# Schooling and Income Effect of Education Development in China: Evidence from the National Compulsory Education Project in China's Poor Areas

Dehua Wang, Jie Zou, Zhonggen Mao*

Whether investing in hardware for basic education is conducive to the long-term development of school-age children is a controversial issue in academic literature. Based on the data from the 2013 Chinese Household Income Project (CHIP), this paper uses the difference-in-differences (DID) method with cross-sectional data to evaluate the policy effects of China's National Compulsory Education Project for Poor Areas launched in 12 central provinces in the mid-1990s. We find that through the construction, renovation, or expansion of primary and junior middle school buildings and purchase of teaching equipment, this project achieved on its goal of "raising education levels": the years of schooling of the beneficiaries after they reached adulthood increased significantly by roughly 0.7 years. However, the project was not found to have an "income effect", i.e., the income of the beneficiary children did not increase with the increase in schoolings. The econometric analysis shows that although the project could raise the income levels of the children in adulthood, it did not significantly increase the likelihood of the beneficiary groups moving to cities as migrant workers. Limited by the late development of secondary and tertiary industries in poor counties, the beneficiaries who stayed locally were less likely to engage in non-farm occupations, and thus their income did not significantly increase. This indicates that policies like the Compulsory Education Project designed to improve school hardware in a short period of time can effectively promote the development of basic education in poverty-stricken areas. Nonetheless, to achieve the fundamental goal of raising income while increasing schooling, it is also necessary to actively guide people moving to cities to work as migrant workers and bolster the development of local secondary and tertiary sectors.

**Keywords:** National Compulsory Education Project, poverty alleviation through education, difference-in-differences with cross-sectional data

* Dehua Wang (email: wangdh@cass.org.cn), Research Fellow, National Academy of Economic Strategy, Chinese Academy of Social Sciences (CASS); Jie Zou (corresponding author, email: zouj18@mails.tsinghua.edu.cn), PhD candidate, School of Economics and Management, Tsinghua University; Zhonggen Mao (email: maogen@swufe.edu.cn), West China Economic Research Center, Southwestern University of Finance and Economics. Fund project: National Natural Science Fund of China (71773139), CASS Innovation Program (2018CJY01-A002), and China Research Institute of Finance and Taxation of Tsinghua University. Our thanks go to Fangwen Lu, attendees at the 4th Camphor Economist Circle (2017, Jinan University), 2017 Young Economist Society (Northeast University of Finance and Economics), 4th China Public Finance Forum (2018, Wuhan University), Tsinghua University China Economy Seminar (2018), Public Policy Quantitative Evaluation Theory and Application Symposium (2019, Xiamen University) and the authors take sole responsibility for this paper.

## I. Introduction

Over the past 70 years, China has spared no effort in developing education, especially basic education, in order to improve human capital, promote economic development and reduce poverty. Educational development has always been a basic national policy. Overall, China has made remarkable progress in this regard: the average years of education for the population aged 15 and above were merely 1.6 in 1949 (Zhai, 2009) and rose to 9.11 in 2011 (Department of Population and Employment Statistics of National Bureau of Statistics, 2013). As Li *et al*. (2017) concludes, the significantly improved human capital is an important reason for the boost in labor productivity since the launch of reform and opening up in 1978, which in turn has driven rapid economic growth.

China has a large population that needs access to compulsory education but lacks funds and proper educational infrastructure due to the relatively backward economic development; this is typical for a resource-constrained country aspiring to develop basic education (Zhai, 2009). How to retrofit shabby school buildings, increase teaching equipment, and rapidly improve the hardware of primary and secondary schools has been a long-standing challenge. In particular, given the regional development imbalances, the Chinese government has rolled out special support policies and leveraged external forces to improve school buildings, teaching equipment and other aspects of school hardware to propel the development of basic education in poverty-stricken areas. This type of policies designed for specific regions is referred to as "big push" policies in the field of social infrastructure by Atolia *et al*. (2017)[1] and as "fight against poverty in education" in this paper.[2] Such policies can have immediate effects such as more spacious and brighter classrooms and more teaching equipment and also lead to improvements in current educational indicators such as increased enrollment and reduced dropout. However, in the long run, how will the improved hardware influence the core indicators such as the increase in human capital and income of individual students? This is a topic rarely addressed in the existing literature.

Based on data from the 2013 Chinese Household Income Project (CHIP), this paper examines how the National Compulsory Education Project for Poor Areas ("Compulsory Education Project" for short), which was implemented in Category-

[1] The "big push" theory, proposed by British development economist Paul N. Rosenstein-Rodan in 1943, thinks that large investments should be made in various sectors concurrently in underdeveloped areas to promote balanced growth of these sectors, thereby promoting the rapid and comprehensive development of the entire economy. Atolia *et al.* (2017) focuses on comparing the advantages and disadvantages of "big push" policies in the fields of economic infrastructure and social infrastructure.

[2] A more general term is "poverty alleviation through education", i.e., taking various measures to develop education to help the poor get out of poverty. This paper proposes the concept of "fight against poverty in education", which refers to dedicated policies supporting the development of basic education in poverty-stricken areas. Obviously, the latter is part of the former.

II regions between the end of 1995 and the end of 1997, influenced the years of education ("schooling effect") and income ("income effect") of the children after they reached adulthood. In 1995, against the lagging basic education in the central and western regions and the difficulty to basically achieve universal coverage of the nine-year compulsory education and eradicate illiteracy among the middle-aged and young people on schedule, the Chinese government launched the Compulsory Education Project designed to improve the hardware of primary and junior middle schools in 852 poor counties of 22 central and western provincial administrative regions. It is China's largest basic education project for poverty alleviation with the widest coverage and most special investment from the central government and a typical policy for "fight against poverty in education". Whether individuals could benefit from the project depends on whether they received primary or junior middle school education in the project counties and whether they were at primary or junior middle school when the project was implemented. By matching the counties in the CHIP with the project counties in Category-II regions and combining the age information of the surveyed individuals in 1995, we used the difference-in-differences (DID) method with cross-sectional data to identify the project's long-term impacts on the increase in human capital and income of the children after they reached adulthood.

