Economic Modelling 121 (2023) 106223

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

# Education and Migrant Health in China

Yuanyuan Chen [a], Haining Wang [b,*], Zhiming Cheng [c], Russell Smyth [d]

[a] School of Finance & Investment, Guangdong University of Finance, Guangzhou, Guangdong, China
[b] Center for Chinese Public Administration Research, School of Government, Sun Yat-sen University, Guangzhou, Guangdong, China
[c] Department of Management, Macquarie Business School, Macquarie University, Sydney, New South Wales, Australia
[d] Department of Economics, Monash Business School, Monash University, Melbourne, Victoria, Australia

ARTICLE INFO

*Keywords:*
Education
Migrant
Health
Compulsory education reform
Higher education expansion
China

ABSTRACT

We examine the causal relationship between education and health among migrants using data from the China Migrants Dynamic Survey. Our identification strategy exploits exogenous changes in compulsory schooling laws and expansion in China's higher education sector. We find that an additional year of education increases the mean self-reported health scores of less-educated migrants by 2.6 percentage points; however, no significant effect is found for better-educated migrants. Heterogeneity analysis suggests that this result is driven by the effect of education on health for women and migrants residing in rural areas. We find that for less-educated migrants, the relationship between education and health is mediated by the positive effects of education on health awareness, healthcare utilization, health behaviors, and income. We conclude by emphasizing the implications of our findings for investing in social policies that result in better health outcomes.

## 1. Introduction

Education, a significant component of human capital, plays a crucial role in influencing an individual's health-related behaviors and outcomes (Eide and Showalter, 2011; Hamad et al., 2018; Silles, 2009). For example, more education is associated with an improved diet, regular exercise, and moderate alcohol consumption (Li and Powdthavee, 2015). More education is also linked to a lower likelihood of obesity (Kim, 2016; von Hippel and Lynch, 2014), less psychological distress, better mental health (Fergusson et al., 2015; Zhang et al., 2015), and better health in old age (Lantz et al., 2001; Ross and Wu, 1996; Zajacova and Hummer, 2009).

Despite the seemingly voluminous literature on education and myriad health outcomes, debate remains about whether the relationship is statistically causal and economically meaningful. In a meta-analysis of studies seeking to establish a causal link between education and health outcomes using compulsory schooling laws to identify the relationship, Hamad et al. (2018) find disagreement concerning whether education and health are causally linked. In a more recent meta-analysis of 4866 estimates gleaned from 99 published studies, Xue et al. (2021) find that the overall effect size of education on health is practically zero, implying that education has no discernible benefits for improving health outcomes. This result is consistent with recent findings by Albarrán et al. (2020), who exploit exogenous variation in compulsory schooling induced by school laws across European countries. These authors find that education has no causal effect on three self-reported health outcomes: general health status, long-standing illnesses/health problems, and engaging in limited activities.

We examine the health returns to education among rural–urban migrants using data from the 2017 China Migrants Dynamic Survey (CMDS). We exploit two education policies—the introduction of the Compulsory Education Law (CEL) between 1986 and 1994 and the 1999 higher education expansion—as natural experiments to identify the causal relationship between education and health status for individuals with varying levels of education. Our main estimates suggest that an additional year of education causes a 2.6 percentage point increase in the mean self-reported health scores of less-educated migrants; however, it has no significant effect on the self-reported health scores of better-educated migrants. This result is robust to a wide range of sensitivity tests. Heterogeneity analysis suggests that this result is driven by the effect of education on health for women and migrants residing in rural areas. We find that wages, health awareness, healthcare utilization, and health behaviors mediate the relationship between education and self-reported health for less-educated migrants.

* Corresponding author. Center for Chinese Public Administration Research, School of Government, Sun Yat-sen University, Guangzhou, Guangdong, China.
*E-mail addresses:* 47-198@gduf.edu.cn (Y. Chen), wanghn36@mail.sysu.edu.cn (H. Wang), zhiming.cheng@mq.edu.au (Z. Cheng), russell.smyth@monash.edu (R. Smyth).

https://doi.org/10.1016/j.econmod.2023.106223
Received 21 June 2022; Received in revised form 19 January 2023; Accepted 30 January 2023
Available online 31 January 2023

We focus specifically on the impact of education on health among migrants because migrants represent a vulnerable population, for which there have been recent calls to give greater attention to the relationship between social policies and health behaviors (Burns et al., 2021). Many studies document that migrants have poorer health outcomes than nonmigrants (Bollini and Siem, 1995). Other studies suggest that education can support a broader holistic inclusion strategy for the socioeconomic integration of migrants into the fabric of host communities (Platform and Maletic, 2016). If investing in education improves health outcomes for migrants, this is likely to have spillover effects to better labor market outcomes and socioeconomic integration into host communities.

In many ways, China is a natural country for a study on education and migrant health. First, exogenous changes in its education policies represent quasi-experiments to study the causal effects of education on health outcomes. Changes in China's educational policies have been used as an identification strategy to study the effects of education on a range of other outcomes, including for migrants (Cheng, 2021; Cheng and Smyth, 2021; Huang et al., 2021; Tang et al., 2020; Xiao et al., 2017). Second, China has undergone large-scale internal migration, which is among the largest in the history of humankind (Zhao, 1999). Third, from a social policy perspective, several studies have identified the need to address the gap in health outcomes between migrant and nonmigrant populations in Chinese cities (Chen, 2011; Hu et al., 2008). Poor health outcomes for migrants in China are often exacerbated by the fact that they typically work in "three D″ jobs—jobs that are dirty, dangerous, and demeaning—and lack access to healthcare services available to nonmigrants (Guo et al., 2014; Hu et al., 2008; Nielsen and Smyth, 2008). Providing better education opportunities for migrants provides an important avenue to potentially reduce the health gap between migrants and nonmigrants.

We provide three important contributions to the quasi-experimental literature that has exploited variation in educational policies to rigorously estimate the health effects of education. First, we find evidence that education improves self-reported health for less-educated individuals; this result comes at a time when studies such as Xue et al. (2021) are beginning to question the health effects of education. Second, we extend the literature to the study of education and health for migrants. Third, we increase our understanding of the causal effect of education on self-reported health using variations in educational policies for an important country outside of Europe and the United States. Hamad et al. (2018: 176–177) specifically identified this gap in the literature, stating that "while self-rated health has been studied numerous times in the U.S. and several European countries, it has not been examined at all in other country settings. Since compulsory schooling laws implementation differed in each country, and because educational attainment may have different effects in different historical or political contexts, findings in one country for a particular outcome are not necessarily generalizable to other settings."

Our findings have three important social policy implications. First, they suggest investing in education for less-educated individuals can generate substantial returns to self-reported health later in life. In this respect, the effects of the health returns that we find to investing in education for less-educated migrants are substantial, equivalent to more than a six standard deviation increase in monthly family income per capita. Second, in terms of targeting scarce resources, our results suggest that targeting female migrants and migrants in rural areas leads to the highest returns to investing in education. Third, reducing health inequalities in China has been identified as an important social policy priority (Zhou et al., 2017). Since we do not find any significant health returns to education among migrants with higher education, our findings imply that education could reduce health inequality among migrants.

## 2. Data and descriptive analysis

We use data from the 2017 CMDS, administrated by the National Health and Family Planning Commission of China.[1] The CMDS is a large-scale, nationally representative cross-sectional survey that employs a stratified, multistage sampling design with a probability proportional to size sampling method.

The sampling frame of the 2017 CMDS was the entire population of migrants aged 15 years old or above who did not hold a local household registration (*hukou*) and who had lived in their current residence for more than a month as of May 1, 2017. The 2017 CMDS covered 169,989 migrants across 31 mainland provinces/autonomous regions/municipalities and the Xinjiang Production and Construction Corps (XPCC).[2] The CMDS collected detailed information on migrants' health status, educational attainment, and a wide range of personal and family characteristics.

To rule out the potentially confounding long-run impact of the 1959–1961 Great Chinese Famine and the most intensive period of the 1966–1976 Cultural Revolution on an individual's health status (Cheng et al., 2021; Hayward et al., 2022; Meng and Zhao, 2021), we restrict our analytical sample to individuals born in or after 1970. Our analytical sample contains 138,020 valid respondents. On average, migrants with junior high school education or below obtained 8.17 years of education, while those with senior high school education or above obtained 13.68 years of education. Among the final sample employed, 12.98 percent completed primary school (6 years) or below, 45.40 percent completed junior high school (9 years), 21.68 percent completed senior high school (12 years), and 19.92 percent held higher education qualifications, i.e., a graduate diploma (15 years), bachelor's (16 years), masters, or Ph.D. Table A1 shows that migrants with senior high school education or above are more likely to be single and a member of the Chinese Communist Party, hold a nonagricultural *hukou*, undertake inter-city migration, have a longer migration spell, and be employed as managers and technicians. Moreover, they have relatively higher family economic status, reflected in higher family income and homeownership rate, and a smaller family size than their less-educated counterparts. In contrast, migrants with a junior high school education or below are more likely to be married, hold an agricultural *hukou*, and be engaged in interprovincial migration. Migrants with a junior high school education or below are more likely to be solo or employer entrepreneurs and engage in business, service, and production-related work.

In the 2017 CMDS, respondents were asked to rate their current health status on a four-point Likert scale, ranging from one (very unhealthy) to four (very healthy). Self-reported health (SRH) status is a comprehensive health measure that presents an overall picture of the respondent's physical health, mental health, disease prevalence, and satisfaction with their health status (Xie and Mo, 2014). It has been well documented that the SRH is strongly associated with various objective health indicators and is a good predictor of subsequent functional decline and mortality among adults (Borrell et al., 2004).

Table 1 presents summary statistics on SRH among migrants by the level of education. Those with higher educational attainment report better SRH. Mann–Whitney tests find statistically significant differences in health status between individuals with and without junior high school education and between individuals with senior high school education and those with higher education qualifications. There is a significant difference in health status between migrants with junior and senior high

[1] For a detailed description of the CMDS, see the project website: https://www.chinaldrk.org.cn.

[2] The XPCC was established in 1954 as a paramilitary and economic organization in the Xinjiang autonomous region. The administrative status of XPCC is on par with the provincial capital. The XPCC administrates seven county-level cities and five towns. Its GDP was 252 billion Chinese yuan, and its population was 3.1 million in 2017.

**Table 1**
Summary statistics on health status by the level of education.

| | Number of observations | Self-reported health status | | Mann–Whitney test | |
|---|---|---|---|---|---|
| | | Mean | Std. Dev. | z-statistic | p-value |
| Primary school education or below | 17,920 | 3.72 | 0.55 | −30.569 | 0.0000 |
| Junior high school education | 62,667 | 3.85 | 0.39 | | |
| Senior high school education | 29,929 | 3.87 | 0.35 | −6.213 | 0.0000 |
| College education or above | 27,504 | 3.89 | 0.32 | | |

school education.[3]

## 3. Methods

We estimate the following health production function:

$$H_i = \alpha + \beta_1 E_i + \beta_2 X_i + \beta_3 OP_i + \beta_4 DP_i + \beta_5 C_i + \varepsilon_i \quad (1)$$

where $H_i$ denotes SRH for individual $i$, and $E_i$ is an individual's years of education. $X_i$ is a vector of control variables consisting of personal, family, and regional characteristics that the existing literature suggests are important correlates of health status; $OP_i$ and $DP_i$ denote origin and destination province fixed effects, respectively, and $\varepsilon_i$ is the error term. We also include cohort effects ($C_i$) in all specifications to control for potential cohort-specific differences in health status, such as cohort size and other unobserved factors that vary uniformly across cohorts. In our baseline estimates, we control for the linear cohort trend in the estimation. In robustness checks, we control for more demanding interactions between the birth province, destination province, and the linear cohort trend to allow the cohort-specific effects to vary across origin and destination provinces. Table A1 in the appendix presents the definitions and descriptive statistics for each control variable in Equation (1).

