China Economic Review 88 (2024) 102293

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

The long-term cognitive and schooling effects of childhood vaccinations in China

Hamid R. Oskorouchi [a,c,*], Alfonso Sousa-Poza [a,c], David E. Bloom [b]

[a] Institute for Health Care & Public Management, University of Hohenheim, Germany
[b] Department of Global Health and Population, Harvard T.H. Chan School of Public Health, USA
[c] IZA, Germany

ARTICLE INFO

JEL:
I12
I18
I21

Keywords:
China
Vaccination
Cognitive skills
Education

ABSTRACT

By exploiting rich retrospective data from the China Health and Retirement Longitudinal Study on childhood immunization, socioeconomics, and health status, we assess the long-term effects of childhood vaccination on cognitive and educational outcomes in China. Applying an instrumental variable approach that resembles an unobserved natural experiment to different sets of control variables and subsamples, we estimate the average and local treatment effects of childhood vaccination. Our results confirm that immunization before the age of 15 has long-term positive and economically meaningful effects on non-health outcomes such as education and cognitive skills. These effects are strong, with vaccinated individuals enjoying about one additional year of schooling and performing better on several cognitive tests later in life. Finally, a causal mediation analysis shows that, although education mediates the effect of childhood immunization on later-life cognitive abilities, other factors (e.g., better child health) are more responsible for these long-term effects.

1. Introduction

Immunization benefits, such as disease prevention and herd immunity, yield high economic returns. A review of 108 studies from 51 low- and middle-income countries amply estimates that vaccinations avert costs of less than 1,000 USD per disability-adjusted life year (Ozawa, Mirelman, Stack, Walker, & Levine, 2012). In addition, a growing number of epidemiological, immunological, and clinical studies indicate positive vaccine effects well beyond the intended disease protection (Benn, Netea, Selin, & Aaby, 2013; Saadatian-Elahi et al., 2016), including heterologous nonspecific protection via immune system training or induction of cross-reactive T-cells (Andersen et al., 2018). Because these nonspecific effects may be relatively substantial, they are just as—or even more—important for child mortality than specific vaccine effects (Benn et al., 2013). Hence, in assessing the effects of childhood vaccination on cognitive and schooling outcomes in late adulthood, we define indirect effects by non-health status outcomes (i.e., education and cognitive abilities), regardless of whether induced via targeted disease prevention (specific effects) or immune system boosting (nonspecific effects). Childhood vaccinations primarily influence cognition and schooling through improved child health, which can raise educational outcomes through increased school attendance and achievement (Nandi et al., 2019). The latter, in turn, can enhance cognitive skills not only in childhood but also in later life, as the well-documented association between schooling and old-age cognitive outcomes suggests (e.g., memory disorders; Glymour, Kawachi, Jencks, & Berkman, 2008). A more direct effect of childhood health on

* Corresponding author at: Fruwirthstr. 48, 70599 Stuttgart, Germany.
E-mail address: hamid.oskorouchi@uni-hohenheim.de (H.R. Oskorouchi).

https://doi.org/10.1016/j.chieco.2024.102293
Received 18 April 2023; Received in revised form 7 October 2024; Accepted 15 October 2024
Available online 20 October 2024

cognitive skills relates to the detrimental effects of illness and stress on the hippocampus, which can negatively affect episodic memory performance in later life (Evans & Schamberg, 2009; Hassevoort, Khan, Hillman, & Cohen, 2016).

Our selected research setting of China provides a unique opportunity to assess childhood vaccination effects on later adulthood cognitive and educational outcomes because, before Mao and the Chinese Communist Party assumed state power in 1949 (commonly termed the "Liberation"), immunization was extremely limited, infectious diseases extremely serious and difficult to control (Liang & Liu, 2019), and preventive medicine almost nonexistent in most of the country (Sidel & Sidel, 1975). In 1950, however, the Ministry of Health oversaw a free smallpox vaccination campaign that achieved around 90 % coverage nationwide within three years, followed in 1953 and 1954 by child vaccination directives against diphtheria toxoid and tuberculosis, respectively, and accelerated research on vaccine development for all high-mortality infections. A key factor in the success of these campaigns was mass mobilization (World Bank, 1984) driven by the policy that medicine should serve the needs of workers, peasants, and soldiers alike (Sidel & Sidel, 1975), with preventive intervention prioritized over therapeutic. Before 1975, and particularly during the Cultural Revolution (1965–1975), childhood immunization among older adults in China was nowhere near universal (Liang & Liu, 2019). This uneven immunization coverage provides the treatment variation necessary to implement econometric techniques. The Chinese setting is doubly interesting because its older populations often score low in the cognitive ability inherently necessary for sound financial (and health) outcomes, while also lacking access to specialist advice on such essential topics (Lei, Smith, Sun, & Zhao, 2014). At the same time, the threat to eldercare provision posed by past fertility reductions and migration patterns that increasingly separate parents geographically from their adult children make the promotion of cognitive health in China particularly important (Smith, Strauss, & Zhao, 2014)—and cognitive health, as discussed, can be bolstered by increased vaccinations.

Our analysis, which draws on the China Health and Retirement Longitudinal Study (CHARLS), adds to the growing body of recent literature on the long-lasting effects of childhood vaccination on educational, economic, and cognitive outcomes in later adulthood (Atwood, 2022; Barteska, Dobkowitz, Olkkola, & Rieser, 2023; Bütikofer & Salvanes, 2020). Research on this topic is very limited for China; to our knowledge, our study is the first to examine the long-term impact of vaccines on education and various cognitive outcomes in later adulthood in China. Moreover, we are unaware of any research, even outside of China, that concentrates on outcomes among elderly individuals. Our analysis shows that vaccinations before the age of 15 have long-term positive and economically meaningful effects on non-health outcomes such as education and cognitive skills. These effects are strong, with vaccinated individuals enjoying about one additional year of schooling and performing better on several cognitive tests later in life. To measure these effects, we exploit a rich set of cognitive ability and schooling completion variables, as well as retrospective information on individual life histories, including vaccinations and socioeconomic and health status during childhood. In estimating the treatment effect of childhood vaccination, we apply an instrumental variable (IV) approach that resembles an unobserved natural experiment to different sets of control variables and subsamples. For everyone in the sample, we construct this IV by counting the number of vaccinated individuals for each year and province (first administrative unit) of birth, each time excluding from the count the birth district (secondary administrative unit) of the observation itself. Contrary to a standard spatial instrumental variable, our IV has limited bias potential because it proxies random variation in immunization campaigns by district. In fact, until the late 1970s, due to limited vaccine supplies, Chinese health authorities targeted their immunization campaign to a subset of districts within each province on a random rotating basis (Yu et al., 2018). That is, larger IV values proxy lower probabilities that an observation received childhood vaccination because the campaign rollout in its district occurred in a later stage.

2. Literature review

Whereas most literature on the effects of childhood vaccination focuses on child mortality rather than other outcomes, some recent studies do address the effects on cognitive and schooling outcomes. For example, in addition to inducing a 7.4 percentage point increase in male school enrollment probability in Bangladesh—albeit with no apparent effect for females (Driessen, Razzaque, Walker, & Canning, 2015)—age-appropriate measles vaccination improved school grades in South Africa and enhanced child anthropometry, cognition, and schooling outcomes in Ethiopia, India, and Vietnam relative to nonvaccinated children (Nandi, Shet, et al., 2019). Likewise, a study of the associations among these three outcomes and *Haemophilus influenzae* type B vaccination in India documents significantly higher scores for vaccinated children on English, mathematics, reading, the Peabody Picture Vocabulary tests, and school grades, and higher z-scores for height for age and body mass index for age (BMIZ; Nandi, Deolalikar, Bloom, & Laxminarayan, 2019). This same pattern is observable among Filipino children given full childhood vaccination against measles; polio; tuberculosis; and diphtheria, pertussis, and tetanus (DPT) (Bloom, Canning, & Shenoy, 2011). Nandi, Kumar, Shet, Bloom, and Laxminarayan (2020) show that schooling attainment among young adults improved (by about 0.25 grades) via exposure to India's Universal Immunization Programme, which was implemented between 1985 and 1990. Similarly, Anekwe, Newell, Tanser, Pillay, and Barnighausen (2015) show that immunization increases years of schooling of young South Africans (6–11 years old) by 0.20 grades.

Not only do the aforementioned findings underscore vaccination's potential long-term benefits for cognition and schooling, but a large body of economics literature documents the financial benefits of good childhood health (e.g., Case, Fertig, & Paxson, 2005; Smith, 2009). Specifically, Case et al. (2005) show that, even after controlling for parental socioeconomic characteristics, those who experience poor health as children not only have poorer health as adults but also have significantly lower educational attainment and lower socioeconomic status. Given the strength of these predominantly beneficial vaccine effects (both specific and nonspecific) on child health, such benefits could be expected to last the lifetime. Yet aside from the previously cited studies, we know of few analyses that assess the longer-lasting effects of vaccines, particularly those on non-health outcomes like cognitive ability and educational achievement.

Some recent studies, however, have analyzed the long-term effects of vaccines on human capital, in particular measles

vaccinations. Barteska et al. (2023) demonstrate that the United States' first large-scale measles vaccination campaign led to higher levels of educational achievement. Their research method utilizes the differences in measles exposure across various states just before the 1967–1968 Measles Eradication Campaign; this effort successfully decreased the reported cases of measles by 90 % within a two-year span. Chuard, Schwandt, Becker, and Haraguchi (2022) applied epidemiological modeling and a reduced-form economic method to assess the lasting effects of early-life exposure to measles, particularly around the time the measles vaccine was introduced in the U. S.. This approach highlighted significant enhancements in education, labor market participation, and health in adulthood for those born post-vaccine introduction, especially in states where measles mortality was previously high. Atwood and Pearlman (2023) show that mass measles vaccination in Mexico in 1973 led to improved long-run educational and labor market outcomes for adult men. Although van den Berg, von Hinke, and Vitt (2023) do not find long-term educational effects of measles vaccination programs in the United Kingdom, they do observe that exposure to the vaccination in early childhood positively affected adult height (but only among those with high genetic endowments for height). Daramola, Hossain, Kazianga, and Fogam (2022) examines the lasting impact of Burkina Faso's "vaccination commando program," initiated in 1984. This initiative successfully vaccinated more than a million children against diseases like measles, yellow fever, and meningitis within a fortnight. Their research indicates that this vaccination program markedly enhanced human capital among those who received the treatment: key findings include substantial reductions in child mortality, notable improvements in child health, and increased rates of primary school enrollment and completion. Furthermore, individuals who were vaccinated as part of this cohort show a higher likelihood of employment, particularly in the formal sector, in their adult years. Finally, Nandi et al. (2020) show that the Universal Immunization Programme in India—which included four childhood vaccines for infants—positively affected the long-term schooling of adults (aged between 20 and 49 years).

Empirical studies exploring the influence of vaccines on human capital in China are rare. Among the few is the research by He, Huang, and Wang (2022). They utilized data from the 2010 China Family Panel Studies and a cohort difference-in-differences approach to assess the effects of a meningitis vaccination initiative from the 1980s on human capital development and economic progress. Their findings indicate a notable enhancement in the long-term accumulation of human capital among those who received the vaccine. Specifically, the study revealed that vaccination led to an average increase of approximately 0.5 years in the educational attainment of the vaccinated group. Furthermore, it increased the likelihood of completing primary, junior high, and high school education by 3.1 %, 10 %, and 22.4 %, respectively.