The empirical results show that the this project delivered on its goal of "raising education levels": the years of education of the participating children after reaching adulthood increased significantly by roughly 0.7 years. However, the project had quite limited effect on "raising income levels", i.e., the income of the participating children after they grew up was not noticeably higher. Econometric analysis shows that although the project could raise the income levels of the children in adulthood when they worked as migrant workers, the probability of them working as migrant workers did not increase, and the chance of staying in hometown doing non-agricultural work even dropped considerably, perhaps because of the lagging secondary and tertiary industries. In other words, policies like the Compulsory Education Project designed to improve the basic education hardware in poor areas in a short period of time have the schooling effect of improving human capital; however, to achieve the objective of raising income levels, it is also necessary to actively guide people to work as migrant workers and bolster the development of local secondary and tertiary industries.

The contributions of this paper mainly include two aspects.

First, the impact of hardware on student development is a highly controversial topic in educational economics and development economics. By evaluating a major and distinctive project aimed at improving the hardware of primary and secondary schools in impoverished areas, this paper provides empirical evidence from China for this topic. Based on the analysis of more than 600,000 students in the United States, the well-known Coleman Report (Coleman *et al.*, 1966) suggests that resources such as hardware have little effect on improving students' academic attainment and long-

term performance. Since then, a series of studies revolving this topic have been carried out in the fields of pedagogy and economics. While earlier research neglects the endogenous problem of hardware investment, some more recent studies by scholars such as Cellini *et al*. (2010), Neilson and Zimmerman (2014), Hong and Zimmer (2016), and Martorell *et al*. (2016) began to look at the exogenous impact of whether U.S. municipal bonds for education were approved. These studies focused on the impact of hardware construction in U.S. urban areas on student performance in the short to medium term. Although the relevant literature adopts a policy evaluation method with a greater emphasis on causality, there is lack of consensus in their conclusions.

Projects similar to the Compulsory Education Project include the school construction project launched by the Julius Rosenwald Fund to help rural black children in South America from the 1910s to the 1930s ("Rosenwald project") and the INPRES project for primary school development in rural impoverished areas implemented by Indonesia in the 1970s. Aaronson and Mazumder (2011) found that the enrollment rate, literacy rate, and years of schooling of rural black participants in the Rosenwald project grew significantly. Research by Duflo (2001, 2004) showed that the years of education and income of the children benefiting from the INPRES project increased by 1.5% and 2.7%, but this came at the cost of lower wages for older children not covered by the project. As can be seen, research based on historical events of school hardware investment in rural areas has drawn more consistent conclusions.

The Compulsory Education Project is unique in terms of organization and implementation. The Rosenwald project and the INPRES project both aimed to build more schools to improve access to education in addition to improving the quality of school hardware. By contrast, the Compulsory Education Project, implemented in combination with the policy of adjusting the layout of primary and secondary schools, in effect saw a decrease in the number of teaching sites and schools in the project counties instead of increasing access to education. For example, the number of primary schools declined from 125872 to 118524. Therefore, compared to the other two projects, it laid greater emphasis on school hardware construction.[1] Moreover, neither the Rosenwald nor INPRES project clearly designated the benefiting and non-benefiting areas. Duflo (2001, 2004) and Aaronson and Mazumder (2011) defined the treatment group and the control group by school density in each area. By contrast, funds of the Compulsory Education Project are earmarked for the project counties, which clearly distinguish the benefiting areas (project counties) and non-benefiting areas (non-project counties). These two characteristics of the Compulsory Education Project make it suitable to be used to evaluate the impact of school hardware construction. Different from the conclusions made by Duflo (2001, 2004) and other

[1] Despite the criticism for canceling teaching sites and merging schools in recent years, the decrease in the number of schools and the increase in average school size are considered achievements in the summary reports on the Compulsory Education Project.

studies, this paper finds that the income effect of the project is not significant; in view of this, a quantitative decomposition method is used to preliminarily discuss the reasons based on the actual conditions in China, leading to findings that carry policy implications.

Second, this paper evaluates the actual performance of China's major fiscal expenditures and special transfer payment projects from the micro-level perspective of beneficiaries. From the angle of public economics, performance evaluation of fiscal expenditures should be centered on their substantial impact on social production and people's wellbeing. For example, education expenditures should be evaluated from the perspective of human capital and labor productivity of the educated. In addition, the Compulsory Education Project was a large special transfer payment from a higher government to governments at lower levels. This paper also studies the substantial performance of large special transfer payments at the individual level.

The structure of this paper is as follows: the second part introduces relevant information about the project and puts forward the hypothesis; the third part describes the identification strategy, data, and preliminary descriptive statistical results; the fourth part is the empirical analysis of the project's schooling effect; the fifth part conducts an empirical analysis of the income effect and further examines the effect by using quantitative decomposition; and the last part discusses the policy implications of the empirical results.

## 2. Research Background and Hypothesis

Unbalanced regional development poses a major challenge for China to develop basic education and popularize nine-year compulsory education. In 1995, the Chinese government divided the country into three categories according to the levels of education and economic development and put forward different requirements for the popularization of nine-year compulsory education in different regions.[1] According to the analysis by Zhu Kaixuan, then director-general of China's National Education Commission, the most outstanding problem in the popularization of compulsory education facing poverty-stricken areas was the shortage of funding; without strong support from the central and local governments, these backward areas would have

[1] Category-I regions refer to 9 economically developed provinces and municipalities including Beijing, Tianjin, Shanghai, Liaoning, Jilin, Jiangsu, Zhejiang, Shandong, and Guangdong; Category-II regions refer to 12 provinces with moderate levels of economic development and key to the popularization of nine-year compulsory education, including Hebei, Shanxi, Heilongjiang, Anhui, Fujian, Jiangxi, Henan, Hubei, Hunan, Hainan, Sichuan, and Shaanxi; Category-III regions cover 9 economically underdeveloped provincial-level regions that pose the greatest difficulties in the popularization of compulsory education, including Inner Mongolia, Guangxi, Guizhou, Yunnan, Tibet, Gansu, Qinghai, Ningxia, and Xinjiang. See: Guidance on Categorization in the Popularization of Nine-year Compulsory Education, *China Education Yearbook 1996*, 133-144.

great difficulty in achieving the goal of universal access to nine-year compulsory education.[1] At the same time, these regions had great development potential, and achieving the goal on schedule would have strategic significance for reducing regional development gaps.