We first employ ordinary least squares (OLS) regression to estimate the relationship between years of education and health status among the subsamples of migrants with junior high school education or below (less-educated) and of those with senior high school education or above (better-educated), respectively. However, educational attainment is not randomly assigned and is correlated with individual, household, and regional characteristics. Although we have included a full set of control variables to mitigate the potential impacts of omitted variables, unobserved heterogeneity, such as cognitive ability and personality traits, may still affect both educational attainment and health status. For instance, an unobserved ability could generate an upward bias in the estimated effect of education on health outcomes if it positively correlates with migrants' schooling years and health status; however, the OLS estimates could also be biased downward if individuals with high discount rates are more likely to choose a relatively lower level of education, which provides a higher marginal rate of return (Belzil and Hansen, 2002). Existing studies suggest that bias from the higher discount rate is more prominent than omitted ability, resulting in OLS underestimating economic returns to education in China (Heckman and Li, 2004).

We employ an instrumental variable (IV) approach to address potential endogeneity by exploiting the CEL implemented between 1986 and 1994 and the 1999 higher education expansion as exogenous changes that affected migrants' educational attainment. Fang et al. (2016) discuss why the CEL is an appropriate IV for educational attainment. At the cohort level, the CEL resulted in different probabilities of completing junior high school education among children aged 16 or below before and after implementing this policy. In contrast, at the province level, the staggered implementation of the CEL across different provinces created a series of natural experiments. In addition, in the 1970s and 1980s, few in China could foresee the significant increase in returns to education since the 1990s, reflecting the development of a free labor market, which replaced a system of state-led job allocation that did not depend primarily on educational attainment (Fang et al., 2016). Thus, parents were unlikely to manipulate their children's participation in compulsory education solely to obtain higher returns to education in a future free labor market, which could not have been readily foreseen in the 1980s. Equally, in the context of the present study, it is improbable that parents would have manipulated their children's participation in compulsory education to benefit future health outcomes.

The higher education expansion policy changed the likelihood of university admission among students who took the university entrance exam before and after 1999. Huang et al. (2022) argue that the 1999 higher education expansion can serve as a natural experiment to identify suitable IVs for educational attainment because it was announced suddenly, without any public consultation, and colleges and universities were only given a few months to prepare for a 47 percent intake surge. The higher education expansion policy has been treated as an exogenous shock in examining the effect of education on socioeconomic development in China (Li et al., 2014; Wang, 2021).

While migrants with junior high school education or below were affected by the CEL differently, they were not influenced by the introduction of the higher education expansion policy since they did not proceed to senior high school education. In contrast, those with senior high school education or above have uniformly completed nine years of compulsory education required by the CEL; thus, they were only affected differently by the higher education expansion policy. Therefore, using these two education policies as natural experiments, we can identify the causal relationship between schooling years and health status for individuals with relatively low and high levels of education.

We estimate the following treatment model to instrument for years of education ($E_i$) in Equation (1):

$$E_i = \alpha + \theta_1 Z_i + \theta_2 X_i + \theta_3 OP_i + \theta_4 DP_i + \theta_5 C_i + \sigma_i \quad (2)$$

where $E_i$ denotes an individual's years of education and $Z_i$ is a vector of instrumental variables. Following previous studies that employ education policies as an identification strategy to draw causal inferences about education on various outcomes (Cheng, 2021; Huang et al., 2021; Wang et al., 2022), we construct two sets of IVs for the lower and higher education subsamples, respectively.

For migrants with a lower level of education, the first IV we employ is the length of exposure to the 1986 CEL. Since the 1980s, China has witnessed rapid economic growth, and the Chinese government has increasingly realized the important role of education in social and economic development. The CEL was promulgated as a national policy on April 12, 1986, specifying that all children aged six must complete at least nine years of compulsory schooling (six-year primary school education and three-year junior high school education). A typical child usually starts school at age 6 and completes their 9-year compulsory education by age 15, when they decide whether to pursue further education or enter the labor market. After the CEL was introduced, the enrolment rate of 7-year-old children increased dramatically from approximately 70 to 96 percent between 1990 and 2000 in rural areas, and the enrolment rate in junior high schools increased from 68 to 98 percent between 1985 and 2003 (Cui et al., 2019). However, substantial geographical and temporal variations existed in implementing the CEL across provinces between 1986 and 1994. Beijing, Zhejiang, Sichuan, Hebei, and Jiangxi were among the first to implement the reform, while Tibet was the last to implement this law. As a result, cohorts born in different years and provinces experienced different levels of CEL exposure.

[3] The Mann–Whitney test results: z-statistic = −14.610, p-value = 0.0000.

This study defines migrants who were aged 16 or older when the CEL was implemented as a nonexposed cohort (exposure = 0), those aged between 7 and 15 as a partially exposed cohort (exposure = 1), and those aged 6 or younger as a fully exposed cohort (exposure = 2). We treat the measure of CEL exposure as a cardinal variable. Prior research suggests that the IV results remain consistent whether the reform exposure measure is treated as ordinal or cardinal (Cui et al., 2019). On average, migrants fully exposed to the CEL obtained 8.54 years of education, while those partially or never exposed to the CEL obtained 7.86 and 7.37 years of education, respectively.

Fig. 1a plots the proportion of those completing junior high school education among migrants with lower levels of education by the length of CEL exposure. We observe significant jumps in the likelihood of junior high school education for migrants exposed to the CEL, and the jump is more visible for fully exposed cohorts. The length of exposure to the CEL positively correlates with the proportion of those completing junior high school education among migrants who were partially exposed to the CEL. Given that the CEL was gradually implemented across different provinces, we further examine the relationship between the length of CEL exposure and the likelihood of having junior high school education by province concerning the different timing of the CEL. Fig. 1b shows that fully exposed cohorts have a considerably higher likelihood of completing junior high school education than cohorts either partially or never exposed to the CEL and that this pattern is consistent across provinces with different timing of the CEL implementation.

The second IV is the quarter of birth, which is widely used in the existing literature (Angrist and Keueger, 1991). In China, children are not allowed to work until 16 years of age, which may affect their schooling decisions. Since the CEL was implemented with different degrees of enforcement strictness and effectiveness across provinces (Li and Zhang, 2017), children who were born in the first three quarters may turn 16 years old before the completion of 9 years of education and, thus, be more likely to leave school to enter the labor market than those who were born in the fourth quarter. We create a dummy variable equal to 1 if a respondent was born in the fourth quarter and 0 otherwise.

For migrants with a higher level of education, we employ three IVs; the first two are based on the fact that urban and rural residents were differentially affected by the higher education expansion in 1999 (Huang and Zhu, 2020). Before the higher education expansion policy, the enrolment rate in universities was less than 3 percent; before 1999, it increased by an average annual rate of 4.7 percent. The Ministry of Education in China suddenly announced a 47 percent increase in university places in early 1999, followed by increases of 38 and 22 percent in 2000 and 2001, respectively, and more modest growth in subsequent years (Huang et al., 2021). In so doing, the Chinese government sought to increase the gross university enrolment rate to 15 percent by 2010 in response to the economic slowdown and rising youth unemployment caused by the 1997 Asian Financial Crisis.

The IVs used in this study are interactions between one's *hukou* status at age 18 (*HK18*: nonagricultural *hukou* = 1, agricultural *hukou* = 0) and each of two variables; namely, a dummy variable indicating whether the respondent was born after 1980 (*Post 1980*) and a birth year linear trend (*Trend*). Theoretically, the first cohort affected by the 1999 higher education expansion was born in 1980. These variables capture heterogeneous effects and cohort trends of the higher education expansion reform on the higher educational attainment of individuals who had an urban or rural *hukou* at 18 years old. The third IV is the linear trend of higher education enrolments between 1988 and 2013, which is matched to each respondent and birth province when an individual was 18 years old and due to take the national university entrance exam. This IV could reflect one's probability of obtaining a higher education qualification conditional on the change in the supply of university admission space. The third IV is similar to those used in other studies that use the changes in admission supply due to higher education expansion across cohorts and provinces as an IV (Fu et al., 2022; Li et al., 2017).

Fig. 2a shows the proportion of having a higher education qualification among those better-educated migrants who were born before and after the higher education expansion. We can see that the likelihood of receiving higher education for both urban and rural *hukou* holders at age 18 shows a significant jump following the higher education expansion policy, and it is more pronounced among those urban *hukou* holders, which is consistent with previous findings based on the 2018 China Family Panel Studies (Wang et al., 2022). Fig. 2b also shows that the number of enrolments in higher education institutions increased dramatically after the expansion, from 35,000 in 1998 to 226,000 in 2013. Thus, our IVs capture the timing and heterogeneous effects of the higher education expansion policy on the educational attainment of migrants with urban and rural origins, which are exogenous to any individual and household circumstances and satisfy the exclusion restriction.

Taking Equation (2) as the first stage of the two-stage least squares (2SLS) model, we estimate the following equation in the second stage:

$$H_i = \alpha + \gamma_1 \widehat{E}_i + \gamma_2 X_i + \gamma_3 OP_i + \gamma_4 DP_i + \gamma_5 C_i + \varepsilon_i \tag{3}$$

where $\widehat{E}_i$ is the fitted value of migrants' schooling from Equation (2), and $\gamma_1$ is the causal effect of education on migrants' health status.

## 4. Main results

We first examine the relationship between education and health for migrants with lower and higher levels of education. Panels A and B in Table 2 present the OLS and IV estimates for the two subsamples, respectively.[4] The OLS results show that having more schooling is positively associated with the health status of migrants with junior high school education or below (model A1); its impact on those with senior high school education or above is insignificant (model B1). Among migrants with a lower level of education, a one standard deviation increase in years of education (1.81) is associated with a 0.078 standard deviation (or 0.034 points) increase in SRH scores. These findings are consistent with significant and decreasing marginal returns to education (Hout, 2012; Kemptner et al., 2011).

The IV estimates from the 2SLS model, which addresses the endogeneity of education, are consistent with the OLS estimates regarding the significance level and direction of coefficients. The first stage results show that the IVs are all significantly correlated with migrants' educational attainment and satisfy the weak instrument and over-identification tests, suggesting their validity as instruments. The IV results in model A2 show that having an additional year of education could increase the SRH of less-educated migrants by 0.10 points, corresponding to 0.231 standard deviations, or a 2.62 percentage point, increase in mean health scores. To put the magnitude of this estimated effect in context, it is equivalent to a 6.13 standard deviation increase in monthly family income per capita.

The size of the effect of education on migrants' SRH is comparable to the health returns to schooling reported in the existing literature. For instance, extensive evidence from randomized controlled trials, twin studies, and quasi-experiments suggests that having one additional year of education could significantly reduce the mortality rate by 2.6–5 percentage points (Lundborg et al., 2016); the likelihood of obesity by 0.6–3 percentage points (Li and Powdthavee, 2015); and smoking by approximately 2–5 percentage points (Amin et al., 2013). Moreover, Xie

[4] One potential concern is that some control variables, such as family income, farmland, homeownership, and urban area, might be highly correlated, which could introduce simultaneity bias into the estimation; however, the correlation coefficients among these variables are relatively low, ranging from −0.0854 to 0.1851. To check whether the correlation among these control variables biases our results, we exclude them from our specifications and reestimate the models in Table 2. The results in Table A2 show that excluding these control variables does not qualitatively change the estimates of the effects of education on migrants' health status.

