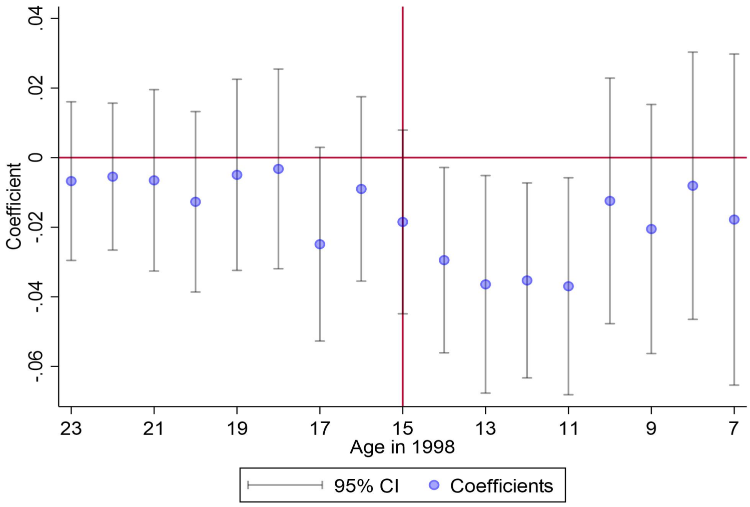


Figure 3. Parallel trend test

5.3 Robustness check

In this section, we further test the robustness of the empirical results through placebo tests, substitution of the core explanatory variable, exclusion of other disaster effects, and an evaluation of the validity of precipitation anomalies as a proxy for flood shocks.

5.3.1 Impact on urban individuals

The negative impacts of natural disasters are particularly severe in less developed regions (World Bank and United Nations, 2010). China's vast rural areas are especially vulnerable to natural disasters due to poor infrastructure and inadequate emergency response capacity. If the deprivation of educational opportunities is caused by flooding, we might expect the impact of flood shocks to be less severe in urban areas. To test this hypothesis, we estimate equation (1) using an urban sample. The regression results, presented in Table 3, show that flooding does not significantly affect urban residents' access to education, which is not surprising given the better infrastructure and relief efforts in cities compared to rural areas.

Table 3. Impact of 1998 floods on education attainment of urban individuals

	Whether to continue high school studies		
Variables	(1)	(2)	(3)
Affected cohort (7 - 15 in 1998) × Flood intensity	0.010	0.010	-0.002
	(0.012)	(0.012)	(0.010)
Observations	82,034	82,034	82,001
R-squared	0.143	0.143	0.145
Controls	NO	YES	YES
Base education × Cohort trend	NO	NO	YES
County FE	YES	YES	YES
Province by cohort FE	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE and Province by cohort FE. Columns (2)-Column (3) sequentially incorporate individual control variables (including gender, and ethnicity) and baseline education level by county with cohort trend interaction terms. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

5.3.2 Placebo test and replacement of precipitation anomalies time period

To further verify the robustness of our empirical results, we create two fictional cohorts. First, we select a sample of individuals aged 25-28 in 1998, treating the 25-26-year-olds as the placebo-treated group and the 27-28-year-olds as the placebo-controlled group. Similarly, we advance the birth year for a sample aged 3-6 in 1998, designating the 3-4-year-olds as the placebo-treated group and the 5-6-year-olds as the placebo-controlled group. As shown in columns (1) and (2) of Table 4, the regression results do not present a significant effect.

In column (3) of Table 4, we replace the flood intensity variable with average daily precipitation anomaly data in 1997 (calculated as the average daily precipitation by county in 1997 minus the average daily precipitation over the past 40 years) , and the regression results remain insignificant. These findings support the robustness of our baseline regression.

Given that the 1998 floods were primarily driven by summer precipitation anomalies, in column (4) of Table 3, we substitute the flood intensity variable with the summer precipitation anomaly for 1998 (calculated by subtracting the county's average daily precipitation for June-August 1998 from the average precipitation over the past 40 years during the same period). The regression results show a slight reduction in effect size but still indicate a significant negative impact.

Table 4. Placebo test and Replacement of flood variable

	Whether to continue high school studies			
	Placebo test			Replacement of precipitation anomaly period
	25 - 26 vs 27 - 28	3 - 4 vs 5 - 6	Fictitious floods occur	6 - 8 precipitation
VARIABLES	(1)	(2)	(3)	(4)
Affected cohort × Flood intensity	-0.008	-0.005	-0.004	-0.004*
	(0.006)	(0.018)	(0.012)	(0.002)
Observations	44,503	28,439	176,132	176,723
R-squared	0.079	0.188	0.169	0.169
Controls	YES	YES	YES	YES
Base education × Cohort trend	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. The treatment group for column (1) is the 25-26 year olds in 1998 and the control group is the 27-28 year olds in 1998. The treatment group for column (2) is the 3-4 year-olds in 1998 and the control group is the 5-6 year-olds in 1998. Column (3) replaces the 1998 precipitation anomaly with 1997 data. Column (4) uses the 1998 summer precipitation anomaly as a proxy variable for flood shocks. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

5.3.3 Excluding other confounding factors

We further rule out the effects of other confounding factors in our results. Among the numerous disasters that occurred from 1998 to 2010, the SARS epidemic (Chae et al., 2020) in 2002 and the Wenchuan earthquake (Feng et al., 2016) in 2008 had nationwide impacts that could have affected educational opportunities for rural individuals. The 2002 SARS epidemic gradually evolved into a nationwide crisis, leading to the closure of primary and secondary schools in China. Since the Chinese government only publishes infection data by province, we included province-fixed effects in our baseline regression to account for the provincial-level impact of SARS. To further isolate the effects of SARS, we excluded samples from Guangdong, Beijing, Shanxi, Inner Mongolia, and Hebei, as these five provinces accounted for nearly 90% of total SARS infections[6]. The results in column (1) of Table 4 indicate that the regression outcomes do not change significantly after removing samples from these provinces.

Another severe disaster nationwide was the 2008 Wenchuan earthquake in Sichuan province, which affected over 40 million people[7] and was the most destructive earthquake in China's history. We identified the counties impacted by the Wenchuan earthquake from government data[8] and Liu and Xu (2021), adding an interaction term between the dummy for affected counties and the dummy for affected cohorts to control for the earthquake's impact. The regression results, shown in column (2) of Table 4, indicate no significant effects.

In addition to natural disasters, changes in education policy also affect our estimates.