The Compulsory Education Project is the widest-ranging special project that supports the development of education in specific poverty-stricken areas with the largest amount of funding from the central government since 1949. Based on the project management methods and experience of international organizations such as the World Bank, this project was carried out in 852 counties in Category-II and Category-III regions, including 592 state-level and 260 provincial-level poor counties.[2] The total investments reached RMB 12.756 billion, of which RMB 3.9 billion came from the central government, RMB 6.27 billion from local governments, and RMB 2.585 billion from urban and rural surcharges and non-fiscal funds, accounting for about 2% of China's GDP in 1995.[3] All the funds were steered towards the project counties.

Given that two phases of the Compulsory Education Project were implemented in Category-III regions, in order to get clean estimates of the effects of the project, this paper looks at the project counties in Category-II regions where only the first phase of the project was carried out. The project in these regions lasted from the end 1995 to the end of 1997, and covered 383 counties a total population of 159 million in 13 provinces (municipalities) in Central China, including 262 state-level and 121 provincial-level poor counties. The total investments included RMB 1.5 billion from the central government, RMB 2.938 billion from provincial, municipal, and county-level governments, and RMB 2.131 billion from non-fiscal sources.[4] The central special funds were allocated to the provincial-level governments, which then allocated the funds to lower-level governments. According to the documents on the allocation of central and provincial funds for the Compulsory Education Project (Finance Department of Ministry of Education and Education and Culture Department of Ministry of Finance, 2002), the funds were generally allocated based on such factors as farmers' per capita net income, fiscal revenue, size of poor population, and whether a county is an ethnic county, with priority given to ethnic minority areas, old

---

[1] Zhu, K.X. (1996). A Major Task in Accelerating the Popularization of Nine-Year Compulsory Education. *Ethnic Education of China* (*Zhongguo Minzu Jiaoyu*), 3, 3-4.

[2] *Analysis on Comprehensive Benefits of the National Compulsory Education Project for Poor Areas*. Beijing: Ocean Press, 4-5.

[3] Finance Department of Ministry of Education and Education and Culture Department of Ministry of Finance. (2002). *Analysis on Comprehensive Benefits of the National Compulsory Education Project for Poor Areas*. Beijing: Ocean Press, 304. In 1995, China posted GDP of RMB 6134 billion (National Bureau of Statistics).

[4] *Analysis on Comprehensive Benefits of the National Compulsory Education Project for Poor Areas*. Beijing: Ocean Press, 298. In 1997, Chongqing was separated from Sichuan Province and became a municipality directly under the administration of the central government, so the number of provincial regions covered by the project increased to 13.

revolutionary base areas, border areas, and large grain-producing counties.

The funds for the Compulsory Education Project in Category-II regions were mainly used for the construction of hardware for primary and junior middle schools such as the expansion and renovation of school buildings and purchase of teaching equipment, desks and chairs, with few spent on the purchase of books and principal and teacher training. Generally, more funds went to primary schools than junior middle schools in order to "satisfy the needs of primary schools first".[1] For example, the total funds of the project earmarked for Hebei Province were about RMB 490 million, of which 51.8% was used for the construction or renovation of primary schools, 21% for that of junior middle schools, 20% to purchase teaching equipment and desks, 4.4% to purchase books, and less than 4% for training principals and teachers.[2] Given that poor infrastructure such as school buildings was a major obstacle to the popularization of nine-year compulsory education in economically underdeveloped areas, such spending arrangements were in alignment with the development status of basic education in Category-II regions at that time.

At the end of 1998, the Ministry of Education and the Ministry of Finance evaluated the implementation of the Compulsory Education Project in each project county. According to the evaluation conclusions (2002), the project "greatly improved the conditions of primary and secondary schools in project counties", "increased access to compulsory education and accelerated the progress in 'two basics'." In terms of objective indicators such as school buildings, enrollment rate, and drop-out rate, the project helped promote educational development in the project counties. However, in economic literature, human capital measured by years of education and the future income growth of the educated are generally used as long-term performance indicators for education. By contrast, the evaluation results by the government (2002) are only about short-term impacts of the project through comparison with the past situations of the project counties, presenting insufficient evidence showing the project's effects on improving human capital. So, can the better hardware brought by the Compulsory Education Project in poverty-stricken counties in Category-II regions truly improve long-term educational performance?

## 3. Data and Evaluation Strategy

The data used in this paper is mainly rural household survey data from the 2013

[1] Announcement of Ministry of Education and Ministry of Finance on the Completion of the National Compulsory Education Project for Poor Areas in Category-II Regions. http://old.moe.gov.cn//publicfiles/business/htmlfiles/moe/moe_355/200409/3844.html.

[2] Finance Department of National Education Commission and Education and Culture Department of Ministry of Finance. (1997). *Guide on the Management of the National Compulsory Education Project for Poor Areas*. Beijing: Higher Education Press, 241-242. The rates for Hebei are calculated based on relevant information, and the situations of other provincial regions are similar.

China Household Income Project (CHIP). The main evaluation strategy is to use difference-in-differences (DID) with cross-sectional data.

The CHIP survey data is one of the most authoritative household survey data in China and has been widely used in research literature.[1] The CHIP has the following characteristics: it can meet the needs of mid- and long-term evaluations, incorporates long-term migrant workers, even if some have changed the nature of their household registration,[2] and contains information such as educational background, occupation, and income of all members in a rural household, as well as the administrative division code of the county of the surveyed households. By comparing these counties with the list of project counties in the Compulsory Education Project, it is easy to identify the beneficiaries.

Similar to other studies (Duflo, 2001; Aaronson and Mazumder, 2011), this paper uses DID with cross-sectional data. To use this method, the area where a child benefiting from the project is located needs to be accurately identified so as to define the treatment group and the control group. This paper matches the counties covered by the Compulsory Education Project with those of the surveyed rural households in CHIP, the lattering totaling 201 in the 2013 CHIP survey. After matching, it is found that 19 of them were project counties in Category-II regions, 10 in Category-III regions, and the remaining 172 were not covered by the project. As mentioned earlier, the 10 counties in Category-III regions are excluded. Therefore, the 19 project counties in Category-II regions are used as the treatment group, and 172 non-project counties as the control group.