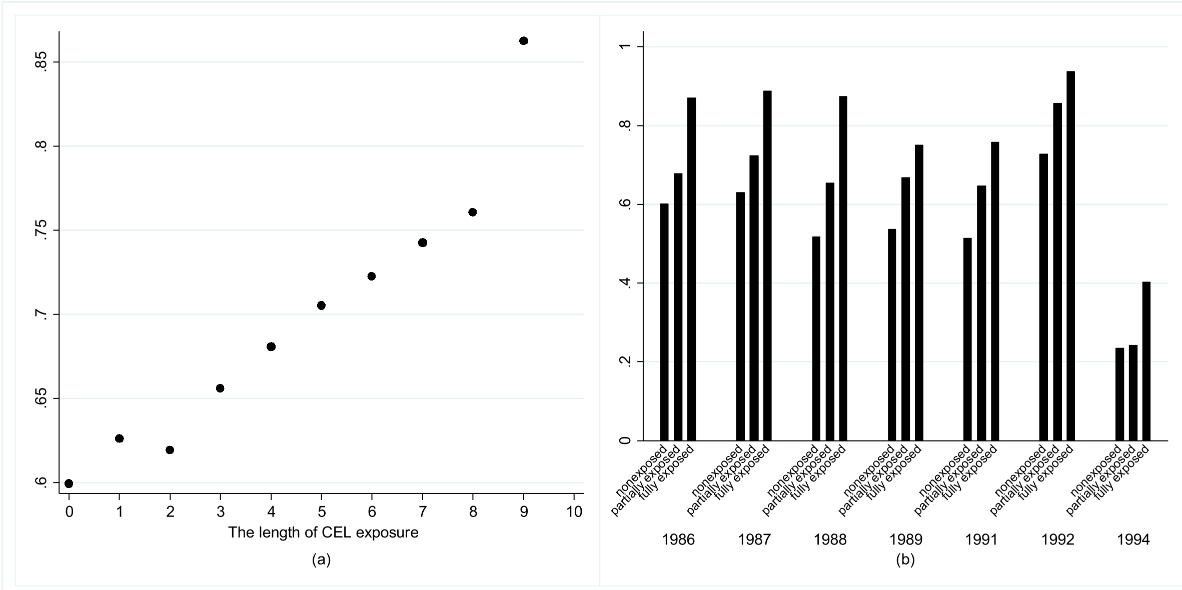


**Fig. 1.** Proportion of migrants with junior high school education by the length of CEL exposure and the timing of the CEL.

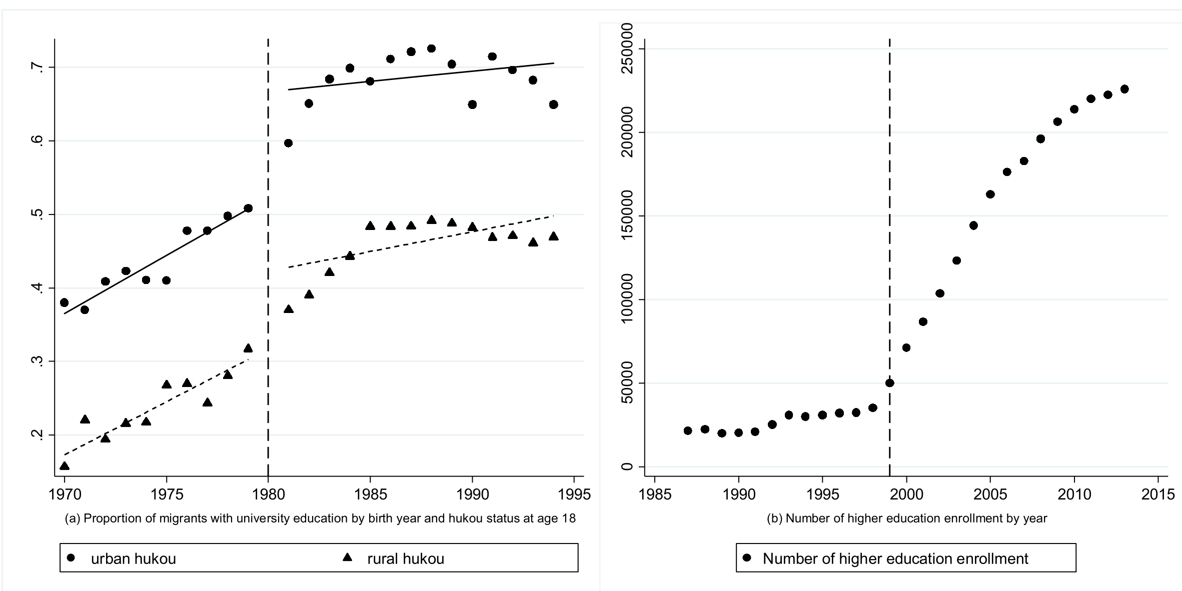


**Fig. 2.** Proportion of migrants with higher education by birth year and *hukou* status at age 18 and the number of higher education enrolment by year.

and Mo (2014) and Huang (2015) find that an extra year of education leads to a 2.33 percentage points reduction in the risk of being overweight and a 2 percentage point decrease in reporting fair or poor health, respectively. The health return rates to education in our study are lower than the rates of wage returns to education in China, which typically range from approximately 8 to 25 percent (Awaworyi Churchill and Mishra, 2018; Dai et al., 2022). For instance, in one study, Yao et al. (2018), who employ a novel set of IVs based on the unique institutional environment in rural China, find that the economic returns to an additional year of schooling among rural migrants in urban China ranged between 7 and 11 percent.

The magnitude of the IV estimates is 5.36 times larger than that of OLS estimates, implying that OLS biases downward the estimated effect of education on health status. The larger IV estimates could be explained by the higher returns for those compliers of the education policies than for the average individual; the IV results can be interpreted as representing the impact on the subgroup of individuals whose educational attainment has been most affected by these policies and changes (Wang et al., 2022). Another possible reason for the result is that the OLS estimates may suffer attenuation bias stemming from the classical measurement error in education (Cui et al., 2019).

One may be concerned that our findings are biased by potential selection into migration, as extensive literature demonstrates that schooling and health conditions have played important roles in migration decisions (Chen, 2011; Ding, 2021); however, this study focuses on the effect of education on the health status of migrants. We do not seek to examine the health returns to education among the general population; thus, migration *per se* should not be regarded as endogenous in the absence of nonmigrant residents in cities in the CMDS sample or any other survey data of migrants (Cheng and Smyth, 2021; Connelly and Zheng, 2003; Messinis, 2013). Moreover, findings from previous studies suggest that with and without adjustment for sample selection bias, the

**Table 2**
Education and health status.

| | Panel A: Junior high school or below | | Panel B: Senior high school or above | |
|---|---|---|---|---|
| | (A1) OLS | (A2) IV | (B1) OLS | (B2) IV |
| Years of education | 0.0187*** | 0.1002*** | 0.0013 | 0.0017 |
| | (10.42) | (2.80) | (1.39) | (0.15) |
| Control variables | Yes | Yes | Yes | Yes |
| Birth province fixed effects | Yes | Yes | Yes | Yes |
| Destination province fixed effects | Yes | Yes | Yes | Yes |
| Cohort trend | Yes | Yes | Yes | Yes |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| adj. $R^2$ | 0.0637 | −0.0295 | 0.0371 | 0.0371 |
| Kleibergen–Paap rk Wald F statistic | | 13.0700 | | 134.1483 |
| Hansen J statistic | | 0.4509 | | 1.2262 |
| p-value | | 0.5019 | | 0.5417 |
| Conditional likelihood-ratio (CLR) test statistic | | 8.48 | | 22.73 |
| CLR test p-value | | 0.0046 | | 0.0000 |
| Lagrange multiplier K test statistic | | 8.31 | | 5.40 |
| K test p-value | | 0.0040 | | 0.0202 |
| K-J test p-value | | 0.0049 | | 0.0000 |
| Anderson–Rubin (AR) test statistic | | 8.81 | | 39.10 |
| AR test p-value | | 0.0122 | | 0.0000 |
| ***First stage results*** | | | | |
| CEL exposure | | 0.1567*** | | |
| | | (4.52) | | |
| Born in the fourth quarter | | 0.0374*** | | |
| | | (2.71) | | |
| Nonagricultural *hukou* before age 18 × Post-1980 birth | | | | 0.3458*** |
| | | | | (6.13) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | | | | −0.0271*** |
| | | | | (−7.21) |
| Linear trend of higher education enrolment | | | | 0.0228*** |
| | | | | (15.14) |
| F test for the joint significance of instruments | | | | |
| F statistic | | 13.07 | | 134.15 |
| p-value | | 0.0000 | | 0.0000 |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1; full results are available from the authors.

estimates are qualitatively similar when examining migration decisions and the economic returns to migrants' education (Connelly et al., 2012; Zhu, 2002). Xing (2014) finds that the selection effect of education among temporary rural migrants is negligible. Other studies find no statistically significant relationship between the level or years of education and the decision to migrate from rural to urban regions in China (Akgüç et al., 2016; Giulietti et al., 2013, 2018). One reason is that rural–urban migration is usually triggered by a need to improve one's individual or household economic conditions rather than potential returns of education, health, and other individual characteristics across different locations (Zhang and Zhao, 2015).

Next, we examine the heterogeneous effects of education on health status between males and females and between urban and rural areas. Panels A and B in Table 3 present the IV estimates of health returns to education by gender and urban/rural origin for less and better-educated subsamples, respectively. The results suggest that females and individuals from rural areas mainly drive the positive effect of schooling on the health status of less-educated migrants. These results are consistent with previous findings that females and individuals from rural or medium and low-income families benefit more from the CEL in terms of employment, health, education, and cognitive ability (Cheng and Smyth, 2021; Huang, 2015; Jiang et al., 2020; Xiao et al., 2017). Higher human capital and improved labor market status could, in turn, improve health status (Wang et al., 2019). Consistent with our main findings, education has no significant impact on the health status of better-educated migrants, regardless of their gender and regional origins.

**Table 3**
Heterogeneity in education effects on health status: IV results.

| | Junior high school or below | Senior high school or above |
|---|---|---|
| **Panel A: By gender** | | |
| Male | 0.0559 | −0.0100 |
| | (0.81) | (−0.61) |
| *N* | 40,364 | 29,359 |
| Female | 0.1009*** | 0.0205 |
| | (3.94) | (1.27) |
| *N* | 40,223 | 28,074 |
| **Panel B: By region** | | |
| Rural area | 0.0998*** | 0.0070 |
| | (2.72) | (0.44) |
| *N* | 74,019 | 42,924 |
| Urban area | 0.1301 | 0.0002 |
| | (1.02) | (0.01) |
| *N* | 6568 | 14,509 |

Notes: z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors.