[6] See official document in https://www.gov.cn/test/2005-06/28/content_10710.htm
[7] See official document in https://www.gov.cn/jrzg/2008-05/31/content_1001011.htm
[8] See official document in https://www.gov.cn/jrzg/2008-07/22/content_1053017.htm

One education policy that had a significant impact on rural education in China during the period 1998-2010 was the policy of "primary school merger program", which was initiated in 2001 to reduce the financial burden on localities and to improve the layout of rural schools, which were scattered and widely dispersed. In the year the policy was implemented, there were 60,000 fewer elementary school nationwide. As the policy progressed, many local governments took the opportunity to reduce their investment in education, making it difficult for rural students to go to school near their homes, and in 2006, the Chinese government slowed down the process of the program. In order to exclude the impact of this policy on the estimation results, we refer to the current common practice of using the number of elementary school in each county in 2006 minus the number of elementary school in each county in 2001 divided by the number of elementary school in each county in 2001 as the intensity variable of the policy in each county to cross-multiply with the dummy variable of the affected cohort and add the interaction term to the baseline regression, and the results of the regression are shown in Table5, Column (3). Flood shocks remain robust.

Column (4) simultaneously controls for the effects of both the 2002 SARS epidemic and the 2008 Wenchuan earthquake and "primary school merger program", and the regression results still remain robust.

Table 5. Estimation results excluding other confounding factors

	SARS	Wenchuan Earthquake	Primary school merger program	Three of them
VARIABLES	(1)	(2)	(3)	(4)
Affected cohort × Flood intensity	-0.015*	-0.015*	-0.038***	-0.037**
	(0.009)	(0.009)	(0.014)	(0.015)
Affected cohort × Earthquake affected counties		0.004		0.009
		(0.04)		(0.056)
Affected cohort × Primary school merger intensity			-0.017	-0.012
			(0.030)	(0.034)
Observations	163,009	176,132	72,976	66,010
R-squared	0.167	0.169	0.170	0.168
Controls	YES	YES	YES	YES
Base education × Cohort trend	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Column (1) excludes the sample of provinces more severely affected by SARS. Column (2) controls for the impact of the Wenchuan earthquake. Column (3) controls for the impact of the primary school merger program. Column (4) controls for the effects of three of them. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

5.3.4 Validity of Flood Impact Intensity Measurements

5.3.4.1 Distance of counties from rivers

One possible concern is that precipitation anomalies may not adequately capture the extent of flood impacts across locations. To rule out this concern, we used ArcGIS to calculate the distance of each county from the nearest river and cross-multiplied it with the flood impact and age cohort interaction terms. We expect a positive regression result for this interaction, as the 1998 floods had two main causes: inland flooding and flooding from dyke breaches in the Yangtze, Songhua, and Nenjiang river basins. Most of the severely affected areas were a result of dyke breaches. Thus, if precipitation anomalies accurately reflect flood impact, individuals in areas farther from the river should experience less flooding.

The regression results are displayed in Table 6. In the first column, the regression for the entire sample shows that the distance interaction term is not significant but exhibits a positive effect. This is primarily because most flood breaches in 1998 occurred in the Yangtze River Basin. In column (2), when we restrict the sample to provinces in the Yangtze River Basin, the negative effect of flood shocks on educational outcomes diminishes as the distance from the river increases. The results remain significant even when we include samples from the Songhua and Nenjiang River basins in column (3). These results further validate our use of precipitation anomalies as a proxy variable for flooding.

Table 6. The role of river distance on flood impact

	Whether to continue high school studies		
	All	Yangtze river basin provinces	Provinces in the Yangtze, Songhua and Nenjiang river basins
VARIABLES	(1)	(2)	(3)
Affected cohort × Flood intensity	-0.021**	-0.023**	-0.025**
	(0.010)	(0.011)	(0.011)
Affected cohort × Flood intensity×River distance	0.003	0.006**	0.005**
	(0.002)	(0.003)	(0.003)
Observations	176,132	56,572	76,581
R-squared	0.169	0.184	0.185
Controls	YES	YES	YES
Base education × Cohort trend	YES	YES	YES
County FE	YES	YES	YES
Province by cohort FE	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Column (1) shows the regression results for all regional samples. Column (2) is restricted to provinces in the Yangtze River basin (we only include provinces through which the main stream of the Yangtze River flows and where flood damage has been reported, including Sichuan, Chongqing, Hunan, Hubei, Jiangxi, Anhui, and Jiangsu provinces). Column (3) adds the Songhua River and Nenjiang River provinces (including Heilongjiang, Jilin, and Inner Mongolia Autonomous Region) to the Yangtze River basin. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

5.3.4.2 Replacement of flood intensity proxy

To further test the reliability of our baseline regression results, we reconstructed the flood shock intensity proxy variable using the Global Active Archive of Large Flood Events published by the University of Colorado[9]. This archive documents large flood events that have occurred worldwide from 1985 to the present, providing information on the latitude and longitude of the flood-affected locations and the causes of the floods. In total, 26 locations in China are recorded as having experienced large flooding in 1998. Notably, 1998 had the highest number of flood locations in China from 1985 to 2021, and all recorded floods during that year were due to heavy precipitation.

Using the latitude and longitude data from the archive, we created a dummy variable indicating whether a county was flood-affected. Specifically, we identified the 25 counties where the 26 flooding events occurred and included the 184 adjacent counties as the treatment group, while designating the remaining counties as the control group. Figure A2 illustrates the treatment group information, showing that the primary locations of the 1998 floods in China largely coincide with the distribution of precipitation anomalies previously mentioned, primarily clustered in the Yangtze River Basin and the Songhua River Basin. Using this new flood shock variable, we re-estimated equation (1), with the regression results displayed in Table 7. In column (1) of Table 7, the regression includes only the counties directly affected by the flood, revealing that the probability of an individual in a flood-affected county continuing to attend high school decreases by 6.8 percentage points. Column (2) expands this to include the surrounding counties, with results indicating a similar decline in high school attendance probability, reduced by 3.1 percentage points.

Table 7. Replacement of flood shock proxy variable regression results

	Whether to continue high school studies	
	The treatment group contains only the counties where the flooding occurred	Including neighboring counties
VARIABLES	(1)	(2)
Affected cohort × Affected counties	-0.068*	-0.031*
	(0.040)	(0.017)
Observations	176,132	176,132
R-squared	0.169	0.169
Controls	YES	YES
Base education × Cohort trend	YES	YES
County FE	YES	YES
Province by cohort FE	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Column (1) shows the results of a regression that includes only the county where the flood occurred as the treatment group, and column (2) includes the counties surrounding the county where the flood occurred as the treatment group as well. ***, **, * indicates significance at the 1%, 5%, 10% levels.

[9] See official document in http://floodobservatory.colorado.edu/Archives/

5.3.5 Other tests

To account for the possibility that some rural children start school even later, we adjust the age of schooling at each stage in Table A2 to test the sensitivity of our cohort age settings. We also conducted robustness tests for setting of different precipitation anomaly, with the relevant results presented in Table A3. All robustness tests are not far from the results of the benchmark regression, these findings further bolster our confidence in the conclusions of this paper.