For the DID method with cross-sectional data, dummy variables similar to those in ordinary DID before and after policy implementation also need to be constructed. Whether an individual could benefit from the project depends on their age. The Compulsory Education Project covered primary and junior middle schools, where students are generally aged 7 to 15 in China. Therefore, individuals who were aged 6 or below in 1995 were considered as the fully benefiting age group, while those aged 16 or above in 1995 had completed junior middle school education before the project was launched and therefore considered as non-benefiting age group. Like the method used by Duflo (2001), this paper sets a dummy variable—*young*, which is similar to ordinary DID before and after policy implementation. Its value is 1 when an individual was aged 2 to 6 in 1995, and 0 when they were aged 16 to 20.[3]

Therefore, a DID model with cross-sectional data can be set up as (1):

---

[1] For the basic information on the CHIP data, please visit: http://ciid.bnu.edu.cn/CHIP/index.asp.

[2] According to the CHIP guide, rural households refer to households where the head has a rural *hukou* registered in the town (neighborhood) in which they are living; household members refer to all people that live in the same house or share all expenses and income. The samples include non-permanent population, including migrant workers, those going to school in other places, and people who have changed the nature of their household registration (from rural to urban). See: Wu (2007).

[3] The general practice is to ensure the age differences of the treatment group and the control group are the same. Therefore, the benefiting age group in this paper is 2-6 years of age instead of 0-6.

$$y_{ij} = a_j + \beta young_i + \lambda program_j \times young_i + X_{ij}\delta + \varepsilon_{ij} \quad (1)$$

In model (1), $y_{ij}$ represents the years of education or income of individual $i$ whose *hukou* (or residence) is in county $j$; project $j$ and young $i$ are dummy variables of whether the individual belongs to the specific project county and the benefiting age group, respectively. The coefficient $\lambda$ of the interaction term of the two dummy variables represents the effect of the project. $X_{ij}$ is a county-level characteristic variable. In order to control the unknown regional factors that may affect the years of education or income, the fixed effect $\alpha_j$ at the county level is controlled in model (1). At this point, project $j$ does not need to join the regression.

The ex-ante parallel trend between the treatment group and the control group is an important premise for applying the DID technique. However, the project counties of the Compulsory Education Project were not selected randomly: most of them were economically and socially underdeveloped. Given their systematic socioeconomic gap with non-project counties, there is a possibility that even if the Compulsory Education Project was not implemented, there may still be systematic differences in the years of education and income levels of residents between them. This paper uses two methods to mitigate this endogenous problem.

First, while controlling the fixed effect at the county level, some interaction terms between county-level regional variables in 1990 and whether they were 2~6 years old are added to control the ex-ante trend.[1] This approach essentially allows the ex-ante characteristic variables of each county to have varying effects on different age groups. We checked the factor variables the provincial-level governments considered when allocating the central government's grants for the Compulsory Education Project, such as farmers' per capita net income, per capita fiscal revenue, population density, and whether a county is an ethnic county, to determine the ex-ante county-level regional variables.[2] If the project was still found to have significant schooling and income effects after the ex-ante trend is controlled, the results will be highly robust.

Second, a placebo test is performed on the ex-ante parallel trend to verify whether there are significant differences in the increase of years of education or income of individuals in project and non-project counties who received compulsory education prior to the project. Similar to the method used by Duflo (2001), individuals aged 16 to 20 in 1995 are redefined as the treatment group and those aged 21 to 25 as the control group, and DID analysis with

[1] Moser and Voena (2009) provided the reasons and principle for using this method. To control the ex-ante trend, county-level variables of the year before the project was implemented should be used. In view of data availability, county-level regional variables in 1990 are selected in this paper.

[2] The Chinese government (2002) introduced the factors on which the provinces allocated funds for the project. Among them, the population density is mainly used to measure the topographic features. The terrain in the central region is increasing, the mountainous areas are sparsely populated, and the plains are more populated.

cross-sectional data is carried out. Since both groups had completed compulsory education by 1995, they were definitely not the beneficiaries of the Compulsory Education Project. Therefore, this is a placebo test.[1] If it is found that there are no systematic differences in years of education and income between the two age groups in the project and non-project counties, this can be used as evidence of a parallel trend.[2]

In addition, the decomposition method is used to examine why the project was found to have no income effect and analyze the impacts of the project on the children's probability of working as migrant workers and engaging in non-agricultural work, so as to explore the deep reason why the project could barely raise income levels.

The CHIP survey data records the years of education after an individual reached adulthood, which can be directly used to measure the level of education. In addition, this paper uses other information in CHIP to calculate the per capita income from agricultural operations and by adding it to income from wages or non-agricultural operations to get the total income of an individual.

---

[1] Like the studies by Duflo (2001) and others, to find an age group that is not affected by the policy to conduct a placebo test is a common practice for parallel trend tests of DID with cross-sectional data. A reviewer argues that dividing into two groups alone is not sufficient to test the parallel trend and suggests that the 16~25 age bracket in 1995 be divided into multiple groups for verification. Following this suggestion, we divide the 16~25 age group into 10 groups for parallel trend tests of the schooling effect (using the 25-year-olds as the control group); it is found that while the interaction term coefficients is significantly positive for the 16-year-old group, the coefficients are not significant for the 17- to 24-year-old groups. The significance of the interaction term coefficient for the 16-year-old group can be explained by the fact that some children in rural areas went to school late. Duflo (2001) also gives yearly interactive term coefficient images of samples aged 0 to 24. The main purpose is to analyze the marginal effect of the benefit period, but objectively it can also provide evidence for the parallel trend test. We also replicate the approach of Duflo (2001) and get similar images. These two auxiliary tests can further verify the validity of the parallel trend test in this paper. Given that it is common to divide the 16-to-25 age group into two groups for the parallel trend test to be clearly contrasted with the principal component regression, this paper uses it as the main method of the parallel trend test. The results of the two parallel trend tests are not included herein due to length limitations. If you are interested, feel free to contact us.