The aforementioned studies focus on educational (and other) outcomes in adulthood. We are not aware of studies that assess long-term effects that last into late adulthood. Furthermore, no study to date has assessed the possible long-term *cognitive* effects of vaccines. Although this lack of empirical evidence on vaccination's long-term effects no doubt relates to the methodological challenges of identifying causal relations between childhood vaccines and later-life outcomes, the randomized trials that should ideally be used to assess such effects are often very difficult to implement, especially many years after the original exposure (Barnighausen, Bloom, Cafiero-Fonseca, & O'Brien, 2014). Even the potential alternative of observational studies may suffer from selection bias if vaccine recipients differ from nonrecipients in ways that relate to the outcome variable independently of vaccination (Fine et al., 2009), prompting several researchers to mitigate this problem by using quasi-experimental propensity score matching (PSM; e.g., Nandi, Shet, et al., 2019; Nandi, Deolalikar, et al., 2019; Bloom et al., 2011). Another method for avoiding selection bias is to exploit natural experiments such as the introduction of universal vaccination programs, which guarantee that access to vaccines is independent of such bias-inducing factors as household wealth, parental education, and health insurance coverage (Andersen et al., 2018). An additional challenge when assessing vaccination's long-term implications is the widespread unavailability of precise childhood vaccination data for older adults. Our analysis thus makes a valuable contribution to the literature, not only by being one of the first to document the long-lasting effects of childhood vaccination on educational and cognitive outcomes, but also by applying a combination of techniques to limit possible biases. Our application of IV regressions addresses nonrandom selection into treatment and possible measurement errors.

3. Materials and methods

3.1. Data

We draw data from CHARLS, a longitudinal survey representative of the Chinese population over age 45, which—in addition to basic demographics—collects information on the socioeconomic determinants of aging, including physical and physiological health. Our analysis uses the latest standard CHARLS wave (2018) and a 2014 retrospective wave that provides family information; work history; and data on childhood education, health, and wealth status. Note that, due to the time-invariant nature of the treatment variable (childhood vaccination) and of the full set of controls, exploiting CHARLS' longitudinal nature is impossible. Moreover, we perform our analyses using the latest waves of CHARLS to maximize the sample size, but, as a robustness check, we also replicate the analysis using CHARLS wave 1. In the main analysis of this study, we retain only the observations for those aged 80 years old or younger (in 2018), but we report the results using the full sample in Appendix A. Excluding the so-called "super healthy" individuals from the sample is common in studies on elderly populations (c.f. Abeliansky & Strulik, 2018) to limit selective mortality–related biases. In fact, especially in the case of China, individuals older than age 80 in our sample surpassed their life expectancy at birth.

Dependent variables. To estimate the long-term effects of childhood vaccination on cognitive abilities and education, we exploit years of schooling and numeracy, mental status, and episodic memory as reflected by cognitive ability measures *(see* Table B1 *in Appendix B for the corresponding survey items)*. Specialized personnel collect these cognition measures, which match those used in internationally comparable surveys of older adults (e.g., the Survey of Health, Ageing and Retirement in Europe, also known as SHARE), with individual daily life reasoning abilities captured by episodic memory in the form of immediate and delayed recall. For

these two tests, an enumerator reads a list of 10 Chinese nouns and then counts how many words on the list an individual can recall immediately and then four minutes later. For our study, we transform these measures from counts to shares of correctly recalled nouns out of the total list (with the corresponding effect interpreted as the probability of correctly recalling all words immediately and four minutes later).

The second set of CHARLS cognitive ability measures, which comprises items from the Telephone Interview of Cognitive Status study, proxies mental status by the ability to (i) name the day's date (year, month, day, and day of the week), (ii) redraw a picture, and (iii) count down from 100 by sevens up to five times (serial 7 subtraction, as a proxy for numeracy). Our econometric models use aggregated numeracy and mental status measures representing the number of successfully accomplished tasks over the entire cognitive ability test (c.f. McArdle, Smith, & Willis, 2011). Lastly, because CHARLS specifies only a categorical variable for the highest educational level achieved rather than exact years of formal education, we compute years of schooling by converting this level into actual years (c.f. Molina, 2016). Table 1, which presents the characteristics of both the restricted (without migrants and individuals who are more than 80 years old: columns 1–3) and the full sample (columns 4–6), reports that an average individual has 4.85 years of schooling and numeracy, mental status, and episodic memory scores of 0.53, 0.69, and 0.32, respectively.

Main independent variable. The treatment in our models is childhood vaccination (i.e., before age 15), determined based on a corresponding item in the 2014 CHARLS retrospective wave, coded as one for vaccinated individuals (before age 15) and zero for unvaccinated; the latter comprise 16 % of our sample (see Table 1). Unfortunately, because CHARLS provides no information on vaccination type and exact timing, we cannot construct variables that proxy vaccine-specific exposure, which would have allowed us to disentangle the biological mechanisms underlying our results. Nonetheless, knowing the timeline of vaccine introduction in China and exploiting the representativeness of our sample for older adults, we can narrow the pool of vaccines responsible for any positive effects on education and cognitive abilities. That is, given China's vaccine licensure timeline (Zheng et al., 2018) and our youngest respondent's 1970 birth year, the treatment group primarily received the antitubercular Bacillus Calmette-Guérin (BCG, 1937); vaccines against plague (1946), yellow fever (1954), poliovirus (1961), and measles (1965); and possibly the combined DPT immunization (1973). Of these vaccines, epidemiological evidence for beneficial nonspecific effects has been widely documented for measles and BCG (Benn et al., 2013). The measles vaccine stands out as a highly effective public health measure, not least because the vaccine plays a vital role in decreasing illness and death rates from various other pathogens. This feature is linked to the specific nature of the measles virus and its effect on the immune system: contracting measles leads to severe immunosuppression, making individuals more vulnerable to other infections, a condition known as "immune amnesia" (Atwood, 2022; Atwood & Pearlman, 2023; Benn et al., 2013). Available epidemiological data also indicate that the BCG vaccine might have significant positive effects beyond its specific use, especially on the immune system (Roth, Garly, Jensen, Nielsen, & Aaby, 2006). Specifically, observational studies have shown a correlation between early BCG vaccination and a decrease in mortality rates. Not every vaccine program potentially impacting our study has demonstrated advantageous nonspecific effects. Unlike the BCG and measles vaccines, which show positive effects, the DTP vaccine appears not to share these benefits and may even produce contrary effects (Benn et al., 2013).

Control variables. We select predictors of childhood vaccination probability with a vector that includes demographic and childhood socioeconomic measures and parental characteristics (see Table B2 for the corresponding CHARLS life history questions). These variables, for which Table 1 details the descriptive statistics, include dummy indicators for parents being illiterate and alive (at the time of the survey), the subject's gender, belonging to an ethnic minority, childhood poverty, age or cohort dummies, and dichotomous variables identifying the community of birth or of residence, depending on migration history. One limitation of CHARLS is that it does not provide specific information on the actual province/district of birth but only reports if an individual has moved from this province/

Table 1
Descriptive Statistics.

	N	Mean	SD	N	Mean	SD	Min	Max	Difference
	Age-restricted, no migrants			*Full sample*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Education (years)	10,580	4.85	4.15	15,884	4.92	4.10	0	16	−0.07
Numeracy	10,155	0.53	0.39	15,071	0.54	0.40	0	1	−0.01
Mental status	9682	0.69	0.30	14,377	0.71	0.30	0	1	−0.02
Episodic memory	9910	0.32	0.18	14,682	0.33	0.18	0	1	−0.01
Vaccination (yes = 1)	10,580	0.84	0.37	15,884	0.85	0.36	0	1	−0.01
IV	10,110	0.85	0.14	15,070	0.84	0.15	0	1	0.00
CR-IV	10,580	2.24	2.33	15,884	1.93	2.32	0	8	0.31
Mother illiterate (yes =1)	10,580	0.87	0.33	15,884	0.85	0.36	0	1	0.03
Father illiterate (yes = 1)	10,580	0.63	0.48	15,884	0.59	0.49	0	1	0.04
Mother alive (yes =1)	10,580	0.22	0.41	15,884	0.25	0.43	0	1	−0.03
Father alive (yes = 1)	10,580	0.11	0.32	15,884	0.14	0.35	0	1	−0.03
Poor in childhood (yes = 1)	10,580	0.41	0.49	15,884	0.40	0.49	0	1	0.01
Age (years)	10,580	61.66	8.18	15,884	61.07	9.96	45	(80 | 105)	0.58
Gender (male = 1)	10,580	0.49	0.50	15,884	0.49	0.50	0	1	0.01
Starved in childhood (yes = 1)	10,580	0.83	0.37	15,884	0.82	0.38	0	1	0.01
Ethnic minority (yes = 1)	10,580	0.08	0.28	15,884	0.08	0.27	0	1	0.00

Notes: Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. SD: standard deviation; IV: instrumental variable for vaccine status; CR-IV: instrumental variable for years of education.

district.

3.2. Methods

The challenge in disentangling childhood vaccinations' long-term nonspecific effects on later-life outcomes is that the parental attributes, of which childhood vaccination status is a function, might also determine better human and health capital in adulthood. Specifically, because the treatment (i.e., vaccination) is not randomly assigned, a naïve ordinary least squares (OLS) model would produce upwardly biased estimates of the vaccination effect on the outcome variables studied. To avoid this bias, we estimate the effect of early-life vaccination on education and later-life cognitive status by employing an OLS estimator and a conservative set of control variables that includes community-of-birth and cohort dummy variables, among others. Alternatively, we estimate this treatment effect by relying on the same control variables and year-province of birth dummies within a Two Stage Least Square (2SLS) estimator. Both approaches ensure the potential outcome's conditional independence by minimizing prevaccination background differences between the treatment and control groups to enable the estimation of a causal effect of childhood vaccination on human capital and later-life cognition.

For all outcome variables, we estimate the effect of childhood vaccination ($Vaccine_i$) on education and later life cognitive abilities (Y_i) employing a bivariate OLS regression, and then we gradually include meaningful control variables. Eqs. 1–2 outline the analytical formulas of the bivariate and multivariate OLS specifications, where Y_i alternatively represents one of the selected outcome variables (i.e., year of education, numeracy, mental status, or episodic memory); CV_i is a vector of individual-level controls that includes dummies for gender, ethnic minority belonging, mother and father being literate and alive, self-reported childhood poverty status, and having suffered hunger during childhood; and ε_i is an idiosyncratic error term. In Eq. 3, we additionally control for both province $\left(\xi_p\right)$ and year-of-birth (ϕ_t) fixed effects and their interaction (where the subscripts p and t refer to the province and the year of birth, respectively). In Eqs. 1–3, ε_i, e_i, and q_i represent idiosyncratic error terms.

$$Y_{i,p,t} = \beta_0 + \beta_1 Vaccine_{i,p,t} + \varepsilon_{i,p,t} \tag{1}$$

$$Y_{i,p,t} = \rho_0 + \rho_1 Vaccine_{i,p,t} + CV'_{i,p,t}\,\delta + e_{i,p,t} \tag{2}$$

$$Y_{i,p,t} = \sigma_0 + \sigma_1 Vaccine_{i,p,t} + CV'_{i,p,t}\,\lambda + \xi_p + \phi_t + \xi_p{}^{*}\phi_t + q_{i,p,t} \tag{3}$$

Finally, in Eq. 4, we estimate a fully saturated model which controls for community-of-birth (c) fixed effects (ψ_c) together with the full set of individual-level control variables and year-of-birth fixed effects (ζ_t), and $w_{i,p,t}$ is an idiosyncratic error term. It should be noted that because of multicollinearity, we are not able to include interaction effects between community- and year-of-birth dummies.

$$Y_{i,p,t} = \tau_0 + \tau_1 Vaccine_{i,p,t} + CV'_{i,p,t}\,\nu + \psi_c + \zeta_t + w_{i,p,t} \tag{4}$$

3.3. Potential sources of bias

In addition to the aforementioned nonrandom selection into treatment, the childhood vaccination coefficient could be biased if errors occur in the reporting of vaccination status. In fact, those with lower episodic memory and mental abilities are presumably more likely to forget childhood events. If the low-ability individuals inaccurately report having been vaccinated, they will be (mistakenly) included in the treatment group, lowering the average observed outcome and thus causing a downward bias. Similarly, when individuals with lower cognitive abilities wrongly report a negative vaccination status, they would belong to the control group. In this latter case, the measurement error will also translate into a downward-biased coefficient if one assumes that, on average, immunized individuals enjoy better health and education. However, extant research on SHARE, the U.S. Health and Retirement Study, and the U.S. Panel Study of Income Dynamics shows that the quality of subjective assessments of older people regarding their childhood conditions and health status is generally high (Havari & Mazzonna, 2015; Smith, 2009).