5.4 Heterogeneity

Table 8 presents the results regarding heterogeneity. We examine the varying impacts of flood shocks on the educational attainment of rural individuals from two comprehensive perspectives: individual and household.

At the individual level, we investigate gender differences in the impacts of the 1998 floods. As shown in column (1) of Table 8, the probability of rural males continuing high school is notably higher than that of females following the flood shock. This disparity is closely linked to the entrenched patriarchal ideology prevalent in rural China, where sons are often viewed as the future support for their parents in old age and the continuation of the family line. As a result, rural families tend to prioritize the education of their sons, thereby exacerbating gender inequality in education post-flood. This finding aligns with the observations made by Paudel and Ryu (2018), who reported that in Nepal, males exhibited better educational outcomes than females after an earthquake in a patriarchal society.

At the household level, we use household id from the 2010 census data to match children's information with their parents' information, exploring how parental education influences rural individuals' exposure to flood shocks. Columns (2), (3), and (4) of Table 8 reveal that rural individuals with more educated parents experienced a less adverse impact on their education following the flood shock, with a particularly pronounced effect related to the father's education. This suggests that the flood shock has weakened intergenerational mobility in education, while higher parental education mitigated the negative effects of flood hazards on their children's educational outcomes.

The heterogeneity results indicate that the impact of floods in rural areas extends beyond simply reducing individuals' years of schooling; it also amplifies gender inequality and hinders intergenerational educational mobility.

Table 8. Heterogeneity

	Whether to continue high school studies			
	Gender	Family education status		
		Mother's education	Father's education	Parent's education
VARIABLES	(1)	(2)	(3)	(4)
Affected cohort× Flood intensity	-0.031***	-0.174***	-0.301***	-0.285***
	(0.01)	(0.022)	(0.030)	(0.032)
Affected cohort × Flood intensity × gender	0.029***			
	(0.005)			
Affected cohort × Flood intensity × Mother's education		0.024***		
		(0.003)		
Affected cohort × Flood intensity ×			0.035***	

Father's education				
			(0.003)	
Affected cohort × Flood intensity × Parent's education				0.037***
				(0.004)
Observations	176,132	99,193	94,505	105,070
R-squared	0.168	0.170	0.172	0.173
Controls	YES	YES	YES	YES
Base education × Cohort trend	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Column (1) focuses on gender differences in flood impacts. Columns (2)-(3) focus on differences in flood impacts due to parents' respective educational attainment, where parents' respective educational attainment is the number of years of education of the individual's parents, respectively. Parental education in column (3) is the average of the father's and mother's years of education. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

6. Mechanisms

In this section, we explore how the 1998 floods in China hindered rural residents' educational opportunities through two primary mechanisms: the deprivation effect and the negative externalities effect of migration. Meanwhile, we explain why the flood's impact on rural education is short-term. On one hand, the floods directly impacted agricultural operations, resulting in reduced financial resources for rural households and negatively affecting individuals' education. On the other hand, in response to the floods, many families sought non-farm employment opportunities, leading to an increase in children being left behind, which further affected their education. What's more, without outside intervention, the effects of the flood could have been long-lasting. We looked at the government's relief measures after the flood to give an explanation for the short-term impact of this flood on education. Therefore, We reveal the potential mechanisms through which flood shocks impacted education, emphasizing the critical roles of government support and parental support in building human capital in rural areas.

6.1 Deprivation effects of the flood

The 1998 floods significantly disrupted agricultural production in China, escalating the economic burden on rural households. Since free compulsory education in China ends after junior high school, many families faced with financial pressure post-disaster may have opted to forgo senior high school education. We assess the economic burden of the 1998 floods on rural households using microdata from rural fixed observation points covering the period from 1995 to 1998[10]. This timeframe is selected due to significant policy changes initiated in 1999, when the Chinese government began a large-scale policy to return farmlands to forests in response to the floods. As the specific pilot areas for this policy are not publicly available, we cannot account for its effects in our analysis. We estimate the following standard difference-in-differences model:

$$Burden_{i,c,p,t} = \alpha_0 + \alpha_1 Flood_{c,p} \times post(t = 1998) + \alpha_2 X_{i,c,p,t} + \lambda_c + \gamma_{t,p} + \varepsilon_{i,c,p,t} \quad (7)$$

[10] The Rural Fixed Observation Point data counts micro-farmers at the end of the year, so the 1998 data captures the impact of flood shocks on farm households.

where $Burden_{i,c,p,t}$ are indicators related to agricultural production, agricultural income, debt status, and expenditures for rural households i in county c of province p in year t. The variable $post(t = 1998)$ represents whether the year is 1998 (1 for yes, o for no). $X_{i,c,p,t}$ are control variables for household characteristics such as years of schooling in the labor force and household size. The remaining variables are consistent with equation (1).

The estimation results for equation (7) are presented in Table 9. Columns (1) and (2) illustrate the floods' negative impact on rural households' agricultural production, revealing significant reductions in year-end seed and grain stocks. Correspondingly, empirical results in Columns (3) and (4) demonstrate a marked decline in rural households' income from grain and seed post-flood. In terms of livelihood stress, Column (5) indicates an increase in living loans taken out by rural households due to the floods, accompanied by higher living expenditures on food and a reduction in cultural and educational spending. This combination of reduced agricultural income and increased debt, alongside essential expenditure pressures, led to decreased investment in children's education. This outcome aligns with previous research by Deuchert and Felfe (2015), which found that economic losses following a disaster reduced parents' investment in their children's education, leading to fewer years of schooling.

Table 9. The impact of the 1998 floods on rural households

	Agricultural production		Agricultural income		Liability	Expenditure	
	Seed stock	Grain stock	Income from grain	Income from seed	Living loan	Purchased grain	Expenditure on culture and education
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post × Flood intensity	-9.667**	-50.72*	-33.38***	-8.600*	251.1**	53.03**	-104.4*
	(3.758)	(27.43)	(12.26)	(4.837)	(125.4)	-22.25	-56.83
Observations	20,528	20,529	20,528	20,528	20,528	20,529	20,529
R-squared	0.118	0.329	0.02	0.37	0.051	0.445	0.111
Control	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES
Province by year FE	YES	YES	YES	YES	YES	YES	YES

Notes: Columns (1)-(2) focus on the impact of floods on household agricultural production, with the dependent variable in column (1) being the year-end seed stock and the dependent variable in column (2) being the year-end grain stock. Columns (3)-(4) focus on the impact of flooding on farm household agricultural income, the dependent variable for column (3) is the farm household income from food production for the year, and the dependent variable for column (4) is the farm household income from seed seedlings for the year. The dependent variable in column (5) is the amount of household living loan. Columns (6)-(7) focus on the impact of floods on agricultural household expenditures, with the dependent variable in column (6) being the farm household's expenditure on food purchases for the year, and the dependent variable in column (7) being the farm household's expenditure on culture and education. All regressions control for County FE, and Province by cohort FE. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

We further explore vulnerable groups affected by flood shocks to illustrate the existence of flood deprivation effects. Given the fragility of agricultural systems, we anticipate that rural residents in areas heavily reliant on agricultural production will be more susceptible to these deprivation effects. We use the development status of primary, secondary and tertiary industries in each region to represent the degree of regional industrial dependence, collecting the shares of outputs from the tertiary, secondary, and primary sectors in total GDP, as well as the shares of employment in these three industries across various regions prior to the 1998 floods. Due to data limitations at the county level for share data, we substitute prefectural-level data for our analysis.