## 5. Mechanisms

This section uses the Sobel–Goodman mediation test to explore the potential mechanisms linking schooling years with health status among less-educated migrants.[5] Specifically, we examine the mediating roles of health-related information, healthcare utilization, health behaviors, and income. Table 4 presents the results. The first channel that we investigate is monthly wages. Panel A in Table 4 shows that an additional year of education significantly increases one's monthly wages by 25.21 percent, which is very close to the upper bound of the economic returns to schooling in previous studies for China (Awaworyi Churchill, 2018). Higher wages could relax family credit constraints and enable migrants to invest more in their health production, such as purchasing higher quality health services, which, in turn, could improve their SRH. Migrants who had higher wages also reported higher SRH scores. A Sobel–Goodman test of the significance of the mediating effect confirms the existence of mediation through wage earnings and suggests that wages explain 1.61 percent of the total effect.

The second channel that we explore is awareness of health-related information. A large body of literature suggests that better-educated people are better able to gather and interpret health-related information (Chaudhry et al., 2006; McMullan, 2006). In the 2017 CMDS, respondents were asked to report whether they were aware of the National Basic Public Health Services program, which seeks to ensure access to health services for all citizens. We create a binary variable set equal to 1 if the respondent was aware of this program and 0 otherwise. The results in Panel B in Table 4 show that having higher educational attainment increases the probability of knowing about this program. Migrants who

[5] The Sobel–Goodman mediation test quantifies the proportion of the total effect explained by the meditating variable in four steps: first, we examine the effect of education on migrants' health status without controlling for the mediator; second, we investigate the effect of education on the mediator; third, we explore the impacts of both education and the mediator on migrants' health status; and fourth, we estimate the significance of the mediating effect using the Sobel–Goodman mediation test. We employ asymmetric confidence intervals for the mediating effect, estimated through bootstrapping, which represents a more accurate method than that based on the conventional theory of confidence limits (Nadkarni et al., 2011).

**Table 4**
Mechanism analysis.

| | (1) SRH | (2) Mediator | (3) SRH |
|---|---|---|---|
| **Panel A: Wage as a potential mediator** | | | |
| Education | 0.1002*** | 0.2521** | 0.0987*** |
| | (2.80) | (2.31) | (2.75) |
| Monthly wages | | | 0.0064*** |
| | | | (3.23) |
| *N* | 80,587 | 80,587 | 80,587 |
| Indirect effect | 0.0016* | (1.88) | |
| Proportion of indirect effect | 0.0161 | | |
| **Panel B: Health-related information as a potential mediator** | | | |
| Education | 0.1002*** | 0.1115*** | 0.0982*** |
| | (2.80) | (3.19) | (2.69) |
| Knowing about the National Public Health Service Program (yes = 1) | | | 0.0207** |
| | | | (2.30) |
| *N* | 80,587 | 80,587 | 80,587 |
| Indirect effect | 0.0023* | (1.87) | |
| Proportion of indirect effect | 0.0230 | | |
| **Panel C: Health record as a potential mediator** | | | |
| Education | 0.1002*** | 0.0809*** | 0.0982*** |
| | (2.80) | (2.69) | (2.72) |
| Having health record at destination (yes = 1) | | | 0.0260*** |
| | | | (3.55) |
| *N* | 80,587 | 80,587 | 80,587 |
| Indirect effect | 0.0021** | (2.15) | |
| Proportion of indirect effect | 0.0210 | | |
| **Panel D: Being vaccinated as a potential mediator** | | | |
| Education | 0.0094** | 0.0186*** | 0.0090** |
| | (2.32) | (4.94) | (2.22) |
| Received any vaccination (yes = 1) | | | 0.0225** |
| | | | (2.22) |
| *N* | 6,223 | 6,223 | 6,223 |
| Indirect effect | 0.0004** | (2.02) | |
| Proportion of indirect effect | 0.0444 | | |
| **Panel E: Sharing personal hygiene products as a potential mediator** | | | |
| Education | 0.0094** | −0.0210*** | 0.0086** |
| | (2.32) | (−5.32) | (2.10) |
| Sharing personal hygiene products (yes = 1) | | | −0.0408*** |
| | | | (−3.34) |
| *N* | 6,223 | 6,223 | 6,223 |
| Indirect effect | 0.0009*** | (2.83) | |
| Proportion of indirect effect | 0.0909 | | |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors. Panels A–C present the IV results. Panels D and E present OLS results because the information about vaccination and sharing groceries is only available for eight cities, and the IVs do not pass the test for weak instruments.

know about this program also reported higher SRH scores. A Sobel–Goodman test of the significance of the mediating effect confirms the existence of mediation via health-related information. The total effect explained is 2.30 percent, slightly higher than the effect mediated via wages.

The third channel that we examine is healthcare utilization. Existing research finds that educational attainment is one of the most important factors that influence the use of healthcare services (Celik and Hotchkiss, 2000). Better-educated individuals are more likely to be aware of the benefits of healthcare and, as a result, are more likely to use preventive healthcare services (Terraneo, 2015). We use two indicators to proxy for the utilization of healthcare services: whether the respondent has a personal health record in the destination community and whether the respondent has received any vaccinations, such as influenza vaccination, hepatitis vaccination, or Japanese encephalitis vaccination. Results in panels C and D in Table 4 suggest that healthcare utilization positively mediates the observed relationship between education and health status. We find that having more schooling increases the likelihood of having a health record in the destination community and being vaccinated, which, in turn, lead to improved SRH. The Sobel–Goodman test further suggests that these two indicators of healthcare utilization are significantly positive and can explain 2.10–4.44 percent of the total effect.

Finally, we examine whether health behavior is a potential mechanism underlying the estimated effect of education on health status. The 2017 CMDS contains survey questions regarding migrants' health-related behaviors centered on whether the respondent shares personal hygiene items, such as face washers, washbowls, shavers, and drinking cups, with other people. We define a dummy variable equal to 1 if the respondent shares any of these personal hygiene products with others and 0 otherwise. The results in panel E show that having a higher level of education reduces the probability of following unhygienic practices, which are conducive to poorer health status. Our results support previous findings that more schooling is associated with improving healthy behaviors (Amin et al., 2013; Cutler and Lleras-Muney, 2010). Moreover, health behaviors could reflect one's health literacy, an underlying mechanism driving the relationship between education and health (Van Der Heide et al., 2013). A Sobel–Goodman test confirms that health behavior positively mediates the relationship between education and migrants' health status and explains 9.09 percent of the total effect.

We also employ structural equation modeling (SEM) to conduct causal mediation analysis as a robustness check on these findings. The results are presented in the appendix in Table A3. The results from the adjusted Baron and Kenny's approach (Iacobucci et al., 2007) show that for each of these mediators, the estimated coefficients in the three steps and Sobel z-test are all significant, suggesting that there is partial mediation. The results from the Zhao et al. (2010) method also point to partial mediation between education and health status among less-educated migrants, as both coefficients in step one and the Monte Carlo test are significant. Moreover, we find that the total indirect effects of schooling years on migrants' health status are statistically significant with effect sizes that constitute between 0.6 and 9.1 percent of the total effect, which are comparable in magnitude to those reported in Table 4.

## 6. Robustness checks

We conduct several robustness checks on our main findings. First, we construct alternative measures of health status and educational attainment. Specifically, to attenuate the risks of over-reporting or under-reporting of health status, we recode this variable as a dummy variable equal to 1 if migrants reported that their health status was "very healthy" and 0 otherwise. Similarly, we recode educational attainment for less and more educated migrants, employing a binary variable indicating whether they have completed junior high school education or hold a university qualification, respectively. We replace the ordinal measure of health status with a dummy variable in panel A in Table 5 and replace the measures of both health status and schooling years with their corresponding dummy variables in panel B. The OLS and IV estimates show that education significantly and positively affects SRH among less-educated migrants, while its impact on health status among better-educated migrants is not significant. The estimated effect suggests that an additional year of compulsory schooling and having junior high school education could increase the probability of being "very healthy" by 1.22–7.13 percentage points and 4.73–32.75 percentage points, respectively. The results are generally consistent with our main findings in Table 2, suggesting that measurement error is not a significant issue biasing the estimates in this study.

Second, we account for a demanding set of region-specific cohort trends, using interaction terms between birth and destination provincial dummies and the linear cohort trend to capture shared trends in our outcome variable and check whether our main estimates are biased by unobserved province-cohort-specific heterogeneity. The results in Table 6 suggest that this more flexible specification does not qualitatively change our main results. The magnitudes and significance of health returns to education are almost identical to those reported in Table 2. These findings help to rule out the possibility that our main estimates are driven by unobserved heterogeneity specific to cohorts

**Table 5**
Robustness checks: Recoding health status and schooling years as dummy variables.

| | Junior high school or below | | Senior high school or above | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| **Panel A: Dummy variable for health status** | | | | |
| Years of education | 0.0122*** | 0.0713*** | 0.0010 | 0.0049 |
| | (11.85) | (2.59) | (1.15) | (0.44) |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| Kleibergen–Paap rk Wald F statistic | | 13.0700 | | 134.1483 |
| Hansen J statistic | | 0.6505 | | 0.9619 |
| p-value | | 0.4199 | | 0.6182 |
| **Panel B: Dummy variables for both health status and education** | | | | |
| 9-year compulsory education/higher education qualification | 0.0473*** | 0.3275** | 0.0032 | 0.0165 |
| | (12.28) | (2.40) | (0.95) | (0.37) |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| Kleibergen–Paap rk Wald F statistic | | 14.0724 | | 101.4977 |
| Hansen J statistic | | 1.2962 | | 1.0343 |
| p-value | | 0.2549 | | 0.5962 |

Notes: t statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors.

**Table 6**
Robustness checks: Controlling for birth and destination province-specific trends.

| | Junior high school or below | | Senior high school or above | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| Years of education | 0.0189*** | 0.1132*** | 0.0011 | 0.0063 |
| | (9.94) | (4.09) | (1.14) | (0.57) |
| Birth province fixed effects | Yes | Yes | Yes | Yes |
| Destination province fixed effects | Yes | Yes | Yes | Yes |
| Cohort trend | Yes | Yes | Yes | Yes |
| Birth province fixed effects × Cohort trend | Yes | Yes | Yes | Yes |
| Destination province fixed effects × Cohort trend | Yes | Yes | Yes | Yes |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| Kleibergen–Paap rk Wald F statistic | | 48.3055 | | 137.3847 |
| Hansen J statistic | | 0.2431 | | 0.9493 |
| p-value | | 0.6220 | | 0.6221 |
| First stage results | | | | |
| CEL exposure | | 0.1794*** | | |
| | | (9.36) | | |
| Quarter of birth (fourth quarter = 1) | | 0.0361*** | | |
| | | (2.88) | | |
| Nonagricultural *hukou* before age 18 × Post-1980 birth | | | | 0.3379*** |
| | | | | (6.14) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | | | | −0.0275*** |
| | | | | (−7.18) |
| Linear trend of higher education enrolment | | | | 0.0233*** |
| | | | | (15.62) |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1; full results are available from the authors.

and provinces.

Third, we examine whether our results are robust to a narrower range of cohorts. To do so, we restrict our sample to migrants born 10 years earlier or later than the first affected cohort, less-educated migrants born between 1975 and 1995, and better-educated migrants born between 1970 and 1990. Employing a narrower range of cohorts is likely to be more homogeneous, which could mitigate potential omitted variable bias due to the unobserved cohort heterogeneity. Models 1–2 and 5–6 in Table 7 suggest that significant and positive health returns to an additional year of education only exist among migrants with lower educational attainment, reaffirming our main findings in Table 2.