Another source of bias in our estimate is the presence of selective mortality (survivor bias). Because healthier individuals are more likely to survive (regardless of vaccination status), our sample likely underrepresents individuals of lower health status. If these individuals are not systematically distributed between the treatment and control groups, the vaccination estimate would not suffer from any survivor bias. However, children enjoying a better family environment (e.g., higher parental education and income) are more likely both to survive and to get vaccinated, resulting in a selected treatment group that causes overestimation of the vaccination effect. The same directional bias should be expected if the immunization campaigns targeted the most vulnerable children, who, despite being vaccinated, do not survive to be included in the survey.

Finally, if a vaccinated child is more likely to live in a community with higher immunization coverage, the childhood vaccination coefficient cannot be interpreted as causal but will be confounded by the indirect effect of the child's peers' immunization status. However, this bias likely understates our estimates because higher immunization coverage would decrease infection risk disproportionally between the treatment and control groups—specifically, in favor of the nonvaccinated children (control group).

3.4. Instrumental variable approach

To address these issues, we estimate our model of childhood vaccination effects on late-life outcomes using a *leave-d-y-out*

instrumental variable approach (where *d* and *y* refer to each observation's district and year of birth, respectively) that exploits the pre-1978 within-province random rollout of vaccinations in China. We construct our IV as the headcount ratio of vaccinated individuals by province of birth and cohort, each time excluding the observations belonging to the same district-cohort from this count.[1]

Eqs. (5) and (6) present the first- and second-stage regressions of our IV identification strategy, respectively. In Eq. (5), we predict individual vaccination status (*Vaccine*) by means of an OLS regression using the instrumental variable (*IV)*; all sociodemographic, child, and parental characteristics (CV) presented in Table 2; and both province (π, Γ) and year-of-birth (χ, Λ) fixed effects. u_i and s_i are individual idiosyncratic error terms. In all OLS models, we compute cluster-robust standard errors at the community level (449 communities) and correct the estimates for nonresponse by employing cross-sectional probability weights. Additionally, we account for the bias in the standard errors of the 2SLS models by correcting the standard errors at the community-of-birth level. Finally, Eq. (6) depicts our second-stage OLS regression.

$$Vaccine_{i,d,p,t} = \theta_0 + \theta_1 IV_{d,p,t} + CV'_{i,d,p,t}\kappa + \pi_p + \chi_t + u_{i,d,p,t} \tag{5}$$

$$Y_{i,d,p,t} = \mu_0 + \mu_1 \widehat{Vaccine}_{i,p,t} + CV'_{i,d,p,t}\, \iota + \Delta_p + \Lambda_t + s_{i,d,p,t} \tag{6}$$

We apply this IV strategy to different samples. In our main specification, we restrict the sample to individuals who never migrated from their community of birth (89.42 % of the total). As a robustness check, we also estimate the same IV model using the full sample, including migrant individuals.

Eq. (7) shows the construction of this instrument, where *IV* represents the instrumental variable; *i* is an individual born in year *t* and in district *d* of province *P*; *j* is an individual born in the same year and province *P* of *i,* but from a different district $k \neq d$; $w_{j,k,p,t}$ and $Vaccine_{j,k,p,t}$ are the vectors of individual (cross-sectional) survey probability weights and vaccination status, respectively; and $N_{k,p,t}$ and $N_{D,p,t}$ are the number of individuals born in a specific district, province, and year, and the total number of districts in province *p*.

$$IV_{d,p,t} = \frac{\sum_{k=1}^{N_{D,p,t}} \sum_{j=1}^{N_{k,p,t}} w_{j,k,p,t}\, Vaccine_{j,k,p,t} - \sum_{i=1}^{N_{d,p,t}} w_{i,d,p,t}\, Vaccine_{i,d,p,t\_}}{\sum_{k=1}^{N_D} \sum_{j=1}^{N_{k,p,t}} w_{j,k,p,t} - \sum_{i=1}^{N_{d,p,t}} w_{i,d,p,t-}} \tag{7}$$

Put simply, the numerator of eq. 7 represents the weighted headcount of vaccinated children, born within a specific year and province, subtracted by an equivalent weighted count computed for child's *i* district of birth. In parallel, the denominator in Eq. 7 denotes the difference between the number of children born in province *P* and year *t* (regardless of their vaccination status) and the identical child count from child's *i* district of birth *D* in year *t*. Thus, the IV embodies a headcount ratio of vaccinated children hailing from province *P* and year *t*, excluding any observations from the district of birth of child *i*.

Because CHARLS does not provide a detailed immigration history module, we rely only on the observations of individuals who did not migrate from their community of birth. Doing so prevents any migration-related measurement error in the IV construction. The proposed instrumental variable not only remains unrelated to each observation's individual characteristics, but it also retains the ability to address misreporting in vaccination status. In fact, this measurement error at the broad aggregate level is likely not correlated with unobservable individual characteristics.

Specifically, our IV proxies random spatial and temporal variation in the unobserved immunization campaign rollouts. In fact, between 1946 and 1949, vaccines gradually became cost-free, and several mass smallpox and cholera vaccination campaigns targeting both urban and poor rural households took place (Yu et al., 2018). Until 1978, due to the lack of a proper cold chain system, only smallpox immunization campaigns were synchronously and universally conducted at the national level because of the relative ease of transport and storage of this vaccine, which could remain effective at relatively high temperatures (up to 8 °C; Yu et al., 2018; WHO, 2005). Cholera and plague vaccines were only used for epidemic control, meaning that no routine vaccination campaigns were offered against these diseases (Yu et al., 2018). In contrast, each provincial department remained responsible for the organization of campaigns of other antigens (BCG, diphtheria, pertussis, polio, and measles; Yu et al., 2018). These campaigns were conducted during the colder winter months according to a "rush-relay" approach: vaccine transportation to the periphery and administration activities were usually completed within a few days (Yu et al., 2018). However, limited vaccine supplies prevented province-wide annual immunization campaigns. Instead, only a randomly selected share of districts within each province was targeted on a rotating basis over a period of five to six years (Yu et al., 2018).

Because we do not have exact information on these rollouts, our IV resembles this unobserved natural experiment, where the proposed IV functions as a proxy for the actual "hidden" exogenous instrument (*Z*).[2] That is to say, individuals who were born in year *t* within district *d* of province *P* are likely to have a diminished probability of being targeted by an immunization campaign when the vaccination coverage in all other districts k within the same province *P* is comparatively higher. Fig. 1 shows the causal direct acyclic graph (solid lines) of the reduced form of our IV strategy, where $Y_{i,k,p,t}$ and $IV_{i,d,p,t}$ represent the individual-level later-life outcome and the *leave-d-y-out* headcount ratio of vaccinated individuals born in a given year *t* and district *k*, respectively. As depicted by the

[1] Figure A1 presents the distribution of birth years for the CHARLS wave 4 sample.

[2] We use the term "hidden" instrument to denote the unobserved randomized vaccination rollout. In fact, although this variable would serve as the optimal candidate for our 2SLS strategy, it remains unobserved. Consequently, we are compelled to utilize an IV that serves as a proxy for this unobserved randomized rollout.

Table 2
First-Stage Regressions, Vaccine-IV (OLS).

	Dependent Variable			
	Vaccination Status			
	(1)	(2)	(3)	(4)
IV	−0.143***	−0.152***	−0.149***	−0.147***
	(−4.32)	(−4.47)	(−4.31)	(−4.25)
Mother illiterate	−0.019*	−0.017	−0.014	−0.018
	(−1.72)	(−1.55)	(−1.22)	(−1.58)
Father illiterate	−0.046***	−0.045***	−0.044***	−0.044***
	(−5.96)	(−5.88)	(−5.68)	(−5.73)
Mother alive	0.007	0.006	0.005	0.004
	(0.69)	(0.64)	(0.53)	(0.37)
Father alive	0.009	0.009	0.010	0.010
	(0.75)	(0.78)	(0.86)	(0.86)
Poor in childhood	−0.013*	−0.012	−0.012	−0.014*
	(−1.84)	(−1.57)	(−1.61)	(−1.87)
Gender (male = 1)	0.049***	0.050***	0.048***	0.049***
	(6.98)	(7.04)	(6.66)	(6.76)
Hunger in childhood	0.051***	0.052***	0.051***	0.049***
	(5.22)	(5.25)	(5.09)	(4.96)
Ethnic minority	−0.016	−0.013	−0.008	−0.009
	(−1.04)	(−0.85)	(−0.53)	(−0.56)
Sample	Education	Numeracy	Mental Status	Episodic Memory
N	10,110	9707	9252	9476
F-stat (partial)	18.77	20.14	18.72	18.21

Notes: Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. Partial F-statistics compares full vs. restricted model. The sample excludes individuals older than 80 years and within-province migrants. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. IV: instrumental variable for vaccination status.

bidirectional dashed arrows, within-district immunization levels and individual outcomes are assumed to be endogenous and the headcount ratios among districts to be correlated, but no feedback loop exists between them. The variable Z is supposed to be exogenous and the only unobserved determinant of cross-group variation. Thus, the proposed IV strategy remains valid only under the following two identifying assumptions: (i) the first-stage regression must not suffer from reverse causality, and (ii) the conditional independence assumption between the hidden instrument (Z) and the error term (u) must hold ($E\left[u_{i,d,p,t}|Z_{y,k,p,t}\right] = 0$) (Sundquist, 2021).

First, we judge first-stage reverse causality due to spatial spillovers to be very unlikely, especially when excluding migrant individuals from the sample. In fact, it is plausible to assume that districts with higher immunization levels each year do not affect the vaccination headcount ratio of neighboring communities over the same period when their hidden instrument Z is switched off (i.e., the district is not eligible for treatment in a specific year). That is, due to the enforcement of random rotation in districts' eligibility, the potential for the "always-takers" to be treated when not eligible remains minimal. Similarly, no existing mechanism justifies a potential spatial spillover from lower to higher immunized areas.

Second, we argue that, after controlling for all relevant observables, cross-group variation in treatment levels can be attributed to the sole random variation in the instrument and to no other confounding factors. Indeed, our IV leverages the variation in D to serve as a proxy for the hidden instrument Z (i.e., the unobserved province-level randomized vaccination rollout), and it is imperative that there are no other unobserved determinants of vaccination status that are common to all observations within each district. If this were not the case, the variation in D could inadvertently proxy some endogenous vaccination determinants in addition to the randomized rollout (U_d), thereby inducing bias in the 2SLS estimates. Such a violation of the 2SLS identifying assumption could occur if, for instance, vaccination rollouts were disrupted or amplified by regional events that would have affected entire districts or provinces. In response to this critique, we underscore that our 2SLS specification directly controls for heterogeneity in the province of birth and for year-of-birth trends. This extensive set of dummy variables concurrently absorbs the potential confounding effect of time-invariant heterogeneity in the quality of institutions at the lowest possible administrative level and of changes over time in national-level socioeconomic status.