We re-estimated equation (1), and the results are presented in Table 10. Our findings indicate that a higher regional share of the tertiary sector correlates with a diminished negative impact of flood shocks on rural residents' access to education. Specifically, for every 1% increase in the share of employees in the regional tertiary sector, the negative impact on education access is reduced by 0.2 percentage points. A similar trend is observed for the secondary sector, with a 1% increase in the share of secondary sector outputs in total GDP linked to a 0.2 percentage point reduction in the adverse effects of floods on education access. Conversely, the impact of the primary sector is markedly different; a higher share of the primary sector exacerbates the negative influence of flood shocks on education outcomes. Our analysis shows that a 1% increase in both the share of primary sector outputs in total GDP and the share of employees corresponds to a 0.1 percentage point increase in the negative impact of flood shocks on educational access. We also examine the heterogeneity in regional per capita cultivated areas within the primary sector, as indicated in Column (7) of Table 10. Our findings reveal that greater dependence on the primary sector negatively influences education access for those affected by floods. This suggests that areas reliant on the primary sector are more vulnerable to flood impacts due to the inherent weaknesses in agricultural systems, thus confirming the existence of the deprivation effect.

Table 10. The impact of the 1998 floods on rural households

	Whether to continue high school studies						
	Tertiary sector		Secondary sector		Primary sector		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Affected cohort×Flood intensity	-0.019	-0.073***	-0.098***	-0.026	0.013	0.023	-0.002
	(0.032)	(0.023)	(0.033)	(0.02)	(0.018)	(0.022)	(0.013)
Affected cohort × Flood intensity× Tertiary sector GDP share	0.000						
	(0.001)						
Affected cohort × Flood intensity × Tertiary sector workers share		0.002***					
		(0.001)					
Affected cohort × Flood intensity × Secondary sector GDP share			0.002***				
			(0.001)				

Affected cohort × Flood intensity × Secondary sector workers share				0.000			
				(0.001)			
Affected cohort × Flood intensity × Primary sector GDP share					-0.001*		
					(0.001)		
Affected cohort × Flood intensity × Primary sector workers share						-0.001*	
						(0.000)	
Affected cohort × Flood intensity × Cropland area per capita							-0.014*
							(0.008)
Observations	169,634	169,634	169,634	169,634	169,634	169,634	169,634
R-squared	0.167	0.167	0.167	0.167	0.167	0.167	0.167
Controls	YES	YES	YES	YES	YES	YES	YES
Base education × Cohort trend	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Columns (1)-(2) focus on the moderating effect of the tertiary sector on flood shocks, where the Tertiary sector GDP share represents the tertiary sector's share of regional GDP, and the Tertiary sector workers share represents the tertiary sector's share of the region's total labor force. Columns (3)-(4) focus on the moderating effect of the secondary sector on flood shocks, where the Secondary sector GDP share represents the share of the tertiary sector in regional GDP, and the Secondary sector workers share represents the share of employees in the secondary sector in the total regional labor force. Columns (5)-(7) focus on the moderating effect of the primary sector on flood shocks, where Primary sector GDP share represents the share of the primary sector in regional GDP, and Primary sector Workers' share represents the share of the primary sector in the total regional labor force. Cropland area per capita represents the per capita area of cropland in the region. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

6.2 Negative externalities of migration

In response to the impacts of natural disasters on agricultural production and personal property, many farmers have turned to non-agricultural work as a means of coping with these challenges. Prior to the 1980s, farmers faced significant barriers to migrating to cities due to China's household registration (*hukou*) system. However, recognizing the need for development, the Chinese government began to relax hukou restrictions, thus allowing farmers greater opportunities to seek work in urban areas. The reform of the household registration system in 1984 eased these strict mobility limitations, particularly in the 1990s, when urban development demands led to large-scale migration of rural laborers to cities. This suggests the potential post-disaster migratory behaviors of rural residents following the floods. Due to data limitations, we cannot establish a causal relationship between the 1998 floods and population migration directly. However, we provide some preliminary evidence of the impact of the floods on migration patterns using the 2000 Census data, which includes detailed migration information at the municipal level. We estimate the following equation:

$$Migration_{i,k} = \beta_0 + \beta_1 Flood_{k,p} + \beta_2 X_{i,k} + \varepsilon_{i,k} \qquad (5)$$

where $Migration_{i,k}$ is a dummy variable representing whether rural individuals in city k, province p migrated in 1998 and subsequent years. $Flood_{k,p}$ is the precipitation anomaly data on the municipal level in city k, province p in 1998. $X_{i,k}$ indicates individual-level control variables, including gender, and ethnicity. $\varepsilon_{i,k}$ is the random error term.

The regression results are summarized in Table 11. Columns (1), (3), (5), and (7) demonstrate the positive impact of floods on migration for different age groups. For all rural residents, the probability of migration increased by 1.7 percentage points for every 1 mm increase in abnormal precipitation. However, for the treatment cohort (ages 7-15 in 1998), the probability of migration rose by only 0.6 percentage points per 1 mm increase in precipitation anomaly, significantly lower than other cohorts. This discrepancy may be attributed to the fact that most individuals in the treatment cohort were still in compulsory education in 2000, a period during which strict limitations on schooling for migrant children existed in urban areas.

We also examine gender differences in migration behavior due to floods. Our findings reveal that males in older birth cohorts (ages 25-60 in 1998) had a higher probability of migration, while females in younger cohorts (ages 7-15 and 16-24 in 1998) were also likely to migrate. This suggests that school-aged girls were more prone to dropping out of school to seek work away from home following the floods.