In addition, in models 3–4 and 7–8, we exclude the first affected cohort from the less-educated sample, namely, migrants who were partially exposed to the CEL and those who were 6 years old when the CEL was enacted, and respondents who turned 18 years old between 1995 and 1998 from the better-educated sample. Then, we reestimate the effect of educational attainment on health status using the reduced samples. We make this adjustment because some less-educated migrants may not have immediately been affected by the implementation of the CEL due to the relatively lower compliance rates at the beginning. For better-educated migrants, some respondents may have retaken the national university entrance examination after the 1999 higher education expansion was implemented. The results suggest that education has a significant and positive effect on SRH among less-educated migrants, while its impact on the SRH among the better-educated migrants is not significant, consistent with Table 2.

Fourth, in our primary analysis, we assume that children typically start their education at age 6 and define migrants between 7 and 15 as the partially exposed cohort in constructing the instrument; however, there was no rigid school entry age before the 1986 CEL. The school enrolment age norms varied by province and *hukou* status, with the age window being between 4.75 and 8.75 years. Moreover, most provinces initially set the mandatory school entry age to 7 under the new law (Chen and Park, 2021); thus, we redefine migrants between 8 and 16 as the partially exposed cohort when the law was implemented, those aged 7 or below as a fully exposed cohort, and those aged 17 or above as a nonexposed cohort. We use this alternative instrument to test the robustness of our results. Model 1 in Table 8 shows that a more prolonged reform exposure is associated with more schooling years, which could significantly promote the SRH among the less-educated migrants.

Fifth, although the CEL has been widely used as an IV in previous studies to examine the pecuniary and nonpecuniary benefits of schooling due to its quasi-experimental nature, the exogeneity of institutional changes have been questioned (Xie and Mo, 2014). This is because the 9-year education program is more likely to be a long-run ambition of the Chinese government than an objective to be immediately realized. The extent of CEL enforcement was influenced by economic growth and fiscal constraints that may also be correlated with the regional healthcare provision. The compliance rates were very low at the beginning of the implementation of the CEL and varied considerably across provinces. By 1990, only three-quarters (76 percent) of counties and 91 percent of the population had realized universal compulsory education (Connelly and Zheng, 2003). To address the potential violation of the exogeneity assumption of the IV based on the CEL, we use another institutional reform in China, namely, the Prohibition on Using Child Labor, introduced in 1991 as an alternative instrument to capture the effect of schooling years on health status. Previous studies adopt a similar IV strategy to examine the marginal payoffs of schooling (Xie and Mo, 2014). The Prohibition on Using Child Labor stipulates that children under the age of 16 cannot be employed by any work unit and has been strictly implemented nationwide. We generate a dummy variable equal to 1 if the respondent was younger than 16 years old when the law was enacted in 1991 and 0 otherwise. The IV estimates in model 2 in Table 8 show that the IVs are strong and valid. We find that migrants regulated by the child labor law are likely to receive more years of schooling and that additional education is associated with better SRH, consistent with the baseline results in Table 2.

Sixth, Bengoa and Rick (2020) find that obtaining an urban *hukou* could significantly improve migrants' objective health conditions and nutritional conditions in the medium term, such as maintaining lower levels of blood pressure and reducing the likelihood of hypertension, while it has no significant effect on SRH status. In our sample, 4.18

**Table 7**
Robustness checks: Using restricted samples.

| | Junior high school or below, born 1975–1995 | | Junior high school or below, dropping the first affected cohort | | Senior high school or above, born 1970–1990 | | Senior high school or above, dropping those who turned 18 years between 1995 and 1998 | |
|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) IV | (3) OLS | (4) IV | (5) OLS | (6) IV | (7) OLS | (8) IV |
| Years of education | 0.0185*** | 0.0652* | 0.0167*** | 0.1134*** | 0.0011 | −0.0207 | 0.0010 | 0.0166 |
| | (8.95) | (1.80) | (7.66) | (4.37) | (1.04) | (−1.02) | (1.02) | (1.38) |
| *N* | 59,518 | 59,518 | 47,796 | 47,796 | 44,794 | 44,794 | 51,596 | 51,596 |
| adj. $R^2$ | 0.0483 | 0.0184 | 0.0642 | −0.0608 | 0.0361 | 0.0275 | 0.0376 | 0.0325 |
| Kleibergen–Paap rk Wald F statistic | | 10.1095 | | 23.9987 | | 25.1189 | | 166.2291 |
| Hansen J statistic | | 4.7203 | | 0.0285 | | 1.2816 | | 1.6224 |
| p-value | | 0.0944 | | 0.8659 | | 0.5269 | | 0.4443 |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors.

**Table 8**
Robustness checks: Employing alternative IVs for the CEL.

| | Junior high school or below | |
|---|---|---|
| | (1) | (2) |
| Years of education | 0.1202*** | 0.1905*** |
| | (2.98) | (4.98) |
| *N* | 80,587 | 80,587 |
| adj. $R^2$ | −0.0808 | −0.3507 |
| Kleibergen–Paap rk Wald F statistic | 14.8034 | 27.7500 |
| Hansen J statistic | 0.1760 | 0.1026 |
| p-value | 0.6749 | 0.7487 |
| First stage results | | |
| CEL exposure (children start primary school at age 7) | 0.1568*** | |
| | (4.83) | |
| Quarter of birth (fourth quarter = 1) | 0.0375*** | 0.0377*** |
| | (2.71) | (2.73) |
| Child labor law (<16 years old in 1991 = 1) | | 0.2120*** |
| | | (6.92) |

Notes: t statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors.

**Table 9**
Robustness check: Controlling for the change in *hukou* status.

| | Panel A: Junior high school or below | | Panel B: Senior high school or above | |
|---|---|---|---|---|
| | (A1) OLS | (A2) IV | (B1) OLS | (B2) IV |
| Years of education | 0.0188*** | 0.0998*** | 0.0013 | 0.0013 |
| | (10.44) | (2.79) | (1.38) | (0.12) |
| Obtaining nonagricultural or resident *hukou* | −0.0220* | −0.0332** | 0.0014 | 0.0014 |
| | (−1.80) | (−2.55) | (0.24) | (0.24) |
| Control variables | Yes | Yes | Yes | Yes |
| Birth province fixed effects | Yes | Yes | Yes | Yes |
| Destination province fixed effects | Yes | Yes | Yes | Yes |
| Cohort trend | Yes | Yes | Yes | Yes |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| adj. $R^2$ | 0.0638 | −0.0284 | 0.0371 | 0.0371 |
| Kleibergen–Paap rk Wald F statistic | | 13.1311 | | 142.8158 |
| Hansen J statistic | | 0.4529 | | 1.7245 |
| p-value | | 0.5009 | | 0.4222 |
| ***First stage results*** | | | | |
| CEL exposure | | 0.1570*** | | |
| | | (4.53) | | |
| Born in the fourth quarter | | 0.0375*** | | |
| | | (2.72) | | |
| Nonagricultural *hukou* before age 18 × Post-1980 birth | | | | 0.3458*** |
| | | | | (6.13) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | | | | −0.0268*** |
| | | | | (−7.35) |
| Linear trend of higher education enrolment | | | | 0.0228*** |
| | | | | (15.14) |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1; full results are available from the authors.

percent of less-educated migrants and 18.63 percent of better-educated migrants changed their *hukou* status from agricultural *hukou* to nonagricultural *hukou* or resident *hukou*. To check whether changing *hukou* status may bias our estimates of the effects of education on health status, we include a dummy variable equal to 1 if migrants have changed their *hukou* status before and 0 otherwise in our specifications. The results in Table 9 show that obtaining a nonagricultural or resident *hukou* significantly affects the health status of migrants with junior high school education or below, while its impact on those with senior high school education or above is not significant; however, including a change in *hukou* status does not qualitatively change the OLS and IV estimates of the effects of migrants' schooling years. The magnitudes and significance of migrants' education are almost identical to those reported in Table 2.

All respondents in our sample reported that they did not hold a local *hukou* in the county of their residence; however, one might be concerned about the possibility of errors in reporting migrants' *hukou* status and associated actual migration status, which might also introduce bias into the estimates of the relationship between education and health. To evaluate whether our main estimates are affected by potential self-reporting bias, following the approach of Wang et al. (2021), we randomly drop 5 percent of migrants in our sample and replicate the IV estimates in models A2 and B2 in Table 2 with a reduced sample 1000 times, respectively. If most migrants accurately report their *hukou* and migration status, the IV estimates using the reduced samples should not seriously deviate from our baseline results. Fig. 3a and b plot the distributions of the estimated coefficients and corresponding t-values for migrants with lower and higher levels of education, respectively. We find that all estimates from the reduced samples center around our IV estimates in Table 2, denoted by the short-dashed line. The upper and lower bounds of the distribution for less-educated migrants share the same sign with our baseline estimates. The health returns to education are significant among migrants with a lower level of education but insignificant for better-educated migrants. The results suggest that potential self-reporting bias is unlikely to impair our findings.

Seventh, migrants with senior high school education or above may also be affected by the introduction of the CEL; some migrants may have reached that education level due to the increase in the number of years

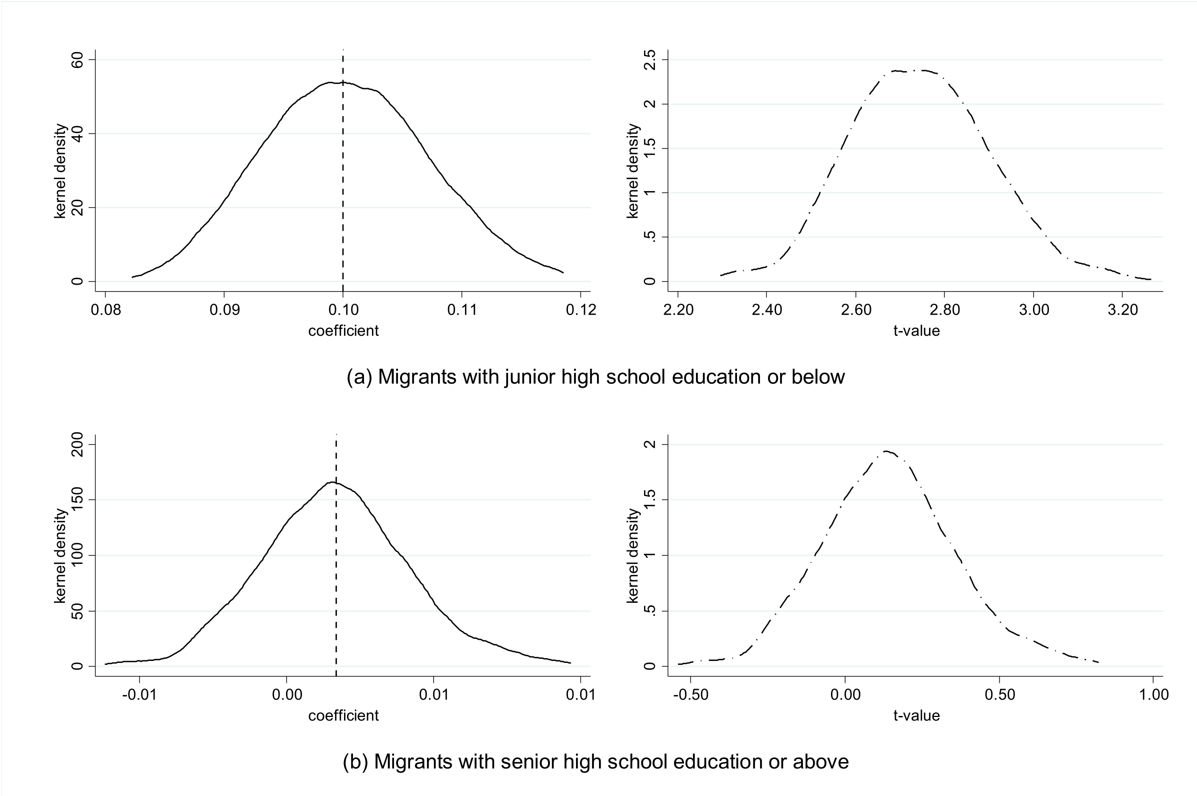


**Fig. 3.** Distribution of estimated coefficients and t-values from reduced samples.

of compulsory schooling. To check this, we include the length of CEL exposure as an additional IV in model B2 in Table 2. Table 10 presents the results. Although the first stage results show that the length of CEL exposure is significantly and positively associated with educational attainment, we find that schooling years still have no significant effect on health status for better-educated migrants, which is consistent with our main findings reported in Table 2.