A falsification test of the IV effect on education and cognitive abilities, using only unvaccinated individuals (presented in the subsequent robustness section), furnishes additional evidence for instrument exogeneity. In fact, if our IV were correlated with other determinants of childhood vaccination status, it would have yielded a statistically significant effect even when employing a sample of unvaccinated individuals.

Our IV is not only highly correlated with individual vaccination status (partial F-statistics always larger than 10), but it also enjoys high variability (mean = 0.85, standard deviation = 0.13) due to its year-of-birth- and district-specific construction. Table 2 presents the results of the first-stage sample-specific regressions. A likelihood ratio test comparing the restricted with the unrestricted model shows partial F-statistics always above the threshold of 10. Moreover, as expected, the IV has a negative effect on the probability of an individual being immunized in childhood.

Admittedly, however, a *leave-d-y-out* instrumental variable could be endogenous in the presence of spatial dependence, such as when the vaccination status of those born in a specific province affects the outcome of the observations for which the specific IV value

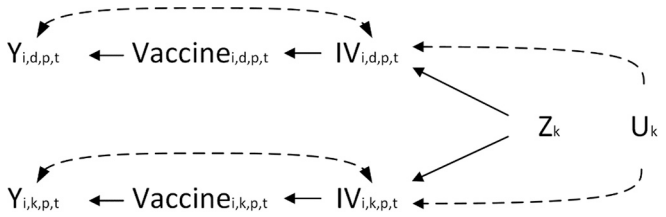

Fig. 1. IV Strategy Direct Acyclic Graph.
Notes: Graph adapted by the authors from Sundquist (2021); solid lines represent the causal path and dashed lines highlight potential violations.

is constructed. As already mentioned, in our setting, this would translate into indirect treatment of both control and treatment group observations (herd immunity spillovers), which would likely cause an underestimation of the true effect if unvaccinated children were more likely to benefit from herd immunity. To limit this possibility, we not only alternate control for various administrative units of birth dummy variables (i.e., province and community), but also construct our IV by excluding the individuals who were born (and reside) in the same district of each CHARLS observation from the IV count.

3.5. Mediation analysis

Although evidence that childhood vaccination improves cognitive abilities is ample (Bloom et al., 2011; Bloom, Canning, & Weston, 2017), how much education mediates this effect in later life is less well understood. We thus implement a causal mediation analysis to disentangle how child immunization translates into higher cognitive capacity via education. Note that this analysis can be interpreted as causal only if there are no unmeasured treatment/outcome and mediator/outcome confounders given the set of control variables. As for the analysis of the effect of childhood immunization, education would be endogenous because of positive selection into treatment caused by parental investment in health and education. The omission of a proxy for innate abilities would additionally confound this relationship. Because these omitted factors positively relate to the mediation analysis outcome variables, we expect a significant upward bias in models that do not instrument education.

To address this double source of endogeneity, we follow the linear causal mediation control function (CF) approach outlined in Frölich and Huber (2017), along with two exogenous instrumental variables. As in the main analysis, we instrument childhood immunization status with a *leave-d-y-out* immunization headcount ratio computed at the district and cohort levels. Conversely, we construct a suitable IV for education by exploiting the 1966–1976 Cultural Revolution (CR) school closures in China. In fact, during the CR era, China underwent several school interruptions: in 1966–1968, education was completely halted; then, in 1968, although primary and middle schools reopened, they completely abandoned the original standardized curricula. Finally, when in 1971 high school admissions resumed, the high school curriculum lasted only two years instead of four as in the original system (which was recovered at the end of the CR in 1976; Chen, Guo, Huang, & Song, 2019). Extant research has employed this quasi-random event to address endogeneity in studying the intergenerational transmission of education (Chen et al., 2019), how school interruptions affect subsequent educational achievements (Meng & Gregory, 2002), and the relationship between rustication and adulthood health (Fan, 2016), among others. Pepper (2000) provides a detailed discussion and analysis of the historical records of school closures during the CR period.

In the present analysis setting, we construct a CR-IV for education by counting the expected years of lost education for the cohorts born between 1948 and 1961 (i.e., school-age individuals during the CR) as reported in Chen et al. (2019); see Tables A.1 (p. 526) and A.2 (p. 527) for urban and rural cohorts, respectively. Here, we formally outline our mediation analysis' empirical strategy, where the following Eqs. (8) and (9) represent the first stage regressions and the endogenous exposure (vaccine) and mediator (educ) are regressed on an exogenous IV along with the full set of controls and province and year-of-birth fixed effects, respectively. Eq. (10) estimates the effect of the exposure on the mediator and, finally, Eq. (11) regresses the outcome variable (one of the proxies of cognitive abilities) on the mediator and the exposure variables together with their interaction while controlling for the first-stage error terms ($\widehat{u}$ and $\widehat{v}$) from (8) and (9). We then compute the control direct, natural indirect, and total effects applying the formulas reported in Valeri and VanderWeele (2013) by means of the *paramed* Stata package (Emsley & Liu, 2013).

$$Vaccine_{i,d,p,t} = \omega_0 + \omega_1 IV_{d,p,t} + CV'_{i,d,p,t}\, g + \Lambda_p + M_t + u_{i,d,p,t} \tag{8}$$

$$Educ_{i,p,t} = \delta_0 + \delta_1 CR\_IV_{i,p,t} + CV'_{i,p,t}\, h + \Sigma_p + P_t + v_{i,p,t} \tag{9}$$

$$Educ_{i,p,t} = \gamma_0 + \gamma_1 \widehat{v}_{i,p,t} + \gamma_2 Vaccine_{i,p,t} + CV'_{i,p,t}\, c + \Psi_p + \Omega_t + \eta_{i,p,t} \tag{10}$$

$$Y_{i,p,t} = \alpha_0 + \alpha_1 \widehat{u}_{i,p,t} + \alpha_2 \widehat{v}_{i,p,t} + \alpha_3 Vaccine_{i,p,t} + \alpha_4 Educ_{i,p,t} + \alpha_5 Vaccine_{i,p,t} * Educ_{i,p,t} + CV'_{i,p,t}\, r + \Phi_p + T_t + j_{i,p,t} \tag{11}$$

Table A1 presents the results of the first-stage regression of education on the CR-IV and shows the statistical relevance of the latter variable (partial F-statistics always above 10). However, the IV exclusion restriction is questionable if one assumes that the CR has had effects beyond educational outcomes. We address this issue by explicitly controlling for cohort-level dummies.

4. Results

Tables 3 and 4[3] show the results for the effect of childhood vaccination on schooling and cognitive abilities among older Chinese. For each relationship, we present the vaccine coefficient estimated by a simple bivariate OLS (column 1), and then we gradually include individual-level control variables (column 2) and dummies for either province, year-of-birth community, and their interaction (column 3), or community-of-birth and cohort dummy variables (column 4). Finally, column 5 of each table reports the 2SLS estimates. As expected, if children of richer and better-educated parents are, on average, more likely to be vaccinated, the effect of childhood vaccine exposure on education and later-life cognitive abilities declines (in absolute value) as we move from a bivariate regression to a more conservative multivariate specification. For example, the association between childhood immunization and years of schooling is 2.4 years in the simple bivariate OLS (Table 3 top, column 1) and only 0.87 years in the most conservative model (column 4). The results for the 2SLS regression (column 4) show that the causal effect of child immunization on education is 1.23 years of schooling. Similarly, the naïve OLS models on the vaccination/cognitive abilities nexus (Tables 3 and 4, columns 1–3, and 6–8) produce upward-biased estimates, which substantially decrease in magnitude when controlling for community-of-birth and cohort dummies (columns 4 and 9 in both Table 3 and Table 4). The corresponding 2SLS models (column 5) yield similar point estimates of the conservative OLS specification. Despite the sizable reduction in treatment magnitude, all treatment estimates remain statistically significant and economically meaningful. In fact, according to our 2SLS models, childhood vaccination increases later-age numeracy, mental status, and episodic memory by 8.6, 6.4, and 3.9 percentage points, respectively.

Table 5 shows the results of the causal mediation analysis of education for the relationship between childhood immunization and later-life cognitive abilities. Column 1 reports the controlled direct and natural indirect effects for three proxies of cognitive abilities (i.e., numeracy, mental status, and episodic memory) obtained from a mediation estimation strategy that only addresses endogeneity in the childhood immunization variable. These results indicate that the education mediator can explain more than a third of the vaccination–cognitive abilities nexus total effect (43 %, 39 %, and 41 % for numeracy, mental status, and episodic memory, respectively). However, when the same model additionally instruments education using a linear CF approach that exploits the CR-IV outlined previously (column 3), the share of indirect over total effect reduces to 28 %, 20 %, and 17 % for numeracy, mental status, and episodic memory, respectively. Such a reduction in the natural indirect effect of education is unsurprising, as this regressor strongly relates to parental investment and innate ability—unobserved factors in our analysis. However, these smaller mediating effects not only remain strongly economically meaningful, but they also signal that much of the long-term childhood immunization benefits translate into better later-life cognitive abilities via other paths in addition to formal education. Assuming that improved child health is the most important mediating factor is plausible. Unfortunately, CHARLS data lacks any objective child health proxy that would allow us to quantify the contribution of this mediator.

4.1. Robustness checks

We provide seven robustness and sensitivity tests to assess the credibility of our main result (Tables 3 and 4). Specifically, we alternatively use (i) a sample that includes individuals older than 80: (ii) domestic migrant individuals; (iii) a different cross-sectional CHARLS wave; or (iv) a regression that controls for a continuous measure of age instead of cohort dummies; we also (v) test the exclusion restriction of the IV for childhood immunization via a falsification test; and (vi) the estimates' sensitivity by looking at coefficient stability relative to R^2 movements; finally, (vii) we compute IV robustness sensitivity statistics to unobserved confounders.

First, we replicate the main analysis including the "super healthy" (above age 80; Tables A6 and A7). This demographic group likely enjoys a better genetic endowment and, simultaneously, has been less exposed to vaccination campaigns during childhood due to its average older age. Thus, its inclusion in the sample could produce understated treatment effects of childhood immunization on cognitive abilities. Conversely, because this same group's ability to access formal education was low (3.4 years of schooling on average), its inclusion could upwardly bias the effect on education. Moreover, the potential for such biases increases when considering the role of selective mortality.

Second, we perform the age-unrestricted analysis with the inclusion of individuals who have migrated at least once from their community of birth (full wave 4; Tables A8 and A9). The main analysis deliberately drops these observations because the inability to identify their community of birth would cause measurement errors in the IV construction. However, if internal migration remains quasi-random in respect to district and provincial immunization rates, this error will not produce any bias but would simply weaken the IV. Thus, in this robustness analysis, we assign a *leave-d-y-out* IV value that corresponds to the community of residence (birth) of the (non)migrant individuals, but without considering their immunization status in the *leave-d-y-out* IV computations.

Third, we re-estimate our models outlined in Eqs. 1–6 using the first wave of CHARLS (Tables A7 and A8) to check if the presence of observations that joined CHARLS at a later stage drives the results. Fourth, we control for a continuous version of the age variable rather than including year-of-birth dummies. All the aforementioned tests produce results in line with our main results in Tables 3 and 4. Fifth, we perform a falsification test by regressing the outcome variable proxies for education and cognitive abilities on the IV using only nontreated individuals. In fact, if the IV produces any effect on the outcome variables for this subpopulation of individuals who were not immunized in childhood, its exogeneity could not be confirmed. The results of this falsification test, reported in Table 6, not only always remain statistically insignificant, but the IV coefficient's magnitude is small. Interestingly, this test also indirectly confirms

[3] Tables A2—A5 provide full regression results.