Additionally, we analyze the impact of floods on the relocation choices of migrants by estimating the following equation::

$$Gap_i = \beta_0 + \beta_1 Flood_{k,p} + \beta_2 X_{i,k} + \varepsilon_{i,k} \qquad (6)$$

where Gap_i represents the difference between relevant indicators for the destination city to which the individual i moved and the origin city. $Flood_{k,p}$ refers to the abnormal precipitation data of the individual's pre-migration city in 1998, with the remaining variables consistent with equation (5).

The results are detailed in Appendix Table A4. The treatment group did not show a significant preference for migration. However, for other birth cohorts, rural individuals tended to migrate to more economically developed areas with higher wages when they experienced greater flood impacts. Notably, there was no significant preference among these individuals for moving to areas with better educational and healthcare conditions.

These findings indicate that floods prompted rural residents to seek employment away from their hometowns, motivated by the damage to agricultural production. This tendency may further increase the incidence of "left-behind" children in flood-affected areas, leading to a loss of human capital. Therefore, the migratory behavior of farmers appears to carry negative externalities for their children's education.

Table 11. The impact of the 1998 floods on migration

	Whether to migrate							
	All ages		7 - 15 years old in 1998		16 - 24 years old in 1998		25 - 60 years old in 1998	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Flood intensity	0.017***	0.017***	0.006***	0.010***	0.065***	0.095***	0.016***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)
Flood intensity × gender		-0.001		-0.008***		-0.059***		0.016***
		(0.001)		(0.001)		(0.004)		(0.001)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	758,579	758,579	140,097	140,097	94,372	94,372	369,514	369,514
R-squared	0.002	0.002	0.001	0.001	0.015	0.014	0.003	0.003

Notes: The dependent variable indicates whether rural individuals migrated in 1998 and subsequent years. Columns (1)-(2) show data for the full sample. Columns (3)-(4) are samples aged 7-15 in 1998. Columns (5)-(6) are samples aged 16-24 in 1998. Columns (7)-(8) are for samples aged 25-60 in 1998. Standard errors are in parentheses. ***, **, * indicates significance at the 1%, 5%, 10% levels.

In our discussion of the 1998 floods, we have observed variations in migration probabilities across different age cohorts following the disaster. The adverse effects of floods on agricultural production may compel parents in rural households to leave their hometowns in search of work in urban areas, thereby increasing the number of left-behind children. To examine this hypothesis, we utilize 2000 census data to explore the relationship between the floods and the prevalence of left-behind children. The rationale for selecting the 2000 census is that the treatment group's age range in that year was 9 to 17 years, which aligns with the definition of left-behind children in China. We estimate the following equation:

$$Out_i = \beta_0 + \beta_1 Flood_{k,p} + \beta_2 X_{i,k} + \varepsilon_{i,k} \qquad (8)$$

where Out_i is a dummy variable indicating whether the parents of individual i were migrant workers while the other variables are consistent with equation (6). Our findings indicate that more severe flood shocks are associated with a higher likelihood of parents working away from home, either individually or simultaneously, as shown in Table 12. In other words, the 1998 floods exacerbated the emergence of left-behind children in rural areas. The literature extensively discusses the implications of having left-behind children on educational attainment, with substantial evidence suggesting that these children tend to experience poorer educational outcomes due to the lack of parental presence (Zhou et al., 2014; Ye et al., 2011).

Table 12. The impact of the 1998 floods on left-behind children

	Father was out for work	Mother was out for work	Parents were out for work
VARIABLES	(1)	(2)	(3)
Flood intensity	0.047***	0.049***	0.025***
	(0.002)	(0.002)	(0.001)
Control	YES	YES	YES
Observations	140,097	140,097	140,097
R-squared	0.005	0.006	0.003

Notes: The dependent variable in column (1) is whether the father works outside (outside the home = 1). The dependent variable in column (2) is whether the mother works outside (outside the home = 1). The dependent variable in column (3) is whether both parents work outside (outside the home = 1). Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

6.3 Government relief for flood shocks

The recovery of an area from a natural disaster heavily relies on the effectiveness of post-disaster response and relief efforts. Our previous findings indicate that the impact of the 1998 floods on rural education access was short-lived, suggesting the potential role of government support in the aftermath. Consequently, we focus on government investment in education following the floods. Using county-level data from the China Education Expenditure Statistics Yearbook 1995 - 2000, we estimate the following standard difference-in-differences model:

$$Investment_{c,t} = \alpha_0 + \alpha_1 Flood_{c,p} \times post(t \geq 1998) + \lambda_c + \gamma_{t,p} + \varepsilon_{t,c,p} \quad (4)$$

where $Investment_{c,t}$ is the government's per capita education expenditure, the education budget per capita, and the share of education expenditure in county c in year t. $post(t \geq 1998)$ is a dummy variable indicating whether the government investment was after 1998, , while the other variables are consistent with those in equation (1).

The estimation results are presented in Table 13. We find that for every 1 mm increase in abnormal regional precipitation, total per capita investment in education and the per capita education budget increased by 0.6% and 0.4%, respectively. Additionally, the share of regional investment in education rose by 0.58 percentage points. These findings suggest that the Chinese government significantly increased its investment in education in the areas affected by the floods. Given that investment and infrastructure development in education generally do not yield immediate results, this may explain why the negative impact of the floods on rural residents' access to education diminished in subsequent cohorts.

Table 13. Education investment of the government

	Total expenditure on education per capita(log)	Budgeted education expenditure per capita(log)	Percentage of budget allocated to education
VARIABLES	(1)	(2)	(3)
Post × Flood intensity	0.006*	0.004*	0.580**
	(0.003)	(0.002)	(0.283)
Observations	7,819	7,819	7,786
R-squared	0.418	0.420	0.727
County FE	YES	YES	YES
Province by year FE	YES	YES	YES

Notes: The dependent variable in column (1) is the logarithm of Total expenditure on education per capita. The dependent variable in column (2) is the logarithm of budgeted education expenditure per capita. The dependent variable in column (3) is the Percentage of budget allocated to education. All regressions control for County FE, and Province by cohort FE. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

7. Disscussion

Our research is most closely aligned with the work of Huang and Dong (2025). Although this article has provided us with many insights, and we have all observed that flood experiences are detrimental to children's human capital accumulation, our article contributes in two key areas. First, as mentioned in the introduction, in our mechanism analysis, we argue that while declines in household production and investment capacities could be major factors leading to the decline in human capital, we also find a significant negative impact from children being left behind. This suggests that for enhancing children's human capital, parental presence and family education might be more crucial than direct financial investments (such as increased income from parents working away). We also discovered that government post-disaster aid and investment could play an important role in mitigating these negative effects. Second, we believe that in order to scientifically estimate such impacts, considerations of sample selectivity (like migration) and the accuracy of flood data are potentially very critical. We thus briefly discussed the risks of ignoring migrant populations and using a database that only includes major floods in the previous section. To gain a more intuitive understanding of these risks, this section will further examine the estimation biases introduced by these two practices.