**Table 10**
Education and health status for better-educated migrants, including an additional IV.

| | IV |
|---|---|
| Years of education | 0.0049 |
| | (0.44) |
| Control variables | Yes |
| Birth province fixed effects | Yes |
| Destination province fixed effects | Yes |
| Cohort trend | Yes |
| *N* | 57,433 |
| adj. $R^2$ | 0.0369 |
| Kleibergen–Paap rk Wald F statistic | 103.0946 |
| Hansen J statistic | 2.1634 |
| p-value | 0.5392 |
| ***First stage results*** | |
| CEL exposure | 0.1506*** |
| | (5.65) |
| Nonagricultural *hukou* before age 18 × Post-1980 birth | 0.2951*** |
| | (5.29) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | −0.0245*** |
| | (−6.58) |
| Linear trend of higher education enrolment | 0.0183*** |
| | (10.27) |
| F test for the joint significance of instruments | |
| F statistic | 103.09 |
| p-value | 0.0000 |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1; full results are available from the authors.

Eighth, we employ a kinky least squares model as an alternative to employing IVs. Unlike the well-known Oster (2019) approach that imposes restrictions on the magnitude of the correlation of the endogenous variable and the structural error term relative to the correlation of the endogenous regressor with other control variables, this method places bounds directly on the endogeneity correlation (Kiviet, 2020). As the results in Table 2 imply that unobservables might bias the OLS estimates downward, the correlation between schooling years and the error term should be negative. Therefore, we set the postulated endogeneity of years of education to be negative, ranging from −0.5 to zero. The kinky least squares estimates, presented in Fig. 4, show that education is positively and significantly associated with health status for less-educated migrants and that the size effect increases with the

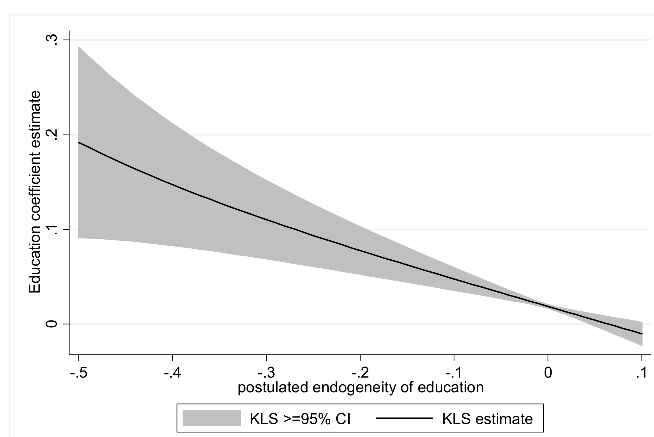


**Fig. 4.** Kinky least squares estimates for the effect of schooling years on health status.

postulated degree of endogeneity. We also employ another approach to quantify the percent bias necessary to invalidate inference from a Rubin causal model framework (Xu et al., 2019). The results show that to invalidate the inference of there being a positive impact of education on the SRH for less-educated migrants, 41.48 percent of the observed cases would have to be replaced with cases with an effect of zero at the 10 percent significance level, and the impact of an omitted confounding variable would have to be 0.0041, which is larger than the impact of gender (0.0037) and family size (0.0012). In sum, the two tests for the sensitivity of causal inference provide strong evidence that our main findings are robust.

Finally, we perform two placebo tests following the approach in the previous literature (Chakraborty and Jayaraman, 2019; Cui et al., 2019). For the first placebo test, for less-educated migrants, we randomly assign the effective year of the CEL to each respondent (1986–1994) and then construct the placebo length of CEL exposure in the same manner as our first IV measure.

We examine the placebo and actual treatment effects on migrants' educational attainment and conduct the random assignment 1000 times. Fig. 5 plots the distributions of OLS estimates of the association between migrants' placebo and actual length of CEL exposure and their years of education. The placebo treatment effect, denoted by the short-dashed line, is very close to zero and is significantly different from the actual treatment effect of CEL exposure, denoted by the solid line.

For the second test, for better-educated migrants, we restrict our sample to migrants born before 1980 who are unaffected by the higher education expansion policy. We assume that the higher education expansion policy was implemented in 1994, five years earlier than the actual policy year, and the first cohort affected by the placebo higher education expansion policy was born in 1975. We construct three similar IVs to those employed in our primary analysis based on the placebo policy year and affected cohorts and examine their impacts on educational attainment. The results in Table 11 show that the three IVs have no significant effects on years of schooling and the probability of obtaining a higher education qualification. The two placebo tests lend further credence to our identification strategy.

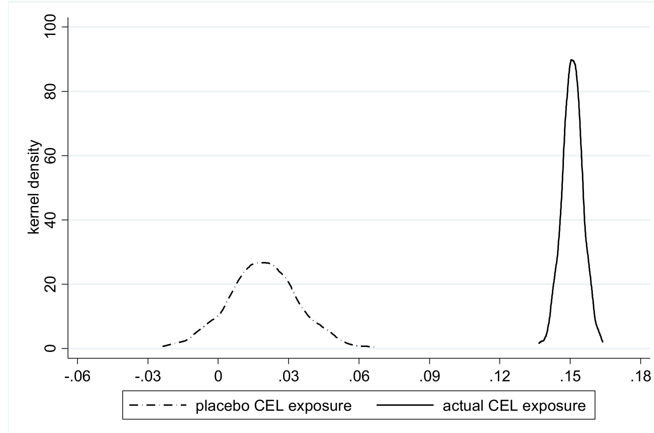


**Fig. 5.** Distribution of placebo and actual treatment effects on migrants' educational attainment.

**Table 11**
Placebo tests for impacts of the higher education policy on migrants' educational attainment.

| | Years of education | Hold higher education qualifications (yes = 1) |
|---|---|---|
| Nonagricultural *hukou* before age 18 × Post-1975 birth | −0.0490 | −0.0183 |
| | (−0.69) | (−0.94) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | 0.0055 | −0.0004 |
| | (0.36) | (−0.10) |
| Placebo linear trend of higher education enrolment | 0.0002 | −0.0004 |
| | (0.06) | (−0.47) |
| *N* | 9,940 | 9,940 |
| adj. $R^2$ | 0.3088 | 0.2583 |

Notes: t or z statistics in parentheses; *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$. All specifications include the full set of controls listed in Table A1 and birth and destination province fixed effects; full results are available from the authors.

## 7. Conclusion

This study provides the first causal estimates of the nonlinear effect of educational attainment on migrants' health status in China. In doing so, we contribute to the knowledge about the causal effect of education on SRH using variations in educational policies outside of Europe and the United States. More generally, we contribute to recent literature that has exploited the exogenous introduction of the CEL and expansion of the higher education sector in China to estimate the causal effect of education on a range of outcomes, including consumption, entrepreneurship, and home ownership, among others (Cheng, 2021; Huang et al., 2021; Wang et al., 2022).

We find that education has a significant and positive impact on health status among migrants with lower education levels, while its impact on better-educated migrants is not significant, suggesting diminishing health returns to education. These results are robust to various sensitivity checks. This finding is consistent with studies suggesting that the nonmarket returns to education are generally highest for people with less ability and fewer opportunities (Heckman et al., 2018).

The mechanism analysis shows that wages, health-related information, healthcare utilization, and health behaviors are channels through which education affects the health status of less-educated migrants. We also find evidence of considerable heterogeneity in the health returns to education across subsamples. Females and individuals from rural areas are likely to benefit more from increased years of schooling. Our findings imply that compulsory education may reduce health inequality between less and better-educated individuals, as well as gender and urban–rural health inequalities in China.

Economists are increasingly interested in the long-term social benefits of investing in education, including health benefits, partly motivated by the need to reduce soaring healthcare costs. For instance, health sector costs in the United States are now the size of the United Kingdom's entire GDP (Courtin et al., 2020). If investing in education can be shown to have a positive causal effect on improving health status, this offers the opportunity to "get ahead of the curve" and reduce healthcare for individuals later in life. Our findings carry important policy implications for reducing health inequality and improving the well-being of migrants in China and other developing countries experiencing massive population migration, such as Bangladesh, India, and Vietnam (Wang et al., 2021). Our findings of decreasing health returns to schooling imply that the Chinese government should invest more resources in compulsory education, especially in basic education in rural areas. Moreover, given that migration is highly selective in terms of education and health, such investment in basic education is likely to have its most substantial effects on the promotion of long-term health status among the most vulnerable groups, especially rural residents with lesser human capital who do not migrate.

## Declaration of competing interest

None

## Data availability

The authors do not have permission to share data.

## Acknowledgments

Yuanyuan Chen acknowledges financial support from the Philosophy and Social Science Planning Project of Guangdong Province, China (grant number: GD21CSH04). Haining Wang acknowledges financial support from the Natural Science Foundation of Guangdong Province, China (grant number: 2021A1515012644). Zhiming Cheng acknowledges financial support from the Macquarie University Research Acceleration Scheme (grant number: 5225220). The auhtors also would like to thank two anonymous referees for their valuable comments and suggestions.