[2] In the econometric analysis of this paper, family background information such as parents' occupations or income is not controlled. This is consistent with the approach used in long-term policy impact assessments by Duflo (2001). The reason is that the model needs to control family background information when the policy was implemented, i.e., in 1995 or before, but micro-level survey data generally only provides family background information at the time of the survey, and there are few retrospective surveys on family background information at long intervals. In the CHIP data, the level of education of parents can be matched by using the household code and the "relationship with the head of household" variable. If the level of education did not change much in 2013 compared with 1995, it could be used as a control variable. We successfully matched only 1201 individuals in this regard. Econometric analysis finds that after controlling parents' level of education and county-level fixed effects, interaction term coefficient of whether individuals were aged 2 to 6 and whether they were in project counties is 0.855, which is significantly positive at the 0.1 level. If the ex-ante trend is further controlled, the number of samples is reduced to 978, and the interaction term coefficient is 0.583, which is significant at the 0.1 level. We believe that the decline in significance here is mainly correlated with a notable decrease in the sample size. As DID with cross-sectional data is used herein, the systemic differences in family backgrounds by region and age group have been absorbed by other variables and will have little impact on the results if not controlled, in addition to the sample size limitations. Therefore, family background information is not controlled in all the analyses herein.

## 4. Schooling Effect of the Compulsory Education Project

Table 1(A) shows the basic regression results of the schooling effect obtained based on the econometric model (1). The explained variables in the table are years of education. The "project counties" dummy variable is added to column (1), which is a standard DID model with cross-sectional data. Column (2) controls county-level fixed effects to solve the problem of missing regional variables. In column (3), interaction terms between county-level characteristic variables such as net income of farmers, per capita fiscal revenue, and per capita industrial output in 1990, and whether an individual was aged 2 to 6 in 1990 are added to control the ex-ante trend. On the basis of column (3), columns (4) and (5) include interaction terms such as population density, whether a county was an ethnic county, and whether an individual was aged 2 to 6; the control variables in column (6) are the same as those in column (5), but the control group is restricted to Category-II regions. In all columns, the interaction term coefficients (*young* × *program*) of whether individuals were aged 2 to 6 and whether they were in project counties are significantly positive at the 0.05 significance level. These results show that the project has a significant schooling effect. The regression coefficients of the interaction terms in each column are relatively stable. The economic implication is that the project increased the average years of education of the beneficiaries by about 0.7 years.

Table 1. Basic Regression Results of the Schooling Effect

| | Years of education | | | | | |
|---|---|---|---|---|---|---|
| | Table 1(A) 2~6 years old/16~20 years old | | | | | |
| | (1) | (2) | (3) | (4) | (5) | Category-II regions |
| Interaction term | 0.693*** (3.23) | 0.753*** (3.51) | 0.694*** (2.86) | 0.679*** (2.79) | 0.625** (2.54) | 0.778*** (2.93) |
| Aged 2~6 | 2.402*** (29.00) | 2.495*** (31.42) | 2.747*** (12.46) | 2.720*** (12.35) | 2.788*** (12.23) | 2.662*** (7.14) |
| Project county | –0.936*** (-5.89) | | | | | |
| County-level fixed effect | No | Yes | Yes | Yes | Yes | Yes |
| Ex-ante trend interaction term | No | No | Yes | Yes | Yes | Yes |
| Constant term | 8.560*** (143.64) | 9.000*** (19.64) | 7.524*** (6.5) | 9.175*** (9.11) | 9.211*** (9.17) | 9.250*** (6.45) |
| N | 5200 | 5200 | 4107 | 4107 | 4043 | 2619 |
| $R^2$ | 0.1653 | 0.3188 | 0.3193 | 0.3214 | 0.3171 | 0.2725 |

Years of education

Table 1(B) 16~20 years old/21~25 years old

| | (1) | (2) | (3) | (4) | (5) | Category-II regions |
|---|---|---|---|---|---|---|
| Interaction term | –0.198<br>(–0.94) | 0.002<br>(0.01) | 0.196<br>(0.86) | 0.199<br>(0.87) | 0.257<br>(1.11) | 0.316<br>(1.25) |
| Aged 16~20 | 0.775***<br>(10.33) | 0.722***<br>(10.28) | 0.404**<br>(2.04) | 0.414**<br>(2.09) | 0.330<br>(1.64) | 0.035<br>(0.10) |
| Project county | –0.738***<br>(–5.35) | | | | | |
| County-level fixed effect | No | Yes | Yes | Yes | Yes | Yes |
| Ex-ante trend interaction term | No | No | Yes | Yes | Yes | Yes |
| N | 5047 | 5047 | 3995 | 3995 | 3940 | 2484 |
| $R^2$ | 0.0356 | 0.2217 | 0.1986 | 0.1991 | 0.1881 | 0.1372 |

Notes: (1) The values in parentheses are the *t*-statistics calculated by "robust" standard errors; (2) ***, **, and * indicate that the regression coefficients pass the significance tests at the 0.01, 0.05, and 0.1 levels, respectively.

Each column in Table 1(B) corresponds to that in Table 1(A). The control variables are the same, while the treatment group and the control group are replaced with the 16~20 and 21~25 age groups who had completed compulsory education before the project to conduct the placebo test. It can be seen that although the interaction term coefficients in Table 1(B) are mostly positive, they are much smaller than those in Table 1(A), and none of them is significant. This indicates that the parallel trend assumption is reasonable: despite the systematic differences between project and non-project counties, there will be no significant difference in the years of education between different age groups without the intervention of the Compulsory Education Project.

To examine the marginal schooling effect on the children benefiting from the project, the dummy variable "aged 2~6" in model (1) is replaced with the years that individuals aged 2~20 benefited from the project. The results are shown in Table 2. As can be seen, the interaction term coefficients of the benefit years and project counties are significantly positive and pass the significance test at the 0.01 level.[1] When the schooling that the children benefited from the project increased by one year, the average years of education after adulthood increased by about 0.07 to 0.09 years. If the samples are limited to Category-II regions, the coefficients will be a bit greater.

[1] Children aged 2 to 6 in 1995 benefited 9 years from the project; for children aged 10 or 15, the benefit years are 6 or 1, respectively.