Building upon the methodological framework established by Huang and Dong (2025), our study adopts their model specification and sample matching approach. We operationalize flood exposure based on individuals' current residences and birthplaces, while limiting our analytical sample to non-migrant populations. To capture flood shocks, we utilize the 1998 major flood events documented in the DFO flood database as our primary proxy. The empirical results, presented in Table 14 (Columns 1-2), reveal notable differences relative to our baseline findings in Table 7. Specifically, Column (1), which restricts the treatment group to counties directly impacted by floods, yields a statistically insignificant coefficient (t-value $\approx$ 1.65) that nearly doubles in magnitude compared to its counterpart in Table 7. Column (2) expands the control group to encompass counties adjacent to flood-affected areas, producing a coefficient of -0.043. This estimate aligns closely with the -0.051 coefficient reported by Huang and Dong (2025), yet remains 38.7% ((0.043 - 0.031)/0.031) higher than our Table 7, Column (2) estimate. In Column (3), we introduce precipitation anomalies as an alternative proxy for flood intensity. The resulting coefficient of -0.020 exceeds our baseline Table 2 estimate by 33.3% ((0.020 - 0.015)/0.015), further substantiating the upward bias observed across alternative specifications.

The empirical findings underscore that neglecting the representativeness of flood databases or overlooking population migration dynamics can result in substantial overestimation of flood impacts. This upward bias may be attributed to two primary factors. First, existing flood databases predominantly document major flood events, potentially amplifying the perceived nationwide consequences of large-scale flooding incidents. Second, disregarding population mobility introduces systematic bias in cohort comparisons. Specifically, within the control cohort, individuals tend to be older, and those with higher human capital are more inclined to migrate both during and following flood events (Wang et al., 2017). In contrast, individuals in the treatment cohort are generally younger, possessing limited capacity for independent migration decisions except when relocating with their parents. Furthermore, educational and institutional constraints often impose greater restrictions on migration for younger populations. Consequently, the failure to account for migration likely introduces a downward bias in the educational attainment of the control cohort, thereby artificially inflating the estimated impact of flood events. Our study makes meaningful progress in addressing these two potential biases.

Table 14. The bias of ignoring migration and data representativeness

	Whether to continue high school studies		
	The treatment group contains only the counties where the flooding occurred	Including neighboring counties	Precipitation anomalies
VARIABLES	(1)	(2)	(3)
Affected cohort × Affected counties	-0.127	-0.043**	
	(0.091)	(0.022)	
Affected cohort×Flood intensity			-0.020*
			(0.011)
Observations	93,310	93,310	82,194
R-squared	0.167	0.167	0.171
Controls	YES	YES	YES
County-linear trend	YES	YES	YES
County FE	YES	YES	YES
Province by cohort FE	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Column (1) shows the results of a regression that includes only the county where the flood occurred as the treatment group, and column (2) includes the counties surrounding the county where the flood occurred as the treatment group as well. Column (3) shows the results that use precipitation anomalies as flood intensity proxy. ***, **, * indicates significance at the 1%, 5%, 10% levels.

8. Conclusion

The 1998 floods in China resulted in catastrophic losses nationwide, disproportionately affecting vast rural areas due to poor infrastructure and inadequate relief efforts. This paper has examined the long-term impact of early life exposure in this disaster on educational outcomes in rural China, revealing the effect of governmental and individual disaster responses on human capital accumulation.

Using a difference-in-differences methodology and an unique identification design, we narrow the estimation bias, finding that rural children exposed to the 1998 floods were less likely to continue their high school education. Specifically, the probability of attending high school decreases by 1.5 percentage points for every additional 1 mm of abnormal precipitation in the area. Heterogeneity analyses reveal that (i) the negative impact on human capital was particularly pronounced among young female adolescents in rural regions and (ii) rural individuals with better-educated parents demonstrated greater resilience to flood shocks. These findings indicate that the 1998 floods exacerbated barriers to intergenerational educational mobility in rural China and intensified educational inequalities between rural men and women.

We propose two explanations for the negative impact of the floods on human capital in rural areas. On one hand, the floods heightened the economic burden on rural households, creating a deprivation effect that led to decreased investment in education. On the other hand, the floods prompted farmers to migrate in search of work. While the potential increase in income from migration could have positive effects, the negative impact of leaving children behind appears to be more dominant. .

Our research also highlight the role of government in the long period. We find that post-disaster investments from the government mitigated the long-term negative effects of the flooding. This finding provides valuable insights for disaster response strategies in developing countries. However, due to data limitations, the mechanisms discussed do not fully explain the effects of the 1998 floods, and further refinement will be pursued in future studies.

Appendix

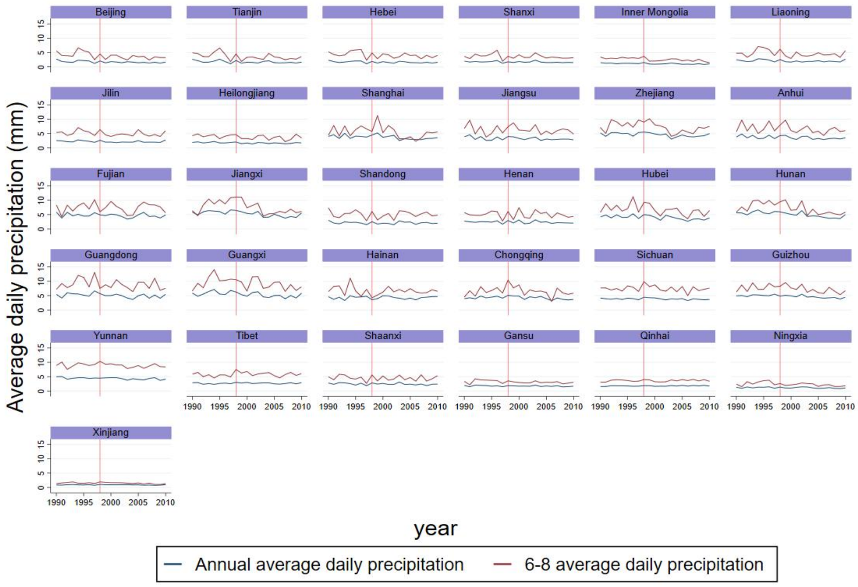


Figure A1. Average daily precipitation in different provinces in China 1990 - 2010

Data source: the European Union and the European Organization of Medium-Range Weather Forecast Centers