Table 3
Effect of Childhood Vaccination on Education and Numeracy (OLS and 2SLS).

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	2.403***	1.673***	1.186***	0.868***	1.228***
	(18.60)	(14.71)	(10.53)	(9.00)	(11.34)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	11,255	10,580	10,580	10,580	10,110
Adj. R-sq	0.043	0.218	0.357	0.443	0.335

	Dependent Variable					
	Numeracy					
	(6)	(7)	(8)		(9)	(10)
	OLS					2SLS
Vaccine	0.150***	0.110***	0.088***		0.059***	0.086***
	(11.83)	(8.94)	(7.61)		(4.89)	(7.28)
Controls[a]	No	Yes	Yes	Yes	Yes	
Year-of-birth	No	No	Yes	Yes	Yes	
Province-of-birth	No	No	Yes	No	Yes	
Year*province	No	No	Yes	No	No	
Community-of-birth	No	No	No	Yes	No	
N	10,802	10,155	10,155	10,155	9707	
Adj. R-sq	0.018	0.097	0.161	0.172	0.132	

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. [a]The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority. The IV specifications drop a small number of observations because of multicollinearity between some cohorts and province of birth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

that the degree of misreporting of the treatment variable is negligible.

Sixth, following Oster (2019), we test the credibility of our treatment effects by exploiting selection on observables to gain insights into selection on unobservables. The intuition of this approach is that, if the coefficient of interest remains stable relative to R^2 when additional controls are added to the regression, the potential for severe bias is limited. Put simply, coefficient instability relative to R^2 is proportional to the degree of omitted variable bias. We thus estimate the threshold value of the proportional selection coefficient (δ) that would make the treatment effect zero. In other words, we assess how important the selection on unobservables (relative to the one on observables) needs to be to yield a null true treatment effect. We implemented this sensitivity analysis with the Stata *psacalc* package (Oster, 2019), which estimates δ by comparing the simple bivariate model ("short regression") to the one that includes the full list of controls ("long regression") given a maximum possible R^2 value (R_{max}). Although R^2 is naturally bounded in the [0;1] interval, assuming an R_{max} of one is overconservative (Oster, 2019). This is particularly the case for the outcome variables of this study, which are governed by several unmeasurable factors (e.g., innate ability) or by determinants quantified with consistent error (e.g., recalled childhood socioeconomics). Following the suggestion of Oster (2019), we thus set R_{max} equal to 1.3R (where R is the R^2 from the long regression). Note that according to the simulations in Oster (2019), a threshold of $R_{max} = 1.3R$ and $\delta \geq 1$ would confirm the treatment effect direction of 90 % of the randomized controlled trial results.

Table 7 reports the results of the coefficient stability analysis for each main outcome variable. Even if, as expected from observational data, our treatment effects vary considerably with the inclusion of control variables, the true treatment effect of vaccination would only be null when the selection on unobservables is more than two times that of observables for all the models. As already mentioned, because a value of $\delta \geq 1$ is set as the boundary (Oster, 2019), these results confirm the credibility of our estimates.

Finally, we run an omitted variable bias framework for sensitivity of IV estimates proposed by Cinelli and Hazlett (2022). This sensitivity test estimates the strength of an omitted confounder of the IV with the outcome variable (or of an indirect effect of the IV on the outcome) such that the IV study results turn statistically not significant (or null). We perform this analysis employing the *sensemakr*

Table 4
Effect of Childhood Vaccination on Mental Status and Episodic Memory (OLS and 2SLS).

	Dependent Variable				
	Mental Status				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.128***	0.094***	0.061***	0.071***	0.064***
	(11.79)	(9.24)	(6.03)	(7.23)	(6.51)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,303	9682	9682	9682	9252
Adj. R-sq	0.023	0.105	0.193	0.138	0.174

	Dependent Variable				
	Episodic Memory				
	(6)	(7)	(8)	(9)	(10)
	OLS				2SLS
Vaccine	0.085***	0.066***	0.041***	0.029***	0.039***
	(12.49)	(10.12)	(7.05)	(4.97)	(6.58)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,539	9910	9910	9910	9476
Adj. R-sq	0.027	0.087	0.205	0.237	0.180

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. [a]The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority. The IV specifications drop a small number of observations because of multicollinearity between some cohort and province of birth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Stata package (Cinelli, Ferwerda, & Hazlett, 2020) within the 2SLS specification that includes both year- and province-of-birth dummy variables (columns 5 of Tables 3 and 4). The results, reported in Table 8, show that a single omitted variable, orthogonal to the full set of covariates employed, that explains more than 4.5–8.6 % of the residual variance of both the treatment and the outcome would make the estimated treatment effect statistically not significant (*RV_q*; column 1).[4] To judge these results contextually, we additionally compute bias-corrected treatment effects assuming that a hypothetical unobserved confounder is one, two, or three times stronger than an observable benchmark control variable. Because of its strong relative explanatory power, we choose the "father illiterate" dummy variable as a benchmark for this exercise. In column 4, we present these corresponding bias-corrected coefficients and conclude that the true treatment effects of childhood vaccination on education and cognitive abilities would become statistically not significant (at 95 % confidence interval) only for a hypothetical confounder that is at least three times stronger than the benchmark.

4.2. Heterogenous effects

To gain insights into the role of childhood vaccination as a driver of lower socioeconomic inequality, we run a series of IV regressions employing stratified subsamples by the economic status of the family of origin (being poor in childhood) and mother and father literacy levels. The findings of the additional analysis, presented in Table A9, indicate that childhood vaccinations have a stronger effect on numeracy and episodic memory in children from lower-income households. We also find that childhood vaccinations have a stronger positive effect on episodic memory for individuals with less-educated parents, although for this subsample the effect of childhood vaccinations is slightly smaller on years of education, numeracy, and mental status.

Finally, we test whether the potential exposure to a more complete set of vaccines produces incremental positive effects on education and cognitive abilities. We do so by stratifying the sample in accordance with the timeline of vaccine licensing in China (see Fig. 2) while allowing the subsamples to be sufficiently large. Table A10 presents the results of this heterogeneity analysis, confirming

[4] Conversely, *RV_qa* in column 2 reports the strength of an omitted variable orthogonal to the covariates necessary to bring the point estimate to zero.

Table 5
Causal Mediation Analysis (linear CF).

	Vaccine IV		Vaccine & Education IV	
	(1)	(2)	(3)	(4)
	Numeracy			
	Estimate	SE	Estimate	SE
Controlled Direct	0.038***	0.008	0.032***	0.008
Natural Indirect	0.029***	0.003	0.013***	0.002
Total	0.066***	0.008	0.046***	0.008
Indirect/Total	43.39 %		28.26 %	
N	15,944		14,158	
	Mental Status			
	Estimate	SE	Estimate	SE
Controlled Direct	0.023***	0.005	0.023***	0.005
Natural Indirect	0.015***	0.002	0.006***	0.001
Total Effect	0.038***	0.005	0.029***	0.005
Indirect/Total	39.47 %		20.68 %	
N	15,944		14,775	
	Episodic Memory			
	Estimate	SE	Estimate	SE
Controlled Direct	0.045***	0.010	0.049***	0.010
Natural Indirect	0.032***	0.003	0.010***	0.001
Total	0.078***	0.011	0.059***	0.011
Indirect/Total	41.02 %		16.95 %	
N	15,944		14,775	

Notes: Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample excludes individuals older than 80 years and within-province migrants. The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority, and cohort- and provincial-level dummy variables. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. SE: standard error.

Table 6
Falsification Test for Vaccine IV (OLS, Control Sample).

	Dependent Variable			
	Education	*Numeracy*	*Mental Status*	*Episodic Memory*
	(1)	(2)	(3)	(4)
IV	−0.736	0.035	−0.034	−0.038
	(−1.54)	(0.60)	(−0.56)	(−1.34)
Controls[a]	Yes	Yes	Yes	Yes
Year & Province of birth	Yes	Yes	Yes	Yes
N	2296	2109	1917	2007
Adj. R-sq	0.319	0.175	0.201	0.193

Notes: IV specification t-statistics computed using a community-level -clustered correction (1000 replications). Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample excludes all individuals immunized before age 15 and older than 80 years and within-province migrants. [a]The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table 7
Coefficient Stability Analysis (Main Sample).

		Dependent Variable			
		Education	*Numeracy*	*Mental Status*	*Episodic Memory*
		(1)	(2)	(3)	(4)
Delta		2.690	2.866	2.369	2.183
R_{MAX}		0.440	0.190	0.240	0.250
Coeff.	Uncontrolled	2.363	0.149	0.127	0.084
	Controlled	1.219	0.083	0.061	0.039
R-sq.	Uncontrolled	0.042	0.019	0.023	0.027
	Controlled	0.339	0.144	0.184	0.190

Notes: Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample excludes individuals older than 80 years and within-province migrants. The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority, and dummies at the cohort (year-of-birth) and province levels.

Table 8
Sensitivity Analysis of Instrumental Variable Estimates (Main Sample).

Dependent Variable	*RV_q*	*RV_qa*		*Coeff.*	*S.E.*
	(1)	(2)	(3)	(4)	(5)
			Bound 1×	0.413	0.204
			Bound 2×	0.329	0.202
Years of Education	10.35 %	8.59 %	Bound 3×	0.245	0.199
			Bound 1×	0.048	0.023
			Bound 2×	0.043	0.023
Numeracy	7.65 %	5.84 %	Bound 3×	0.037	0.023
			Bound 1×	0.054	0.018
			Bound 2×	0.049	0.017
Mental Status	7.37 %	5.47 %	Bound 3×	0.045	0.017
			Bound 1×	0.026	0.010
Episodic Memory	6.39 %	4.49 %	Bound 2×	0.023	0.010
			Bound 3×	0.020	0.010

Notes: Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample excludes individuals older than 80 years and within-province migrants. The vector-of-control variables include father and mother illiterate, father and mother alive, poor in childhood, gender, hunger in childhood, ethnic minority, and dummies at the cohort (year-of-birth) and province levels. Benchmark control variable: father being illiterate. *RV_q* and *RV_qa* represent the robustness values for testing the null hypotheses that the treatment effect equals zero or it becomes statistically not significant, respectively.

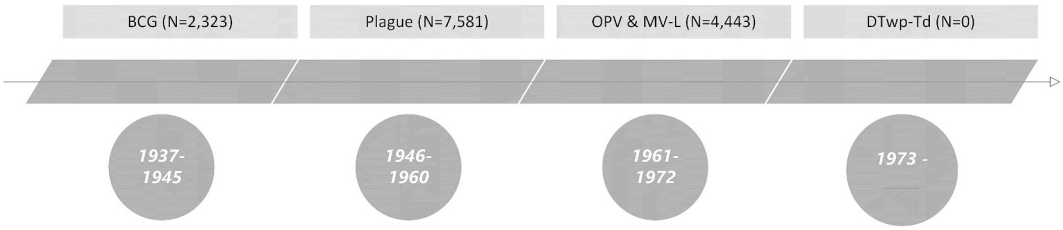

Fig. 2. Subsample Stratification According to the Timeline of Vaccine Licensing in China.
Notes: Adapted by the authors from Zheng et al. (2018). BCG: Bacillus Calmette–Guérin; DTwp-Td: Diphtheria and Tetanus toxoids and whole-cell pertussis; MV: Measles Vaccine; OPV: Oral Poliovirus Vaccine.

that cohorts potentially exposed to a more comprehensive set of vaccines enjoy higher levels of education and cognitive abilities. Note that this result is robust to the inclusion of both province and year-of-birth dummies.