Table 2. Impact of Benefit Years on Years of Education after Adulthood

| | Full sample | | | Category-II regions | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| Interaction term | 0.0907*** (3.53) | 0.0874*** (3.49) | 0.0741*** (2.58) | 0.098*** (3.63) | 0.094*** (3.56) | 0.084*** (2.65) |
| Benefit year | 0.283*** (29.13) | 0.294*** (31.99) | 0.335*** (12.59) | 0.276*** (22.00) | 0.288*** (23.86) | 0.334*** (7.07) |
| Project county | −1.215*** (-6.40) | | | −0.875*** (−4.40) | | |
| County-level fixed effect | No | Yes | Yes | No | Yes | Yes |
| Ex-ante trend interaction term | No | No | Yes | No | No | Yes |
| N | 9649 | 9649 | 7444 | 6191 | 6191 | 4648 |
| $R^2$ | 0.0914 | 0.2510 | 0.2452 | 0.095 | 0.205 | 0.194 |

Notes: the values in parentheses are the *t*-statistics calculated by "robust" standard errors; *** indicates that the regression coefficients pass the significance tests at the 0.01 level.

## 5. Income Effect of the Compulsory Education Project and Quantitative Decomposition

### *5.1. Basic Results*

Children aged 2 to 6 in 1995 were 20 to 24 years old when the CHIP was conducted in 2013, whose employment was generally still unstable. Children aged 7 to 10 in 1995 in the project counties also benefited from the project and were aged 25 to 28 in 2013, when they should be stable forces in the labor market. In view of this, in the following analysis, the 7~10 age group in 1995 is defined as the benefiting age group, while the non-benefiting age group is still the 16~20 age group.[1] The benchmark regression results of the income effect are shown in Table 3. As can be seen from the first four columns in Table 3, no matter whether the logarithmic value of the explained variable "income" includes income from agricultural operations or whether the samples are restricted to Category-II regions, the interaction term coefficients are not significant. In other words, there is no indication that the project had an income effect. Given that per capita household expenditures are highly correlated with household income and are unlikely to be underreported, the last two columns in Table 3 will be replaced by

[1] We also conducted relevant analysis by using the 2~6 age group, and the basic conclusions are the same as in Table 5.

explanatory variables with household expenditures. It can be seen from the table that the interaction term coefficients are still not significant, that is, from the perspective of household consumption, the project was not found to have the effect of raising income levels.

Table 3. Benchmark Regression Results of the Income Effect on the 7~10 Age Group

| | Income (excluding income from agricultural operations) | | Total income | | Household expenditures | |
|---|---|---|---|---|---|---|
| | Full sample | Category-II regions | Full sample | Category-II regions | Full sample | Category-II regions |
| Interaction term | 0.096<br>(1.10) | 0.0869<br>(0.93) | 0.109<br>(1.26) | 0.104<br>(1.13) | –0.057<br>(–1.04) | –0.0258<br>(–0.44) |
| Project county | –0.028<br>(–0.43) | –0.0748<br>(–0.63) | 0.060<br>(0.87) | –0.00446<br>(–0.04) | 0.177***<br>(3.36) | 0.0417<br>(0.45) |
| County-level fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Ex-ante trend interaction term | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 2583 | 1529 | 2449 | 1470 | 3308 | 2075 |
| $R^2$ | 0.1825 | 0.184 | 0.2156 | 0.231 | 0.3095 | 0.208 |

Notes: the values in parentheses are the *t*-statistics calculated by "robust" standard errors; *** indicates that the regression coefficients pass the significance tests at the 0.01 level.

Wu (2007) points out that rural residents with higher education and income levels are more likely to change their household registration (*hukou*) from "rural" to "urban", implicating a high degree of upward selectivity. The above econometric analysis finds that the project had no income effect, partly because some individuals who benefited from the project changed their *hukou* to "urban" in adulthood, resulting in a downward bias for the income effect. The CHIP rural household samples used in this paper include those who worked as migrant workers or changed their *hukou* from "rural" to "urban". Rural residents with higher levels of education are more likely to change their *hukou*, but this should have quite limited impact on the income effect evaluation results hereinof.

### *5.2. Decomposition Analysis*

Why did the project increase the years of education but play no role in raising the income levels of the beneficiaries as adults? To understand this is of great significance for poverty alleviation through education. In this paper, the quantitative decomposition method is used from the perspective of the income structure and employment

characteristics of the beneficiaries to shed light on this issue.

Generally, there are two main employment options for rural residents: to work as migrant workers or stay in their hometown engaging in agricultural or non-agricultural operations. The former means competing in the national labor market, while the latter means being restricted by the local labor market. These two groups of people vary greatly in terms of income structure and employment characteristics and therefore should be analyzed separately. According to the definition in the CHIP questionnaire, the sample individuals in Table 4(A) are people who left their counties to work as migrant workers in 2013. The regression results show that the interaction terms (aged 7 to 10 years project counties) are significantly positive at the 0.05 and 0.1 levels, and the income effect of the project should be about 20%. The samples of the treatment group and control group in Table 4(B) are those who did not work as migrant workers in 2013. The regression results show that whether the total income or income from non-agricultural operations is used as the only factor, the interaction term coefficients are negative but not significant.

Table 4. Regression Results of the Income Effect across Groups

| | (A) Income of migrant workers | | | | (B) Income of non-migrant workers | | | |
|---|---|---|---|---|---|---|---|---|
| | Income (excluding income from agricultural operations) | | Total income | | Income (excluding income from agricultural operations) | | Total income | |
| | Full sample | Category-II regions | Full sample | Category-II regions | Full sample | Category-II regions | Full sample | Category-II regions |
| Interaction term | 0.195** (2.06) | 0.187* (1.90) | 0.207** (2.35) | 0.194** (2.10) | –0.159 (–0.44) | –0.191 (–0.46) | –0.239 (–0.63) | –0.308 (–0.74) |
| Aged 7~10 | –0.053 (–0.64) | –0.131 (–0.87) | –0.056 (–0.70) | –0.103 (–0.71) | 0.024 (0.18) | 0.001 (0.00) | 0.030 (0.24) | 0.085 (0.23) |
| County-level fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Ex-ante trend interaction term | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1724 | 1147 | 1714 | 1141 | 834 | 564 | 825 | 525 |
| $R^2$ | 0.187 | 0.182 | 0.176 | 0.167 | 0.346 | 0.334 | 0.343 | 0.356 |

Note: the values in parentheses are the *t*-statistics calculated by "robust" standard errors; ** and * indicate that the regression coefficients pass the significance tests at the 0.05 and 0.1 levels, respectively.