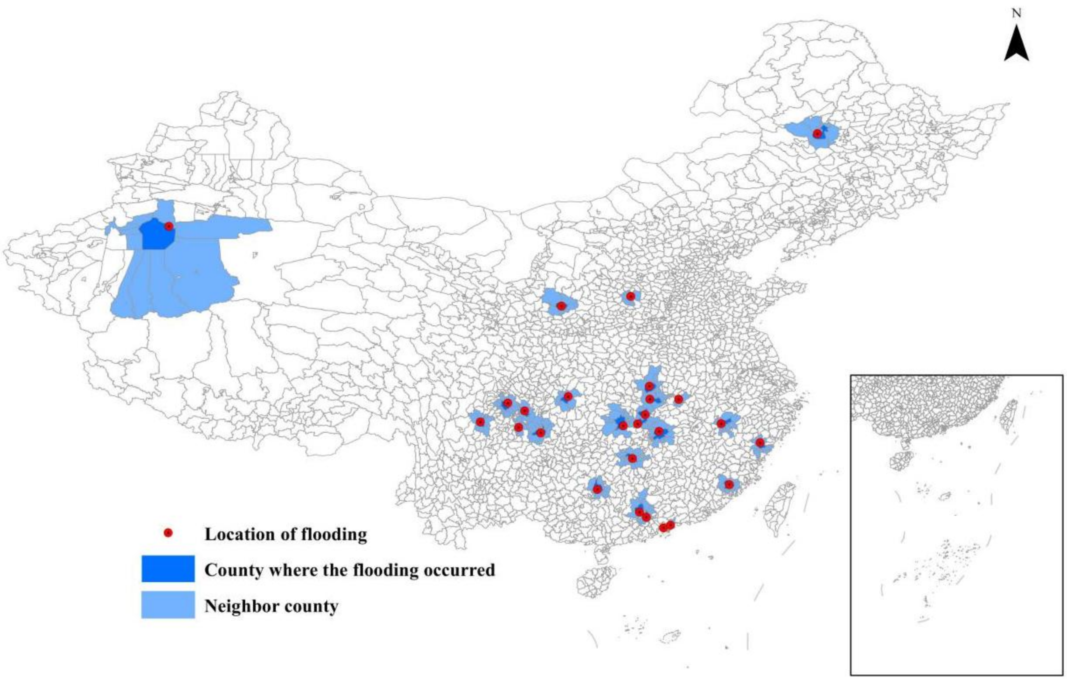


Figure A2. Locations of large floods in China in 1998 as recorded in the Global Active Archive of Large Flood Events

Data source: DFO - Flood Observatory, University of Colorado, USA. http://floodobservatory.colorado.edu/ Archives/

Table A1. Excluding samples that left the place of hukou

VARIABLES	Whether to continue high school studies (1)	(2)	(3)
Affected cohort (7 - 15 in 1998) × Flood intensity	-0.017*	-0.018*	-0.016*
	(0.010)	(0.010)	(0.009)
Observations	160,129	160,129	159,975
R-squared	0.172	0.173	0.173
Controls	NO	YES	YES
Base education × Cohort trend	NO	NO	YES
County FE	YES	YES	YES
Province by cohort FE	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. All regressions control for County FE and Province by cohort FE. Columns (2)-Column (3) sequentially incorporate individual control variables (including gender, and ethnicity) and Base education × Cohort trend. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

Table A2. Cohort age-setting sensitivity

	Whether to continue high school studies			
	Primary School Entry Age +1	Junior high school graduation age +1	Primary School Entry Age +2	Junior high school graduation age +2
VARIABLES	(1)	(2)	(3)	(4)
Affected cohort × Flood intensity	-0.016*	-0.014*	-0.016**	-0.019**
	(0.008)	(0.008)	(0.008)	(0.008)
Observations	166,753	166,753	154,046	154,046
R-squared	0.162	0.162	0.156	0.156
Controls	YES	YES	YES	YES
Base education × Cohort trend	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. The initial cohort ages for the treatment groups in columns (1) and (3) are 8 and 9 years, respectively. The cutoff cohort ages for the column (2) and column (4) treatment groups were 16 and 17 years, respectively. All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

Table A3. Difference between 1998 precipitation and different past years

	Whether to continue high school studies			
	The past 30 years	The past 20 years	The past 10 years	Relative precipitation anomaly
Variables	(1)	(2)	(3)	(4)
Affected cohort (7 - 15 in 1998) × Flood intensity	-0.019**	-0.017*	-0.023**	-0.113**
	(0.009)	(0.009)	(0.010)	(0.051)
Observations	176,132	176,132	176,132	176,132
R-squared	0.169	0.169	0.169	0.169
Controls	NO	YES	YES	YES
Base education × Cohort trend	NO	NO	YES	YES
County FE	YES	YES	YES	YES
Province by cohort FE	YES	YES	YES	YES

Notes: The dependent variable indicates whether the individual has attended high school as of 2010. Column (1) flood intensity proxy uses 1998 average daily precipitation minus the average precipitation for the past 30 years. Column (2) flood intensity proxy uses 1998 average daily precipitation minus the average precipitation of the last 20 years. Column (3) Flood intensity proxy variable uses the 1998 average daily precipitation minus the average precipitation of the last 10 years. Column (4) flood intensity proxy variable uses the difference between the average daily precipitation for 1998 and the average precipitation for the past 40 years divided by the average precipitation for the past 40 years(relative precipitation anomaly indicator). All regressions control for County FE, Province by cohort FE, and Base education × Cohort trend. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

Table A4. Migration options

	All age	7 - 15 years old in 1998	16 - 24 years old in 1998	25 - 60 years old in 1998
	(1)	(2)	(3)	(4)
Panel A: GDP gap between out-migration and in-migration sites				
Flood intensity	74.045***	42.240	58.930***	87.795***
	(12.043)	(37.462)	(16.457)	(20.673)
Control	YES	YES	YES	YES
Observations	26,243	2,146	11,931	10,498
R-squared	0.003	0.002	0.003	0.003
Panel B: Wage gap between out-migration and in-migration sites				
Flood intensity	719.8***	280.9	876.2***	648.2***
	(61.10)	(197.5)	(92.12)	(94.66)
Control	YES	YES	YES	YES
Observations	27,035	2,192	12,267	10,852
R-squared	0.006	0.011	0.008	0.005
Panel C: Educational gap between out-migration and in-migration sites				
Flood intensity	-2.952	-1.331	-12.22***	7.913*
	(2.708)	(8.521)	(3.867)	(4.516)
Control	YES	YES	YES	YES
Observations	27,122	2,196	12,318	10,881
R-squared	0.003	0.006	0.004	0.001
Panel D: Medical condition gap between out-migration and in-migration sites				
Flood intensity	0.573	5.967	-1.940	3.356
	(2.693)	(7.839)	(3.841)	(4.556)
Control	YES	YES	YES	YES
Observations	27,259	2,206	12,367	10,950
R-squared	0.002	0.004	0.003	0.001

Notes: The dependent variable in panel A is the GDP gap between out-migration and in-migration sites. The dependent variable in panel B is the wage gap between out-migration and in-migration sites. The dependent variable in panel C is the gap in the number of middle schools between out-migration and in-migration sites. The dependent variable in panel C is the number of hospitals between out-migration and in-migration sites. Standard errors, in parentheses, are clustered at the county level. ***, **, * indicates significance at the 1%, 5%, 10% levels.