5. Discussion and conclusions

By analyzing data from the China Health and Retirement Longitudinal Study, we find that adults over age 45 who were vaccinated before the age of 15 tend to perform better on standardized cognitive tests and complete more years of schooling than similar unvaccinated individuals. These effects are relatively strong, with vaccinated individuals enjoying about one additional year of schooling and performing substantially better on several cognitive tests, including 8.6, 6.4, and 3.9 percentage points higher numeracy, mental status, and episodic memory scores, respectively. These findings remain robust and statistically significant regardless of subsample and estimation strategy.

Our results thus reinforce the existing evidence of vaccination's health, cognition, and schooling effects in myriad countries, such as the additional 0.2–0.3 school grades by age 12 among Indian children age-appropriately vaccinated for measles (Nandi, Shet, et al., 2019). In particular, our analysis highlights that such benefits, which in China translate into approximately one extra year of schooling, are long-lasting and persist into adulthood. In addition, by comparing the effect size of our results with those in other studies that use CHARLS and assessing how other conditions or shocks measured at their extensive margin affect the episodic memory and mental status of older Chinese adults, we demonstrate that early vaccination's positive effects on cognitive outcomes are substantial.

In a study of widowhood's effect on cognitive functions in older Chinese, Zhang and colleagues conclude that continual widowhood status (i.e., across every data wave) decreases the episodic memory score by 0.15 points, while having any functional physical limitation lowers it by 0.17 points (Zhang, Li, Hongwei, & Liu, 2019). By constructing a similar index for comparative purposes (i.e., the actual number of recalled words instead of their share), this study shows that early vaccination increases the score by 0.39 points, thereby offsetting the joint negative effect of widowhood status and functional limitation by 120 %. Another cross-sectional study of China (Zhang et al. (2019), after adjusting for sociodemographic and cardiovascular diseases risk factors, identifies a negative association between untreated diabetes and episodic memory, with a decrease of 0.19 and 0.47 for the full and the 45–69-year-old samples, respectively. The magnitude of our results can also be considered significant when compared with findings on factors that positively

affect cognitive abilities; for example, a 0.18 points higher Serial Seven numeracy score induced by playing Mahjong or chess (Kesavayuth, Liang, & Zikos, 2018), compared with our finding of a 0.86 points improvement from childhood vaccination. By comparing the size of the vaccination effect on memory status in our study with the results reported by Pan and Chee (2020), we can infer that among older Chinese, childhood vaccination is twice as beneficial as engaging in one extra social activity (e.g., voluntary work, interaction with friends, educational courses, or sports). Lastly, our results on education relate to the findings of Nandi et al. (2020) on the effect of childhood exposure to the Indian Universal Immunization Programme on schooling of young adults. In their paper, these effects range from 0.18 to 0.29 more schooling grades, depending on the econometric specification. In our study, the same results are comparatively larger (approximately one additional year of schooling); this difference in magnitude can be justified only marginally by the treatment heterogeneity between the studies. In fact, the treatment group in Nandi et al. (2020) was exposed to a similar set of vaccines as the treatment group in this study (i.e., *measles*, BCG, *poliovirus*, and DPT), but likely with different timings and doses. Our findings align closely with those of Atwood and Pearlman (2023), who demonstrated that the widespread measles vaccination campaign in Mexico in 1973 led to an increase in schooling years of roughly 0.5 to 0.8 years among adult males.

Convergence in schooling might actually explain the relatively large effects found in our analysis of Chinese elderly. That is, as with income growth, countries with lower average education would experience faster schooling growth than relatively highly educated countries. This finding implies that the potential (in years of schooling) of a given investment in education is greater in less educated populations. Currently, China and India do not diverge much in terms of gross secondary schooling enrollment. However, the average years of schooling in the Indian study treatment sample equals 10.29 years, whereas the figure for the sample of treated Chinese in our study is only 5.25 years. This gap mainly relates to the different generations analyzed in the two studies. Finally, we perform a causal mediation analysis to quantify how much of the total effect of childhood immunization on later-life cognitive abilities is transmitted through higher education levels. Even though other (unmeasured) channels than education (e.g., better child health) explain much of the total effect, education has an economically meaningful and statistically significant mediating role in the vaccine–cognitive ability nexus.

Despite such significant effects, however, we consider our results conservative, not only because we do not account for herd immunity—the indirect protection from infectious disease provided by treated individuals to untreated controls—but also because we cannot directly control for the selective mortality that would dilute childhood vaccination's beneficial effects. Admittedly, however, a limitation of this study is its inability to assess which vaccines give rise to the beneficial effects observed in our analysis. In fact, as in other similar surveys (e.g., SHARE), we lack data on the exact vaccine administered during childhood. Nevertheless, based on the timeline of vaccine introductions in China, we can narrow down the possible vaccines to a selection (notably BCG and measles) that have mostly been shown to have positive nonspecific effects *(*Benn et al., 2013*)*.

Given the approximately 8 % increase in earnings induced by each extra year of schooling in China (Giles, Park, & Wang, 2019), the vaccination effect on cognitive abilities and educational outcomes imply economic benefits, particularly if—as our results suggest—childhood vaccines improve cognitive aging, thereby preventing or postponing cognitive decline into dementia (Petersen et al., 2009). Childhood vaccination may also help to reduce dementia's economic burden while reducing health and wealth inequities by lowering morbidity and mortality (Andre et al., 2008). In fact, in an analysis of our data segmented by the poverty status of parents, the impact of childhood vaccination on education, mental health, and episodic memory is marginally greater for children from poorer households. This result is also supported by He et al. (2022), who show that individuals with poorer economic and social backgrounds gain higher human capital returns from meningitis vaccination. Thus, our results indicate that the beneficial effects of vaccines on cognition and educational outcomes may even amplify inequality reduction and improve labor market outcomes (Viinikainen et al., 2020)—especially in the face of the 19 million children worldwide who received no routine immunizations in 2018 (WHO, 2019). Finally, improved human capital due to vaccines may also reduce risky behaviors, such as excessive alcohol consumption (Huang, Li, Liu, & Ruofei, 2023).

The recent COVID-19 pandemic has proven vaccine hesitancy to be a global problem, with a significant portion of the world population unwilling to get immunized. This hesitancy arises despite ample research underlying not only the safety of vaccines but also the beneficial long-term effects with regard to general health and even an individual's socioeconomic and cognitive status. We thus hope that, at a time when growing public vaccination hesitancy is lowering herd immunity in both high- and middle-low-income countries (*The Lancet Child & Adolescent Health.,* 2019), the robust evidence provided here (using vigorous quasi-experimental methods) of vaccination's long-term beneficial effects will mitigate reluctance and increase the demand for immunization by improving nonbinding vaccine recommendations (Lawler, 2020).

Funding acknowledgement

Our research was supported by the National Institute on Aging of the National Institutes of Health (Award Number P30AG024409) and also by the Value of Vaccination Research Network (VoVRN), which is funded by a grant from the Bill & Melinda Gates Foundation (Grant OPP1158136). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health, the VoVRN, or the Bill & Melinda Gates Foundation. We are grateful to Arindam Nandi for his comments and suggestions on an earlier version of this manuscript, and to Jeff Adams and Ravi Sadhu for their able and helpful assistance.

Data availability

Data can be downloaded upon registration from https://charls.pku.edu.cn/en/

Appendix A. Additional analyses and full results

Table A1
First-Stage Regressions, Education-IV (OLS).

	Dependent Variable		
	Years of Education		
	(1)	(2)	(3)
CR-IV	−0.037***	−0.039***	−0.038***
	(−3.45)	(−3.48)	(−3.51)
Mother illiterate	−0.631***	−0.655***	−0.642***
	(−6.91)	(−6.77)	(−6.85)
Father illiterate	−1.015***	−1.002***	−1.003***
	(−17.60)	(−16.57)	(−17.06)
Mother alive	0.337***	0.333***	0.331***
	(5.05)	(4.78)	(4.89)
Father alive	0.355***	0.304***	0.365***
	(3.89)	(3.20)	(3.92)
Poor in childhood	−0.365***	−0.373***	−0.378***
	(−7.24)	(−7.06)	(−7.33)
Gender (male = 1)	1.801***	1.790***	1.810***
	(33.04)	(31.22)	(32.39)
Hunger in childhood	−0.088	−0.077	−0.082
	(−1.30)	(−1.09)	(−1.19)
Ethnic minority	−0.191**	−0.114	−0.209**
	(−2.19)	(−1.22)	(−2.36)
Sample	Numeracy	Mental Status	Episodic Memory
N	10,244	9768	9990
F-stat (partial)	11.82	12.04	12.32

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. All models additionally control for dummies at the cohort (year of birth) and community levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A2
Effect of Childhood Vaccination on Education (Full Results, OLS, and 2SLS).

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	2.403***	1.673***	1.186***	0.868***	1.228***
	(18.60)	(14.71)	(10.53)	(9.00)	(11.34)
Mother illiterate		2.488***	2.626***	2.686***	2.646***
		(24.99)	(28.39)	(29.51)	(27.24)
Father illiterate		−1.338***	−0.614***	−0.072	−0.623***
		(−8.04)	(−4.35)	(−0.56)	(−4.24)
Mother alive		−1.605***	−1.207***	−0.803***	−1.260***
		(−16.37)	(−12.79)	(−10.10)	(−13.81)
Father alive		0.863***	0.144	0.097	0.183*
		(7.23)	(1.36)	(0.95)	(1.68)
Poor in childhood		0.557***	0.015	0.121	0.014
		(4.35)	(0.12)	(1.02)	(0.11)
Gender (male = 1)		−0.852***	−0.812***	−0.683***	−0.841***
		(−9.30)	(−10.09)	(−9.04)	(−10.34)
Hunger in childhood		−0.205*	−0.007	−0.004	−0.042
		(−1.67)	(−0.07)	(−0.04)	(−0.38)
Ethnic minority		−0.272	−0.979***	−0.091	−0.959***
		(−1.05)	(−4.23)	(−0.42)	1.228***
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	11,255	10,580	10,580	10,580	10,110
Adj. R-sq	0.043	0.218	0.357	0.443	0.335

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. All models additionally control for dummies at the cohort (year of birth) and community levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A3
Effect of Childhood Vaccination on Numeracy (Full Results, OLS, and 2SLS).

	Dependent Variable				
	Numeracy				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.150***	0.110***	0.088***	0.059***	0.086***
	(11.83)	(8.94)	(7.61)	(4.89)	(7.28)
Mother illiterate		0.182***	0.188***	0.190***	0.187***
		(19.69)	(19.82)	(20.09)	(20.05)
Father illiterate		−0.036***	−0.008	0.005	−0.008
		(−2.77)	(−0.68)	(0.38)	(−0.61)
Mother alive		−0.081***	−0.064***	−0.041***	−0.062***
		(−8.44)	(−7.08)	(−4.29)	(−6.63)
Father alive		0.049***	0.023**	0.020*	0.024**
		(4.37)	(2.00)	(1.70)	(2.00)
Poor in childhood		0.019	−0.004	0.010	−0.005
		(1.27)	(−0.28)	(0.67)	(−0.32)
Gender (male = 1)		−0.049***	−0.047***	−0.036***	−0.050***
		(−5.63)	(−5.36)	(−4.29)	(−5.83)
Hunger in childhood		0.004	0.014	0.011	0.020*
		(0.29)	(1.22)	(0.90)	(1.65)
Ethnic minority		−0.088***	−0.078***	−0.016	−0.080***
		(−3.59)	(−3.24)	(−0.62)	(−3.49)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,802	10,155	10,155	10,155	9707
Adj. R-sq	0.018	0.097	0.161	0.172	0.140

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. All models additionally control for dummies at the cohort (year of birth) and community levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A4
Effect of Childhood Vaccination on Mental Status (Full Results, OLS, and 2SLS).