By comparing the results of Tables 4(A) and 4(B), it can be found that as long as the beneficiaries left their hometown and participated in the national labor market,

more years of education indeed increased their income. However, for those who stayed in their hometown, the increase in years of education did not lead to higher income, indicating that the project was not found to have an income effect. Then, did the project increase the probability of working outside their hometown? Table 5 sets the dummy variable "whether to leave hometown to work or do business" as an explained variable. The results show that the interaction term coefficients (aged 7 to 10 × project county) are not significant no matter whether the samples are restricted to Category-II regions. This implies that compared with the control group, the beneficiaries did not choose to work outside their hometown even though they received more years of education.

Table 5. Impact on the Probability of Going out for Non-Farm Jobs

| | Table 5(A) Full sample | | | Table 5(B) Category-II regions | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| Interaction term | –0.024<br>(–0.85) | 0.009<br>(0.30) | 0.040<br>(1.20) | –0.0138<br>(–0.46) | 0.0329<br>(1.08) | 0.0430<br>(1.23) |
| Project country | 0.127***<br>(8.69) | 0.107***<br>(7.92) | 0.019<br>(0.50) | 0.116***<br>(6.63) | 0.0830***<br>(4.96) | 0.0213<br>(0.35) |
| Item | 0.196***<br>(8.13) | | | 0.121***<br>(4.78) | | |
| County-level fixed effect | No | Yes | Yes | No | Yes | Yes |
| Ex-ante trend interaction term | No | No | Yes | No | No | Yes |
| N | 4484 | 4484 | 3449 | 2838 | 2838 | 2133 |
| $R^2$ | 0.0401 | 0.2747 | 0.2487 | 0.0350 | 0.214 | 0.162 |

Notes: (1) The values in parentheses are the *t*-statistics calculated by "robust" standard errors; (2) ***, **, and * indicate that the regression coefficients pass the significance tests at the 0.01, 0.05, and 0.1 levels, respectively.

We then analyze the employment choices of those who stayed in their hometown and were surveyed by the CHIP questionnaire. Non-agricultural occupations include occupations with wage income and non-agricultural businesses. Table 6 sets three dummy variables: "whether engaged in non-agricultural occupations", "whether engaged in non-agricultural occupations with wage income", and "whether engaged in non-agricultural production" as explanatory variables. The results show that the interaction term coefficients in the first four columns are all significantly negative at least at the 0.1 significance level. Although the interaction term coefficients in the last two columns are negative, they are not significant. This indicates that if the beneficiaries stayed in their hometown, the probability of engaging in non-agricultural occupations would be reduced. The main reason is that the opportunities for non-

agricultural occupations with wage income decreased, while whether they were employed in non-agricultural businesses had little impact.

Table 6. Impact on the Probability of Being Employed in Local Non-Agricultural Occupations

| | Whether engaged in non-agricultural occupations | | Whether engaged in non-agricultural occupations with wage income | | Whether engaged in non-agricultural production | |
|---|---|---|---|---|---|---|
| | Full sample | Category-II regions | Full sample | Category-II regions | Full sample | Category-II regions |
| Interaction term | –0.217** (–2.71) | –0.138* (–1.82) | –0.231*** (–2.79) | –0.176* (–1.88) | –0.110 (–1.23) | –0.013 (–0.13) |
| Aged 7~10 | –0.0463 (–0.79) | –0.195 (–1.61) | –0.00221 (–0.04) | –0.114 (–0.89) | –0.0634 (–0.73) | –0.325* (–1.83) |
| County-level fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Ex-ante trend interaction term | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1769 | 1014 | 1512 | 861 | 976 | 650 |
| $R^2$ | 0.280 | 0.222 | 0.342 | 0.261 | 0.331 | 0.239 |

Notes: (1) The values in parentheses are the *t*-statistics calculated by "robust" standard errors; (2) ***, **, and * indicate that the regression coefficients pass the significance tests at the 0.01, 0.05, and 0.1 levels, respectively.

Zhao (2006) finds that the improvement in the return on education in rural areas mainly depends on the rapid development of the non-agricultural economy. Currently, in rural China, the income of non-agricultural jobs is relatively high, which is the main factor influencing the increase of rural residents' income. As can be seen from Table 6, although the project increased the years of education for the beneficiaries, it did not raise, but even reduced, their opportunities to engage in non-agricultural occupations, probably because of the lagging development of local secondary and tertiary sectors. The decrease in non-agricultural employment opportunities for children staying in their hometown is one of the reasons why the project did not have an income effect. Our findings are consistent with those of Zhao (2006).

As we see it, this can be attributed to the lagging development of non-agricultural sectors in the project counties, as shown in Figure 1 which compares the average added value of the three sectors in the 372 project counties with that in the 1207 non-project counties. As can be seen in Figure 3, the added value of the three sectors in project counties was lower than that of non-project counties. However, the gaps in added value of the secondary and tertiary sectors with non-project counties widened over time. Zhen and Ling (2017) pointed out that the development of non-agricultural sectors had a greater impact on the income of the younger generation (2~6 age group) than the older generation (16~20 age group). After 2000, the secondary and tertiary sectors of

the project counties fell behind, unable to provide sufficient non-agricultural jobs. Such impact was greater on the beneficiaries (2~6 age group) in the project counties. As a result, the project had the effect of increasing the levels of education but not income.[1]