References

[1]. Bui A T, Dungey M, Nguyen C V, et al. The impact of natural disasters on household income, expenditure, poverty and inequality: evidence from Vietnam[J]. Applied Economics, 2014, 46(15): 1751-1766.

[2]. Caruso G D. The legacy of natural disasters: The intergenerational impact of 100 years of disasters in Latin America[J]. Journal of Development Economics, 2017, 127: 209-233.

[3]. Caruso G, Miller S. Long run effects and intergenerational transmission of natural disasters: A case study on the 1970 Ancash Earthquake[J]. Journal of Development Economics, 2015, 117: 134-150.

[4]. Chae M, Liang W, Xue S. Pandemics and Intergenerational Mobility of Education: Evidence from the 2003 Severe Acute Respiratory Syndrome (SARS) Epidemic in China[J]. GLO Discussion Paper, 2020, No. 779.

[5]. Chan K W, Zhang L. The hukou system and rural-urban migration in China: Processes and changes[J]. The China Quarterly, 1999, 160: 818-855.

[6]. Chen Y, Fan Z, Gu X, et al. Arrival of young talent: The send-down movement and rural education in China[J]. American Economic Review, 2020, 110(11): 3393-3430.

[7]. Deuchert E, Felfe C. The tempest: Short-and long-term consequences of a natural disaster for children′'s development[J]. European Economic Review, 2015, 80: 280-294.

[8]. Dong W, Ming Y, Deng Y, et al. Recent wetting trend over Taklamakan and Gobi Desert dominated by internal variability[J]. Nature Communications, 2024, 15(1): 4379.

[9]. Duflo E. Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa[J]. The World Bank Economic Review, 2003, 17(1): 1-25.

[10]. Eskander S M S U, Barbier E B. Adaptation to natural disasters through the agricultural land rental market: evidence from Bangladesh[J]. Land Economics, 2023, 99(1): 141-160.

[11]. Feng S, Lu J, Nolen P, et al. The effect of the Wenchuan earthquake and government aid on rural households[J]. IFPRI book chapters, 2016: 11-34.

[12]. He Daixin,Liu Shiyang. Transfer payments and vanishing soft budget constraints: evidence from the 1998 flood[J]. The Journal of World Economy, 2024, (02):126-151

[13]. Heger, Martin Philipp, and Eric Neumayer. The impact of the Indian Ocean tsunami on Aceh's long-term economic growth[J]. Journal of Development Economics, 2019, 141: 102365.

[14]. Huang Z, Dong X. When the levee breaks: The impact of floods on educational outcomes in China[J]. Journal of Development Economics, 2025, 174: 103450.

[15]. Hydrographic Bureau of the Ministry of Water Resources of China, and Hydrographic Bureau of the Yangtze River Water Conservancy Commission of the Ministry of Water Resources: The torrential rainfall and flooding of the Yangtze River in 1998[M]. China Water Conservancy and Hydropower Press, 2002.

[16]. Jensen R. Agricultural volatility and investments in children[J]. American Economic Review, 2000, 90(2): 399-404.

[17]. Jingzhong Y, Lu P. Differentiated childhoods: impacts of rural labor migration on left-behind children

in China[J]. The Journal of Peasant Studies, 2011, 38(2): 355-377.

[18]. Liu X, Xu Y. Unexpected opportunity for girls: Earthquake, disaster relief and female education in China's poor counties[J]. China Economic Review, 2021, 70: 101701.

[19]. Lu W, Yang P, Zheng S, et al. Natural disasters and high-stakes exam performance: Evidence from the 2008 Wenchuan earthquake[J]. China Economic Review, 2023, 77: 101899.

[20]. Mueller V, Quisumbing A. How resilient are labour markets to natural disasters? The case of the 1998 Bangladesh flood[J]. Journal of Development Studies, 2011, 47(12): 1954-1971.

[21]. Nakamura E, Sigurdsson J, Steinsson J. The gift of moving: Intergenerational consequences of a mobility shock[J]. The Review of Economic Studies, 2022, 89(3): 1557-1592.

[22]. Opper I M, Park R J, Husted L. The effect of natural disasters on human capital in the United States[J]. Nature human behaviour, 2023, 7(9): 1442-1453.

[23]. Pathak P, Schündeln M. Social hierarchies and the allocation of development aid: Evidence from the 2015 earthquake in Nepal[J]. Journal of Public Economics, 2022, 209: 104607.

[24]. Paudel J, Ryu H. Natural disasters and human capital: The case of Nepal's earthquake[J]. World Development, 2018, 111: 1-12.

[25]. Tian X, Gong J, Zhai Z. Natural disasters and human capital accumulation: Evidence from the 1976 Tangshan earthquake[J]. Economics of Education Review, 2022, 90: 102304.

[26]. Van den Berg M. Household income strategies and natural disasters: Dynamic livelihoods in rural Nicaragua[J]. Ecological Economics, 2010, 69(3): 592-602.

[27]. Wang J, Yang J, Li B. Pain of disasters: The educational cost of exogenous shocks evidence from Tangshan Earthquake in 1976[J]. China Economic Review, 2017, 46: 27-49.

[28]. Warr P, Aung L L. Poverty and inequality impact of a natural disaster: Myanmar's 2008 cyclone Nargis[J]. World Development, 2019, 122: 446-461.

[29]. Yu W, Hu J. The long-term effects of natural disasters on human capital accumulation: a quasi-natural experiment based on the Yellow River floodplain area[J]. Environment and Development Economics, 2024: 1-17.

[30]. World Bank, United Nations. Natural hazards, unnatural disasters: the economics of effective prevention[M]. The World Bank, 2010.

[31]. Xing, Chunbing. Educational Expansion, Migration and Urban-Rural Educational Gap--Taking University Expansion as an Example[J]. China Economic Quarterly, 2014, 13(01):207-232.

[32]. Zhang, Zhenyu, and Pan Zhang. The long-term impact of natural disasters on human capital: Evidence from the 1975 Zhumadian flood[J]. International journal of disaster risk reduction 91 (2023): 103671.

[33]. Zhou M, Murphy R, Tao R. Effects of parents' migration on the education of children left behind in rural China[J]. Population and Development Review, 2014, 40(2): 273-292.