	Dependent Variable				
	Numeracy				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.128***	0.094***	0.061***	0.071***	0.064***
	(11.79)	(9.24)	(6.03)	(7.23)	(6.51)
Mother illiterate		0.117***	0.130***	0.124***	0.128***
		(16.10)	(18.70)	(17.29)	(18.20)
Father illiterate		−0.052***	−0.011	−0.036***	−0.016*
		(−5.10)	(−1.25)	(−3.68)	(−1.68)
Mother alive		−0.078***	−0.055***	−0.068***	−0.059***
		(−11.22)	(−8.06)	(−10.00)	(−8.69)
Father alive		0.061***	0.014**	0.022***	0.018**
		(8.17)	(2.01)	(2.90)	(2.52)
Poor in childhood		0.019*	−0.005	−0.013	−0.011
		(1.89)	(−0.50)	(−1.25)	(−1.10)
Gender (male = 1)		−0.041***	−0.040***	−0.043***	−0.041***
		(−6.01)	(−5.98)	(−6.41)	(−6.34)
Hunger in childhood		−0.009	0.005	−0.008	0.005
		(−0.99)	(0.55)	(−0.85)	(0.55)
Ethnic minority		−0.043*	−0.067***	−0.048**	−0.068***
		(−1.95)	(−3.00)	(−2.20)	(−2.92)
Controls[a]	No	Yes	Yes	Yes	Yes

(*continued on next page*)

Table A4 (*continued*)

	Dependent Variable				
	Numeracy				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,303	9682	9682	9682	9251
Adj. R-sq	0.023	0.105	0.193	0.138	0.174

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. All models additionally control for dummies at the cohort (year of birth) and community levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A5
Effect of Childhood Vaccination on Episodic Memory (Full Results, OLS, and 2SLS).

	Dependent Variable				
	Episodic Memory				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.085***	0.066***	0.041***	0.029***	0.039***
	(12.49)	(10.12)	(7.05)	(4.97)	(6.58)
Mother illiterate		0.012***	0.017***	0.021***	0.019***
		(3.08)	(4.51)	(5.87)	(5.17)
Father illiterate		−0.045***	−0.020***	−0.014**	−0.021***
		(−6.79)	(−3.04)	(−2.13)	(−3.21)
Mother alive		−0.051***	−0.038***	−0.032***	−0.041***
		(−11.23)	(−8.78)	(−7.14)	(−9.14)
Father alive		0.045***	0.010*	0.003	0.009*
		(8.32)	(1.79)	(0.53)	(1.75)
Poor in childhood		0.025***	−0.002	0.003	−0.002
		(3.38)	(−0.27)	(0.41)	(−0.30)
Gender (male = 1)		−0.022***	−0.021***	−0.018***	−0.021***
		(−5.14)	(−5.44)	(−4.32)	(−4.98)
Hunger in childhood		0.010*	0.013**	0.017***	0.014**
		(1.86)	(2.38)	(3.36)	(2.57)
Ethnic minority		0.023*	0.018	−0.003	0.015
		(1.83)	(1.34)	(−0.30)	(1.11)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,539	9910	9910	9910	9476
Adj. R-sq	0.027	0.087	0.205	0.237	0.180

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes individuals older than 80 years and within-province migrants. All models additionally control for dummies at the cohort (year of birth) and community levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A6
Effect of Childhood Vaccination on Education and Numeracy (OLS and 2SLS; Age-Unrestricted Sample).

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	2.493***	1.717***	1.134***	0.814***	1.169***
	(20.28)	(15.90)	(10.40)	(8.77)	(1.76)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes

(*continued on next page*)

Table A6 (*continued*)

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	11,742	11,039	11,039	11,039	10,477
Adj. R-sq	0.048	0.229	0.372	0.453	0.348

	Dependent Variable				
	Numeracy				
	(6)	(7)	(8)	(9)	(10)
	OLS				2SLS
Vaccine	0.160***	0.118***	0.084***	0.058***	0.084***
	(13.13)	(10.22)	(7.45)	(5.10)	(7.50)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	11,116	10,461	10,461	10,461	9967
Adj. R-sq	0.022	0.104	0.179	0.186	0.154

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes within-province migrants. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A7
Effect of Childhood Vaccination on Mental Status and Episodic Memory (OLS and 2SLS; Age-Unrestricted Sample).

	Dependent Variable				
	Mental Status				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.137***	0.101***	0.060***	0.070***	0.065***
	(13.02)	(10.22)	(6.00)	(7.33)	(6.74)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,622	9981	9981	9981	9504
Adj. R-sq	0.027	0.112	0.216	0.160	0.193

	Dependent Variable				
	Episodic Memory				
	(6)	(7)	(8)	(9)	(10)
	OLS				2SLS
Vaccine	0.091***	0.070***	0.040***	0.028***	0.036***
	(13.14)	(10.70)	(6.91)	(4.80)	(6.39)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	10,864	10,218	10,218	10,218	9730
Adj. R-sq	0.031	0.096	0.232	0.264	0.207

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. The sample excludes within-province migrants. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A8
Effect of Childhood Vaccination on Education and Numeracy (OLS and IV; Full Sample, Wave 4).

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				IV
Vaccine	2.308***	1.617***	1.187***	0.842***	1.135***
	(16.19)	(13.26)	(11.49)	(9.46)	(10.60)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	17,697	16,344	16,344	16,344	15,983
Adj. R-sq	0.040	0.200	0.336	0.416	0.290

	Dependent Variable				
	Numeracy				
	(6)	(7)	(8)	(9)	(10)
	OLS				IV
Vaccine	0.173***	0.131***	0.096***	0.074***	0.097***
	(13.42)	(11.96)	(9.30)	(6.86)	(9.35)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	16,242	15,071	15,071	15,071	14,348
Adj. R-sq	0.024	0.111	0.179	0.188	0.152

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A9
Effect of Childhood Vaccination on Mental Status and Episodic Memory (OLS and 2SLS; Full Sample, Wave 4).

	Dependent Variable				
	Mental Status				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.136***	0.101***	0.069***	0.076***	0.069***
	(13.64)	(11.55)	(8.09)	(9.22)	(8.47)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	15,498	14,377	14,377	14,377	13,689
Adj. R-sq	0.026	0.111	0.218	0.146	0.184

	Dependent Variable				
	Episodic Memory				
	(6)	(7)	(8)	(9)	(10)
	OLS				2SLS
Vaccine	0.098***	0.079***	0.047***	0.037***	0.044***
	(11.49)	(11.71)	(8.94)	(6.16)	(8.18)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No

(*continued on next page*)

Table A9 (*continued*)

	Dependent Variable					
	Episodic Memory					
	(6)	(7)	(8)		(9)	(10)
	OLS					2SLS
N	15,821	14,682	14,682	14,682	13,982	
Adj. R-sq	0.033	0.103	0.232	0.285	0.205	

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A10
Effect of Childhood Vaccination on Education and Numeracy (OLS and 2SLS; Full Sample, Wave 1).

	Dependent Variable				
	Years of Education				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	2.471***	1.747***	1.239***	0.849***	1.191***
	(17.46)	(14.17)	(11.05)	(8.84)	(10.85)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	12,156	11,496	11,496	11,496	10,915
Adj. R-sq	0.045	0.228	0.368	0.466	0.332

	Dependent Variable					
	Numeracy					
	(6)	(7)	(8)		(9)	(10)
	OLS					2SLS
Vaccine	0.140***	0.108***	0.073***		0.054***	0.075***
	(10.57)	(8.24)	(5.74)		(4.14)	(5.93)
Controls[a]	No	Yes	Yes	Yes	Yes	
Year-of-birth	No	No	Yes	Yes	Yes	
Province-of-birth	No	No	Yes	No	Yes	
Year*province	No	No	Yes	No	No	
Community-of-birth	No	No	No	Yes	No	
N	11,293	10,687	10,687	10,687	10,168	
Adj. R-sq	0.017	0.100	0.185	0.209	0.148	

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 1 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A11
Effect of Childhood Vaccination on Mental Status and Episodic Memory (OLS and 2SLS; Full Sample, Wave 1).

	Dependent Variable				
	Mental Status				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS
Vaccine	0.133***	0.103***	0.071***	0.078***	0.066***
	(11.17)	(9.08)	(7.23)	(7.09)	(6.56)
Controls[a]	No	Yes	Yes	Yes	Yes
Year-of-birth	No	No	Yes	Yes	Yes
Province-of-birth	No	No	Yes	No	Yes
Year*province	No	No	Yes	No	No
Community-of-birth	No	No	No	Yes	No
N	11,018	10,418	10,418	10,418	9897
Adj. R-sq	0.029	0.115	0.238	0.144	0.196

(*continued on next page*)

Table A11 (*continued*)

	Dependent Variable				
	Mental Status				
	(1)	(2)	(3)	(4)	(5)
	OLS				2SLS

	Dependent Variable					
	Episodic Memory					
	(6)	(7)	(8)		(9)	(10)
	OLS					2SLS
Vaccine	0.060***	0.046***	0.025***		0.019***	0.026***
	(10.01)	(7.45)	(4.28)		(3.69)	(4.46)
Controls[a]	No	Yes	Yes	Yes	Yes	
Year-of-birth	No	No	Yes	Yes	Yes	
Province-of-birth	No	No	Yes	No	Yes	
Year*province	No	No	Yes	No	No	
Community-of-birth	No	No	No	Yes	No	
N	10,276	9718	9718	9718	9223	
Adj. R-sq	0.016	0.059	0.163	0.253	0.114	

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 1 (regular) data and corrected for nonresponse rate using cross-sectional probability sampling weights. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A12
Effect of Childhood Vaccination on Education and Cognitive Abilities for Subsamples of Individual Poor in Childhood and with Mother or Father Illiterate (2SLS).

	Poor in childhood subsample			
	Years of Education	Numeracy	Mental Status	Episodic Memory
	(1)	(2)	(3)	(4)
Vaccine	1.356***	0.065***	0.043***	0.082***
	(8.01)	(4.27)	(4.90)	(4.64)
N	4100	3732	3848	3941
adj. R-sq	0.320	0.165	0.157	0.156

	Mother illiterate subsample			
	(5)	(6)	(7)	(8)
Vaccine	1.197***	0.067***	0.039***	0.085***
	(11.07)	(6.20)	(6.60)	(6.94)
N	8826	8067	8275	8479
adj. R-sq	0.316	0.165	0.157	0.140

	Father illiterate subsample			
	(9)	(10)	(11)	(12)
Vaccine	1.091***	0.066***	0.038***	0.075***
	(9.45)	(5.44)	(5.85)	(5.69)
N	6390	5764	5941	6110
adj. R-sq	0.304	0.162	0.159	0.141

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample includes within-province migrants and individuals older than 80 years. All regressions control for vector-of-control variables including father and mother illiterate, father and mother alive, poor in childhood, age, gender, hunger in childhood, ethnic minority, and dummies at provincial and year-of-birth levels. Specifications 1–4, 5–8, and 9–12 do not control for poor in childhood, mother illiterate, and father illiterate, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

Table A13
Effect of Childhood Vaccination on Education and Cognitive Abilities by Age Groups (2SLS).