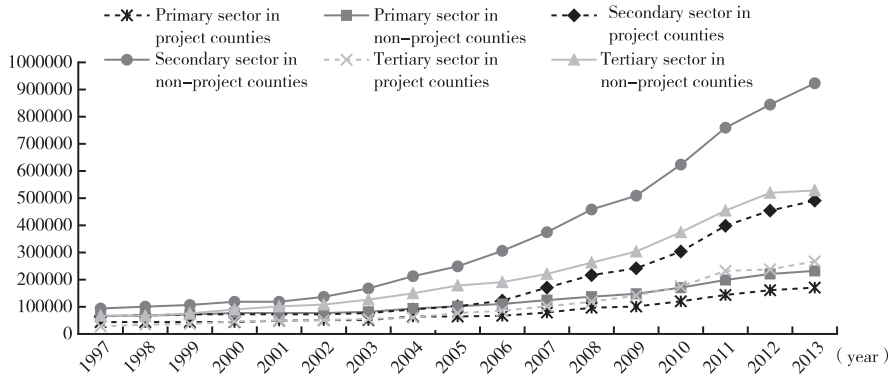


*Figure 1*. Trends of Added Value by Sector in Project Counties and Non-Project Counties (RMB 1000)
Notes: Data on the added value of the three sectors in each county comes from provincial statistical yearbooks and *China Statistical Yearbook for Regional Economy*; the project counties in the figure refer to the 372 counties in Category-II regions where the Compulsory Education Project was implemented, and there were 1207 non-project counties.

## 6. Conclusions and Policy Implications

In order to prop up the development of education in economically underdeveloped areas, developing countries often adopt the "big push" strategy of introducing special support policies and leveraging external forces to rapidly improve school buildings, teaching equipment and other hardware. However, the existing studies pay little attention to the long-term effects of such "big push" development strategy. The Compulsory Education Project in China carried out in the mid-1990s was a typical "big push" in the field of social infrastructure. Using data from the 2013 CHIP rural household survey, this paper analyzes the schooling and income effects of the project in an effort to examine the long-term effects of "big push" strategies, providing empirical evidence from China.

Quantitative analysis finds that the project had a significant schooling effect, increasing the level of education of the beneficiaries by about 0.7 years. However, the project was found to have no effect of raising their income levels. On the face of it, these findings run counter to the classic human capital theory. However, econometric analysis of those leaving their hometown for work finds that the project could indeed increase the income of the beneficiaries, which is consistent with the classic human

[1] Thanks to the reviewers for providing us with this perspective

capital theory. Generally, the project was not found to have the effect of raising income levels because it failed to increase the beneficiaries' probability of going out for work as adults; moreover, those who stayed in their hometown saw a significant decline in the probability of engaging in non-agricultural occupations. The possible reason is that the development of the secondary and tertiary sectors in the project counties was staggering over the years, unable to provide sufficient non-agricultural jobs for local residents. These results indicate that in China, the value of human capital improvements can only be reflected in non-agricultural secondary and tertiary labor markets. The finding that the project had income effect is not at odds with the classic human capital theory; the underlying reason is the lack of a labor market that can give full play to human capital.

This paper has obvious policy implications. Taking the "big push" strategy to promoting the development of education in poverty-stricken areas can have positive effects. However, poverty alleviation through education requires policy coordination to provide more non-agricultural jobs for people in poverty-stricken areas to bring into full play the improved human capital. The measures include actively guiding farmers to work as migrant workers and integrate into the market; promoting the development of the secondary and tertiary sectors in poor areas through industrial policies, combining poverty alleviation through education with industrial development, and providing more non-agricultural jobs for local residents.

## References

Aaronson, D., & Mazumder, B. (2011). The Impact of Rosenwald Schools on Black Achievement. *Journal of Political Economy*, 119 (5), 821-888.

Atolia, M., Li, B., Marto, R., & Melina, G. (2017). Investing in Public Infrastructure; Roads or Schools? IMF Working Papers 17.

Coleman, J., Campbell. E, *et al*. (1966). Equality of Educational Opportunity. Washington, DC: U.S. Government Printing Office.

Duflo, S. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91 (4), 795-813.

Duflo, S. (2004). The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Project in Indonesia. *Journal of Development Economics*, 74 (1), 163-197.

Finance Department of Ministry of Education, & Education and Culture Department of Ministry of Finance. (2002). *Analysis on Comprehensive Benefits of the National*

*Compulsory Education Project for Poor Areas*. Beijing: Ocean Press. (in Chinese)

Finance Department of National Education Commission, & Culture and Education Department of Ministry of Finance. (1997). *Guide on the Management of the National Compulsory Education Project for Poor Areas*. Beijing: Higher Education Press. (in Chinese)

Hong, K., & Zimmer, R. (2016). Does Investing in School Capital Infrastructure Improve Student Achievement? *Economics of Education Review*, 53,143-158.

Li, H. P., Loyalka, P., Rozelle, S., & Wu, B. (2017). Human Capital and China's Future Growth. *Journal of Economic Perspectives*, 31 (1), 25-48.

Martorell, P., Stange, K., & McFarlin, I. (2016). Investing in Schools: Capital Spending, Facility Conditions, and Student Achievement. *Journal of Public Economics*, 140, 13-29.

Moser, P., & Voena, A. (2009). Compulsory Licensing: Evidence from the Trading with the Enemy Act. *American Economic Review*, 102 (1), 396-427.

Neilson, C., & Zimmerman, S. (2014). The Effect of School Construction on Test Scores, School Enrollment, and Home Prices. *Journal of Public Economics*, 120, 18-31.

Rosenstein-Rodan, P. (1943). Problems of Industrialization of Eastern and Southeastern Europe. *Economic Journal*, 53 (210-211), 202-211.

Wu, X. G. (2007). China's Household Registration System and Intergenerational Occupational Mobility. *Sociological Studies* (*Shehuixue Yanjiu*), 6, 38-65.

Zhai, B. (2009). A Miracle in Human History: China's Endeavor to Promote Education for All in 30 Years of Reform and Opening Up. *Educational Research* (*Jiaoyu Yanjiu*), 1, 3-11.

Zhao, L. (2006). Research on Return on Education in Rural China. *Social Sciences in China* (*Zhongguo Shehui Kexue*), 3, 98-109.

Zhen, X., & Ling, C. (2017). Impact of Rural Labor Mobility on Rural Income and Income Gaps: From the Perspective of Labor Heterogeneity. *China Economic Quarterly* (*Jingjixue Jikan*), 16(3), 1073-1096.