## Appendix

**Table A1**
Summary statistics for control variables

| Variable | Definition | Junior high school or below | | Senior high school or above | |
|---|---|---|---|---|---|
| | | Mean | S.D. | Mean | S.D. |
| Male | Male = 1, female = 0 | 0.50 | 0.50 | 0.51 | 0.50 |
| Han ethnicity | Han = 1, ethnic minorities = 0 | 0.89 | 0.32 | 0.93 | 0.26 |
| *Hukou* status | | | | | |
| Agricultural *hukou* | Reference group | 0.89 | 0.31 | 0.63 | 0.48 |
| Nonagricultural *hukou* | | 0.05 | 0.22 | 0.26 | 0.44 |
| Resident *hukou* | | 0.06 | 0.24 | 0.10 | 0.31 |
| Marital status | | | | | |
| Single | Reference group | 0.13 | 0.34 | 0.22 | 0.41 |
| Married | | 0.85 | 0.36 | 0.77 | 0.42 |
| Others | | 0.02 | 0.15 | 0.01 | 0.12 |
| CCP member | Member of the Chinese Communist Party (yes = 1; no = 0) | 0.01 | 0.12 | 0.09 | 0.29 |
| Type of migration | | | | | |
| Inter-province | Reference group | 0.52 | 0.50 | 0.45 | 0.50 |
| Inter-city | | 0.30 | 0.46 | 0.37 | 0.48 |
| Inter-county | | 0.17 | 0.38 | 0.18 | 0.39 |
| Years of migration | Years since migrating to current residence | 11.31 | 7.33 | 9.53 | 6.25 |
| Employment | | | | | |
| Employee | Reference group | 0.44 | 0.50 | 0.59 | 0.49 |
| Employer/self-employed | | 0.38 | 0.49 | 0.25 | 0.43 |
| Others | | 0.01 | 0.11 | 0.02 | 0.15 |
| No work | | 0.16 | 0.37 | 0.13 | 0.34 |
| Occupation | | | | | |
| Manager & technician | Reference group | 0.03 | 0.17 | 0.21 | 0.41 |
| Business & service worker | | 0.53 | 0.50 | 0.47 | 0.50 |
| Production worker | | 0.22 | 0.41 | 0.14 | 0.34 |
| Others | | 0.06 | 0.24 | 0.05 | 0.21 |
| No work | | 0.16 | 0.37 | 0.13 | 0.34 |
| Working hour | Average working hours per week+1, in logarithm | 3.38 | 1.56 | 3.38 | 1.39 |
| Family size | Number of family members | 3.36 | 1.14 | 2.94 | 1.16 |
| Children | Number of children | 1.37 | 0.85 | 0.90 | 0.75 |
| Parental migration | At least one of the respondent's parents was a migrant = 1, otherwise = 0 | 0.22 | 0.41 | 0.27 | 0.44 |
| Family income | Average monthly family income per capita, in logarithm | 7.47 | 0.75 | 7.86 | 0.75 |
| Farmland | Having contracted farmland in origins = 1, otherwise = 0 | 0.50 | 0.50 | 0.33 | 0.47 |
| Homeownership | Own the house that currently lives in = 1, otherwise = 0 | 0.22 | 0.41 | 0.36 | 0.48 |
| Cohort trend | A birth year linear trend equals the birth year – 1969 | 13.28 | 7.80 | 16.54 | 6.08 |
| Urban area | Urban area = 1, rural area = 0 | 0.67 | 0.47 | 0.83 | 0.38 |

**Table A2**
Education and health status-controlling for a restricted control variable set

| | Panel A: Junior high school or below | | Panel B: Senior high school or above | |
|---|---|---|---|---|
| | (A1) OLS | (A2) IV | (B1) OLS | (B2) IV |
| Years of education | 0.0196*** | 0.1083*** | 0.0023** | 0.0076 |
| | (10.58) | (3.16) | (2.41) | (0.72) |
| Control variables | Yes | Yes | Yes | Yes |
| Birth province fixed effects | Yes | Yes | Yes | Yes |
| Destination province fixed effects | Yes | Yes | Yes | Yes |
| Cohort trend | Yes | Yes | Yes | Yes |
| *N* | 80,587 | 80,587 | 57,433 | 57,433 |
| adj. $R^2$ | 0.0603 | −0.0512 | 0.0351 | 0.0345 |
| Kleibergen–Paap rk Wald F statistic | | 15.1566 | | 152.7049 |
| Hansen J statistic | | 0.3850 | | 0.9867 |
| p-value | | 0.5349 | | 0.6106 |
| ***First stage results*** | | | | |
| CEL exposure | | 0.1671*** | | |
| | | (4.86) | | |
| Born in the fourth quarter | | 0.0406*** | | |

(*continued on next page*)

**Table A2** (*continued*)

| | Panel A: Junior high school or below | | Panel B: Senior high school or above | |
|---|---|---|---|---|
| | (A1) OLS | (A2) IV | (B1) OLS | (B2) IV |
| | | (3.00) | | |
| Nonagricultural *hukou* before age 18 × Post-1980 birth | | | | 0.3988*** |
| | | | | (6.94) |
| Nonagricultural *hukou* before age 18 × Birth year linear trend | | | | −0.0297*** |
| | | | | (−7.68) |
| Linear trend of higher education enrolment | | | | 0.0247*** |
| | | | | (15.77) |

Notes: t or z statistics in parentheses; *p < 0.10, **p < 0.05, and ***p < 0.01. All specifications include the full set of controls listed in Table A1 except for family income, farmland, homeownership, and urban area; full results are available from the authors.

**Table A3**
SEM-based medication analysis

| | (1) Delta | (2) Sobel | (3) Monte Carlo |
|---|---|---|---|
| **Panel A: Wage as a potential mediator** | | | |
| Indirect effect | 0.000 | 0.000 | 0.000 |
| Standard error | 0.000 | 0.000 | 0.000 |
| z-value | 4.068 | 4.068 | 4.045 |
| p-value | 0.000 | 0.000 | 0.000 |
| Confidence interval | [0.000, 0.001] | [0.000, 0.001] | [0.000, 0.001] |
| *Baron and Kenny approach to testing mediation* | | | |
| Step 1: X- > M | β = 0.007 | p = 0.000 | |
| Step 2: M- > Y | β = 0.062 | p = 0.000 | |
| Step 3: X- > Y | β = 0.078 | p = 0.000 | |
| *Zhao, Lynch & Chen's approach to testing mediation* | | | |
| Step 1: X- > Y | β = 0.078 | p = 0.000 | |
| Proportion of indirect effect | 0.006 | | |
| **Panel B: Health-related information as a potential mediator** | | | |
| Indirect effect | 0.003 | 0.003 | 0.003 |
| Standard error | 0.000 | 0.000 | 0.000 |
| z-value | 10.367 | 10.372 | 10.382 |
| p-value | 0.000 | 0.000 | 0.000 |
| Confidence interval | [0.003, 0.004] | [0.003, 0.004] | [0.003, 0.004] |
| *Baron and Kenny approach to testing mediation* | | | |
| Step 1: X- > M | β = 0.071 | p = 0.000 | |
| Step 2: M- > Y | β = 0.044 | p = 0.000 | |
| Step 3: X- > Y | β = 0.075 | p = 0.000 | |
| *Zhao, Lynch & Chen's approach to testing mediation* | | | |
| Step 1: X- > Y | β = 0.075 | p = 0.000 | |
| Proportion of indirect effect | 0.040 | | |
| **Panel C: Health record as a potential mediator** | | | |
| Indirect effect | 0.001 | 0.001 | 0.001 |
| Standard error | 0.000 | 0.000 | 0.000 |
| z-value | 6.461 | 6.462 | 6.451 |
| p-value | 0.000 | 0.000 | 0.000 |
| Confidence interval | [0.001, 0.001] | [0.001, 0.001] | [0.001, 0.001] |
| *Baron and Kenny approach to testing mediation* | | | |
| Step 1: X- > M | β = 0.032 | p = 0.000 | |
| Step 2: M- > Y | β = 0.036 | p = 0.000 | |
| Step 3: X- > Y | β = 0.077 | p = 0.000 | |
| *Zhao, Lynch & Chen's approach to testing mediation* | | | |
| Step 1: X- > Y | β = 0.077 | p = 0.000 | |
| Proportion of indirect effect | 0.014 | | |
| **Panel D: Being vaccinated as a potential mediator** | | | |
| Indirect effect | 0.002 | 0.002 | 0.002 |
| Standard error | 0.001 | 0.001 | 0.001 |
| z-value | 1.945 | 1.946 | 1.915 |
| p-value | 0.052 | 0.052 | 0.055 |
| Confidence interval | [0.000, 0.003] | [0.000, 0.003] | [0.000, 0.003] |
| *Baron and Kenny approach to testing mediation* | | | |
| Step 1: X- > M | β = 0.064 | p = 0.000 | |
| Step 2: M- > Y | β = 0.027 | p = 0.032 | |
| Step 3: X- > Y | β = 0.037 | p = 0.007 | |
| *Zhao, Lynch & Chen's approach to testing mediation* | | | |
| Step 1: X- > Y | β = 0.037 | p = 0.007 | |
| Proportion of indirect effect | 0.044 | | |
| **Panel E: Sharing personal hygiene products as a potential mediator** | | | |
| Indirect effect | 0.004 | 0.004 | 0.004 |
| Standard error | 0.001 | 0.001 | 0.001 |
| z-value | 3.043 | 3.045 | 3.014 |
| p-value | 0.002 | 0.002 | 0.003 |

(*continued on next page*)

**Table A3** (*continued*)

| | (1) Delta | (2) Sobel | (3) Monte Carlo |
|---|---|---|---|
| Confidence interval | [0.001, 0.006] | [0.001, 0.006] | [0.001, 0.006] |
| *Baron and Kenny approach to testing mediation* | | | |
| Step 1: X- > M | β = −0.077 | p = 0.000 | |
| Step 2: M- > Y | β = −0.047 | p = 0.000 | |
| Step 3: X- > Y | β = 0.036 | p = 0.010 | |
| *Zhao, Lynch & Chen's approach to testing mediation* | | | |
| Step 1: X- > Y | β = 0.036 | p = 0.010 | |
| Proportion of indirect effect | 0.091 | | |

## References


Akgüç, M., Liu, X., Tani, M., Zimmermann, K.F., 2016. Risk attitudes and migration. China Econ. Rev. 37, 166–176.
Albarrán, P., Hidalgo-Hidalgo, M., Iturbe-Ormaetxe, I., 2020. Education and adult health: is there a causal effect? Soc. Sci. Med. 249, 112830.
Amin, V., Behrman, J.R., Spector, T.D., 2013. Does more schooling improve health outcomes and health related behaviors? Evidence from UK twins. Econ. Educ. Rev. 35, 134–148.
Angrist, J.D., Keueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Q. J. Econ. 106 (4), 979–1014.
Awaworyi Churchill, S., Mishra, V., 2018. Returns to education in China: a meta-analysis. Appl. Econ. 50 (54), 5903–5919.
Belzil, C., Hansen, J., 2002. Unobserved ability and the return to schooling. Econometrica 70 (5), 2075–2091.
Bengoa, M., Rick, C., 2020. Chinese Hukou Policy and Rural-to-Urban Migrants' Health: Evidence from Matching Methods. Eastern Econ Journal 46, 224–259.
Bollini, P., Siem, H., 1995. No real progress towards equity: health of migrants and ethnic minorities on the eve of the year 2000. Soc. Sci. Med. 41 (6), 819–828.
Borrell, C., Muntaner, C., Benach, J., Artazcoz, L.A., 2004. Social class and self-reported health status among men and women: what is the role of work organisation, household material standards and household labour? Soc. Sci. Med. 58 (10), 1869–1887.
Burns, R., Zhang, C.X., Patel, P., Eley, I., Campos-Matos, I., Aldridge, R.W., 2021. Migration health research in the United Kingdom: a scoping review. Journal of Migration and Health 4, 100061.
Celik, Y., Hotchkiss, D.R., 2000. The socioeconomic determinants of maternal health care utilization in Turkey. Soc. Sci. Med. 50 (12), 1797–1806.
Chakraborty, T., Jayaraman, R., 2019. School feeding and learning achievement: Evidence from India's midday meal program. Journal of Development Economics 139, 249–265.
Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., Morton, S.C., Shekelle, P.G., 2006. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann. Intern. Med. 144 (10), 742–752.
Chen, J., 2011. Internal migration and health: Re-examining the healthy migrant phenomenon in China. Soc. Sci. Med. 72 (8), 1294–1301.
Chen, J., Park, A., 2021. School entry age and educational attainment in developing countries: evidence from China's compulsory education law. J. Comp. Econ. 49 (3), 715–732.
Cheng, Z., 2021. Education and consumption: evidence from migrants in Chinese cities. J. Bus. Res. 127, 206–215.
Cheng, Z., Guo, W., Hayward, M., Smyth, R., Wang, H., 2021. Childhood adversity and the propensity for entrepreneurship: a quasi-experimental study of the Great Chinese Famine. J. Bus. Ventur. 36 (1), 106063, 1–17.
Cheng, Z., Smyth, R., 2021. Education and migrant entrepreneurship in urban China. J. Econ. Behav. Organ. 188, 506–529.
Connelly, R., Roberts, K., Zheng, Z., 2012. The role of children in the migration decisions of rural Chinese women. J. Contemp. China 21 (73), 93–111.
Connelly, R., Zheng, Z., 2003. Determinants of school enrollment and completion of 10 to 18 year olds in China. Econ. Educ. Rev. 22 (4), 379–388.
Courtin, E., Kim, S., Song, S., Yu, W., Muennig, P., 2020. Can social policies improve health? A systematic review and meta-analysis of 38 randomized trials. Milbank Q. 98 (2), 297–371.
Cui, Y., Liu, H., Zhao, L., 2019. Mother's education and child development: evidence from the compulsory school reform in China. J. Comp. Econ. 47 (3), 669–692.
Cutler, D.M., Lleras-Muney, A., 2010. Understanding differences in health behaviors by education. J. Health Econ. 29 (1), 1–28.
Dai, F., Cai, F., Zhu, Y., 2022. Returns to higher education in China–evidence from the 1999 higher education expansion using a fuzzy regression discontinuity. Appl. Econ. Lett. 29 (6), 489–494.
Ding, X., 2021. College education and internal migration in China. China Econ. Rev. 69, 101649.
Eide, E.R., Showalter, M.H., 2011. Estimating the relation between health and education: what do we know and what do we need to know? Econ. Educ. Rev. 30 (5), 778–791.
Fang, H., Eggleston, K.N., Rizzo, J.A., Rozelle, S., Zeckhauser, R.J., 2016. The returns to education in China: evidence from the 1986 compulsory education law. In: Eggleston, K. (Ed.), Policy Challenges from Demographic Change in China and India. Brookings Institution Press, Washington, D.C.
Fergusson, D.M., McLeod, G.F., Horwood, L.J., 2015. Leaving school without qualifications and mental health problems to age 30. Soc. Psychiatr. Psychiatr. Epidemiol. 50 (3), 469–478.
Fu, H., Ge, R., Huang, J., Shi, X., 2022. The effect of education on health and health behaviors: evidence from the college enrollment expansion in China. China Econ. Rev. 72, 101768.
Giulietti, C., Wahba, J., Zenou, Y., 2018. Strong versus weak ties in migration. Eur. Econ. Rev. 104, 111–137.
Giulietti, C., Wahba, J., Zimmermann, K.F., 2013. Entrepreneurship of the left-behind. In: Giulietti, C., Tatsiramos, K., Zimmermann, K.F. (Eds.), *Labor Market Issues In China*: 65–92. Emerald Group Publishing Limited, Bingley.
Guo, F., Cheng, Z., Hugo, G., Gao, W., 2014. Wages and employment status of China's migrant workers. In: Wang, M.Y., Kee, P., Gao, J. (Eds.), *Transforming Chinese Cities*: 95–112. Routledge, London.
Hamad, R., Elser, H., Tran, D.C., Rehkopf, D.H., Goodman, S.N., 2018. How and why studies disagree about the effects of education on health: a systematic review and meta-analysis of studies of compulsory schooling laws. Soc. Sci. Med. 212, 168–178.
Hayward, M., Cheng, Z., Wang, B.Z., 2022. Disrupted education, underdogs and the propensity for entrepreneurship: evidence from China's sent-down youth program. J. Bus. Res.151, 33-39.
Heckman, J.J., Humphries, J.E., Veramendi, G., 2018. The nonmarket benefits of education and ability. J. Hum. Cap. 12 (2), 282–304.
Heckman, J.J., Li, X., 2004. Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000. Pac. Econ. Rev. 9 (3), 155–171.
Hout, M., 2012. Social and economic returns to college education in the United States. Annu. Rev. Sociol. 38 (1), 379–400.
Hu, X., Cook, S., Salazar, M.A., 2008. Internal migration and health in China. Lancet 372 (9651), 1717–1719.
Huang, B., Tani, M., Wei, Y., Zhu, Y., 2022. Returns to Education in China: Evidence from the Great Higher Education Expansion. China Economic Review, 101804.
Huang, B., Tani, M., Zhu, Y., 2021. Does higher education make you more entrepreneurial? Causal evidence from China. J. Bus. Res. 135, 543–558.
Huang, B., Zhu, Y., 2020. Higher Education Expansion, the Hukou System, and Returns to Education in China. IZA Discussion. *Paper No. 12954*.
Huang, W., 2015. Understanding the Effects of Education on Health: Evidence from China.
Iacobucci, D., Saldanha, N., Deng, X., 2007. A meditation on mediation: evidence that structural equations models perform better than regressions. J. Consum. Psychol. 17 (2), 139–153.
Jiang, W., Lu, Y., Xie, H., 2020. Education and mental health: evidence and mechanisms. J. Econ. Behav. Organ. 180, 407–437.
Kemptner, D., Jürges, H., Reinhold, S., 2011. Changes in compulsory schooling and the causal effect of education on health: evidence from Germany. J. Health Econ. 30 (2), 340–354.
Kim, Y.-J., 2016. The long-run effect of education on obesity in the US. Econ. Hum. Biol. 21, 100–109.
Kiviet, J.F., 2020. Testing the impossible: identifying exclusion restrictions. J. Econom. 218 (2), 294–316.
Lantz, P.M., Lynch, J.W., House, J.S., Lepkowski, J.M., Mero, R.P., Musick, M.A., Williams, D.R., 2001. Socioeconomic disparities in health change in a longitudinal study of US adults: the role of health-risk behaviors. Soc. Sci. Med. 53 (1), 29–40.
Li, B., Zhang, H., 2017. Does population control lead to better child quality? Evidence from China's one-child policy enforcement. J. Comp. Econ. 45 (2), 246–260.
Li, H., Ma, Y., Meng, L., Qiao, X., Shi, X., 2017. Skill complementarities and returns to higher education: evidence from college enrollment expansion in China. China Econ. Rev. 46, 10–26.
Li, J., Powdthavee, N., 2015. Does more education lead to better health habits? Evidence from the school reforms in Australia. Soc. Sci. Med. 127, 83–91.
Li, S., Whalley, J., Xing, C., 2014. China's higher education expansion and unemployment of college graduates. China Econ. Rev. 30, 567–582.
Lundborg, P., Lyttkens, C.H., Nystedt, P., 2016. The effect of schooling on mortality: new evidence from 50,000 Swedish twins. Demography 53 (4), 1135–1168.
McMullan, M., 2006. Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. Patient Educ. Counsel. 63 (1–2), 24–28.
Meng, X., Zhao, G., 2021. The long shadow of a large scale education interruption: the intergenerational effect. Lab. Econ. 71, 102008.

Messinis, G., 2013. Returns to education and urban-migrant wage differentials in China: IV quantile treatment effects. China Econ. Rev. 26, 39–55.
Nadkarni, A., Acosta, D., Rodriguez, G., Prince, M., Ferri, C.P., 2011. The psychological impact of heavy drinking among the elderly on their co-residents: the 10/66 group population based survey in the Dominican Republic. Drug Alcohol Depend. 114 (1), 82–86.
Nielsen, I., Smyth, R., 2008. The rhetoric and the reality of social protection for China's migrant workers. In: Nielsen, I., Smyth, R. (Eds.), *Migration And Social Protection in China*: 3–13. World Scientific, Singapore.
Oster, E., 2019. Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics 37 (2), 187–204.
Platform, L.L., Maletic, A., 2016. Integrating Refugees and Migrants through Education. European Civil Society for Education, Brussels.
Ross, C.E., Wu, C.-L., 1996. Education, age, and the cumulative advantage in health. J. Health Soc. Behav. 104–120.
Silles, M.A., 2009. The causal effect of education on health: evidence from the United Kingdom. Econ. Educ. Rev. 28 (1), 122–128.
Tang, C., Zhao, L., Zhao, Z., 2020. Does free education help combat child labor? The effect of a free compulsory education reform in rural China. J. Popul. Econ. 33 (2), 601–631.
Terraneo, M., 2015. Inequities in health care utilization by people aged 50+: evidence from 12 European countries. Soc. Sci. Med. 126, 154–163.
Van Der Heide, I., Wang, J., Droomers, M., Spreeuwenberg, P., Rademakers, J., Uiters, E., 2013. The relationship between health, education, and health literacy: results from the Dutch adult literacy and life skills survey. J. Health Commun. 18 (Suppl. 1), 172–184.
von Hippel, P.T., Lynch, J.L., 2014. Why are educated adults slim—causation or selection? Soc. Sci. Med. 105, 131–139.
Wang, H., Cheng, Z., Smyth, R., 2019. Health outcomes, health inequality and Mandarin proficiency in urban China. China Econ. Rev. 56, 101305.
Wang, H., Cheng, Z., Smyth, R., Sun, G., Li, J., Wang, W., 2022. University education, homeownership and housing wealth. China Econ. Rev. 71, 101742.
Wang, H., Cheng, Z., Wang, B.Z., Chen, Y., 2021. Childhood left-behind experience and labour market outcomes in China. J. Bus. Res. 132, 196–207.
Wang, Y., 2021. Who benefits more from the college expansion policy? Evidence from China. Res. Soc. Stratif. Mobil. 71, 100566.
Xiao, Y., Li, L., Zhao, L., 2017. Education on the cheap: the long-run effects of a free compulsory education reform in rural China. J. Comp. Econ. 45 (3), 544–562.
Xie, S., Mo, T., 2014. The impact of education on health in China. China Econ. Rev. 29, 1–18.
Xing, C., 2014. Migration, self-selection and income distributions: evidence from rural and urban China. Econ. Transit. 22 (3), 539–576.
Xu, R., Frank, K.A., Maroulis, S.J., Rosenberg, J.M., 2019. Konfound: command to quantify robustness of causal inferences. STATA J. 19 (3), 523–550.
Xue, X., Cheng, M., Zhang, W., 2021. Does education really improve health? A meta-analysis. J. Econ. Surv. 35 (1), 71–105.
Yao, Y., Chen, G.S., Salim, R., Yu, X., 2018. Schooling returns for migrant workers in China: estimations from the perspective of the institutional environment in a rural setting. China Econ. Rev. 51, 240–256.
Zajacova, A., Hummer, R., 2009. Gender differences in education gradients in all-cause mortality for white and black adults born 1906–1965. Soc. Sci. Med. 69 (4), 529–537.
Zhang, J., Zhao, Z., 2015. Social-family network and self-employment: evidence from temporary rural–urban migrants in China. IZA J. Labor Dev. 4 (1), 1–21.
Zhang, W., Chen, H., Feng, Q., 2015. Education and psychological distress of older Chinese: exploring the longitudinal relationship and its subgroup variations. J. Aging Health 27 (7), 1170–1198.
Zhao, X., Lynch Jr., J.G., Chen, Q., 2010. Reconsidering Baron and Kenny: myths and truths about mediation analysis. J. Consum. Res. 37 (2), 197–206.
Zhao, Y., 1999. Leaving the countryside: rural-to-urban migration decisions in China. Am. Econ. Rev. 89 (2), 281–286.
Zhou, Z., Fang, Y., Zhou, Z., Li, D., Wang, D., Li, Y., Lu, L., Gao, J., Chen, G., 2017. Assessing income-related health inequality and horizontal inequity in China. Soc. Indicat. Res. 132 (1), 241–256.
Zhu, N., 2002. The impacts of income gaps on migration decisions in China. China Econ. Rev. 13 (2), 213–230.