	Years of Education			Numeracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Vaccine	0.889***	1.556***	0.870***	0.051**	0.107***	0.141***

(*continued on next page*)

Table A13 (*continued*)

	Years of Education			Numeracy		
	(1)	(2)	(3)	(4)	(5)	(6)
	(3.99)	(10.43)	(2.96)	(2.30)	(7.20)	(5.07)
Sample	1937–1945	1946–1960	1961–1970	1937–1945	1946–1960	1961–1970
N	2306	7530	4428	2174	7281	4253
adj. R-sq	0.262	0.272	0.221	0.149	0.106	0.124

	Mental Status			Episodic Memory		
	(7)	(8)	(9)	(10)	(11)	(12)
Vaccine	0.057***	0.082***	0.048**	0.025**	0.052***	0.051***
	(3.18)	(6.95)	(2.22)	(2.18)	(7.04)	(4.86)
Sample	1937–1945	1946–1960	1961–1970	1937–1945	1946–1960	1961–1970
N	2025	6957	4149	2091	7121	4196
adj. R-sq	0.174	0.139	0.100	0.100	0.093	0.091

Notes: T-statistics clustered at the community level in parentheses. Estimates obtained using CHARLS waves 3 (life history) and 4 (regular) data. The sample includes within-province migrants and individuals older than 80 years. All regressions control for vector-of-control variables including father and mother illiterate, father and mother alive, poor in childhood, age, gender, hunger in childhood, ethnic minority, and dummies at community and year-of-birth levels. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Adj. R-sq: adjusted R^2.

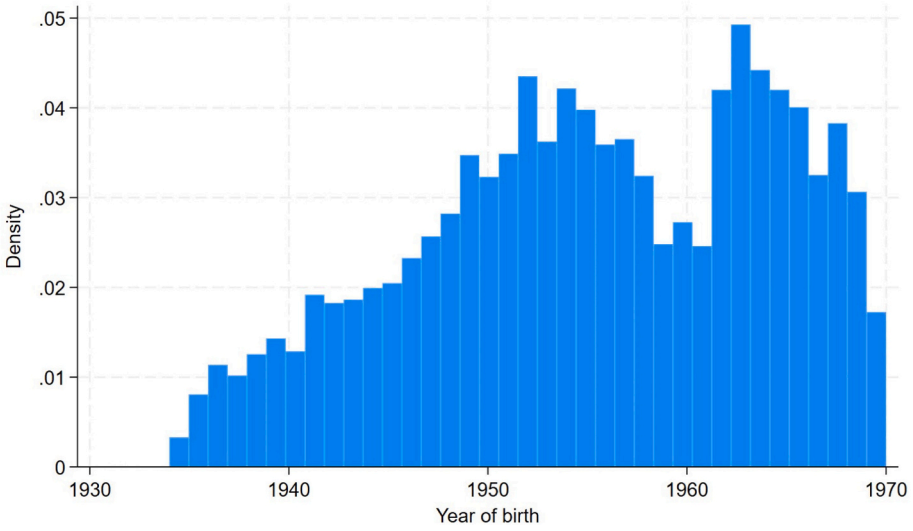

Fig. A1. Distribution of Birth Years (CHARLS Wave 4).
Notes: Drawn by author using CHARLS wave 4 data.

Appendix B. Selected survey tools

Table B1
Outcome Variables Survey Questions.

Dependent Variables		
Variable	Survey Question	Description
Years of schooling	What is the highest level of education you have completed?	Transformation based on Chinese education system
Numeracy (Serial 7)	Let's try some subtraction of numbers this time. What does 100 minus 7 equal? …and 7 from that? (5 times)	Share of correct answers
Mental Status		Share of correct answers
Name year	Please tell me today's date.	1 if year is correct
Name month	Please tell me today's date.	1 if month is correct

(*continued on next page*)

Table B1 (*continued*)

Dependent Variables		
Variable	Survey Question	Description
Name weekday	Please tell me today's date.	1 if weekday is correct
Name day	Please tell me today's date.	1 if day is correct
Episodic memory		Share of recalled words
Immediate work recall	We are going to read a list consisting of 10 words, and we would like you to memorize as many as you can.	
Delayed word recall	Please tell me any of the words that you remember now (4 min later).	

Table B2
Control Variables Survey Questions.

Control Variables		
Variable	Survey Question	Description
Mother alive	Is your biological mother alive?	1 if mother alive
Father alive	Is your biological father alive?	1 if father alive
Mother illiterate	What is the highest level of education your biological mother completed?	1 if no formal education (vs. any other)
Father illiterate	What is the highest level of education your biological father completed?	1 if no formal education (vs. any other)
Ethnic minority	What ethnicity is [the name of the respondent]?	Han (vs. any other)
Poor in childhood	When you were a child before age 17, compared with the average family in the same community/village at that time, how was your family's financial situation?	1 if worse or a lot worse off than others
Hunger in childhood	At what age ranges did this (your family did not have enough food to eat) happen?	1 if hunger before age 15
Age	Imputed age from exact date of birth	Age in years

References

Abeliansky, A. L., & Strulik, H. (2018). Hungry children age faster. *Economics and Human Biology, 29*, 211–220. https://doi.org/10.1016/j.ehb.2018.03.005

Andersen, A., Fisker, A. B., Rodrigues, A., Martins, C., Ravn, H., Lund, N., … Aaby, P. (2018). National immunization campaigns with oral polio vaccine reduce all-cause mortality: A natural experiment within seven randomized trials. *Frontiers in Public Health, 6*, 13. https://doi.org/10.3389/fpubh.2018.00013

Andre, F. E., Booy, R., Bock, H. L., Clemens, J., Datta, S. K., John, T. J., … Schmitt, H. J. (2008). Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bulletin of the World Health Organization, 86*(2), 140–146. https://doi.org/10.2471/blt.07.040089

Anekwe, T. D., Newell, M. L., Tanser, F., Pillay, D., & Barnighausen, T. (2015). The causal effect of childhood measles vaccination on educational attainment: A mother fixed-effects study in rural South Africa. *Vaccine, 33*(38), 5020–5026. https://doi.org/10.1016/j.vaccine.2015.04.072

Atwood, A. (2022). The long-term effects of measles caccination on earnings and employment. *American Economic Journal: Economic Policy, 14*(2), 34–60. https://doi.org/10.1257/pol.20190509

Atwood, A., & Pearlman, S. (2023). *The long term benefits of vaccination campaigns: Evidence from measles*. Pennsylvania. Mimeo: Haverford College.

Barnighausen, T., Bloom, D. E., Cafiero-Fonseca, E. T., & O'Brien, J. C. (2014). Valuing vaccination. *Proceedings of the National Academy of Sciences, 111*(34), 12313–12319. https://doi.org/10.1073/pnas.1400475111

Barteska, P., Dobkowitz, S., Olkkola, M., & Rieser, M. (2023). Mass vaccination and educational attainment: Evidence from the 1967–68 measles eradication campaign. *Journal of Health Economics, 92*, Article 102828.

Benn, C. S., Netea, M. G., Selin, L. K., & Aaby, P. (2013). A small jab - a big effect: Nonspecific immunomodulation by vaccines. *Trends in Immunology, 34*(9), 431–439. https://doi.org/10.1016/j.it.2013.04.004

van den Berg, G. J., von Hinke, S., & Vitt, N. (2023). *Early life exposure to measles and later-life outcomes: Evidence from the introduction of a vaccine*. UK: School of Economics, University of Bristol.

Bloom, D. E., Canning, D., & Shenoy, E. S. (2011). The effect of vaccination on children's physical and cognitive development in the Philippines. *Applied Economics, 44*(21), 2777–2783. https://doi.org/10.1080/00036846.2011.566203

Bloom, D. E., Canning, D., & Weston, M. (2017). The value of vaccination. In *Fighting the diseases of poverty* (pp. 214–238). New York: Routledge.

Bütikofer, A., & Salvanes, K. G. (2020). Disease control and inequality reduction: Evidence from a tuberculosis testing and vaccination campaign. *The Review of Economic Studies, 87*(5), 2087–2125.

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics, 24*(2), 365–389. https://doi.org/10.1016/j.jhealeco.2004.09.008

Chen, Y., Guo, Y., Huang, J., & Song, Y. (2019). Intergenerational transmission of education in China: New evidence from the Chinese Cultural Revolution. *Review of Development Economics, 23*(1), 501–527. https://doi.org/10.1111/rode.12558

Chuard, C., Schwandt, H., Becker, A. D., & Haraguchi, M. (2022). *Economic vs. epidemiological approaches to measuring the human capital impacts of infectious disease elimination*. Cambridge, Massachusetts: National Bureau of Economic Research.

Cinelli, C., Ferwerda, J., & Hazlett, C. (2020). *Sensemakr: Sensitivity analysis tools for OLS in R and Stata*. Available at SSRN 3588978.

Cinelli, C., & Hazlett, C. (2022). *An omitted variable bias framework for sensitivity analysis of instrumental variables*. Available at SSRN 4217915.

Daramola, R., Hossain, M. S., Kazianga, H., & Fogam, A. K. N. (2022). *The lasting effects of early Childhood interventions: The Sankaraàs National Vaccination Commando Program in Burkina Faso. edited by USA. Presented at the 2023 ASSA meeting. New Orleans. Mimeo*.

Driessen, J., Razzaque, A., Walker, D., & Canning, D. (2015). The effect of childhood measles vaccination on school enrolment in Matlab, Bangladesh. *Applied Economics, 47*(55), 6019–6040. https://doi.org/10.1080/00036846.2015.1061647

Emsley, R., & Liu, H. (2013). PARAMED: Stata module to perform causal mediation analysis using parametric regression models. In *Statistical software components S457581*. Boston College Department of Economics.

Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences, 106*(16), 6545–6549. https://doi.org/10.1073/pnas.0811910106

Fan, W. (2016). Turning point or selection? The effect of rustication on subsequent health for the Chinese cultural revolution cohort. *Social Science & Medicine, 157*, 68–77. https://doi.org/10.1016/j.socscimed.2016.03.044
Fine, P. E., Williams, T. N., Aaby, P., Kallander, K., Moulton, L. H., Flanagan, K. L., … Vaccines Working Group on Non-specific Effects of. (2009). Epidemiological studies of the "non-specific effects" of vaccines: Data collection in observational studies. *Tropical Medicine and Internation Health, 14*(9), 969–976. https://doi.org/10.1111/j.1365-3156.2009.02301.x
Frölich, M., & Huber, M. (2017). Direct and indirect treatment effects–causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79*(5), 1645–1666.
Giles, J., Park, A., & Wang, M. (2019). The great proletarian cultural revolution, disruptions to education, and the returns to schooling in urban China. *Economic Development and Cultural Change, 68*(1), 131–164. https://doi.org/10.1086/700568
Glymour, M. M., Kawachi, I., Jencks, C. S., & Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology and Community Health, 62*(6), 532–537. https://doi.org/10.1136/jech.2006.059469
Hassevoort, K. M., Khan, N. A., Hillman, C. H., & Cohen, N. J. (2016). Childhood markers of health behavior relate to hippocampal health, memory, and academic performance. *Mind, Brain, and Education, 10*(3), 162–170. https://doi.org/10.1111/mbe.12108
Havari, E., & Mazzonna, F. (2015). Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE. *European Journal of Population, 31*(3), 233–257. https://doi.org/10.1007/s10680-014-9332-y
He, F., Huang, W., & Wang, J. (2022). Vaccination, human capital investment, and economic growth. In *20th Quarterly Meeting of the China Labor Economists Forum, Guangdong, China, March 26*.
Huang, C., Li, C., Liu, F., & Ruofei, X. (2023). Vaccination and risky behaviors: Evidence from the hepatitis B vaccination campaign in China. *Journal of Population Economics, 36*, 2549–2580.
Kesavayuth, D., Liang, Y., & Zikos, V. (2018). An active lifestyle and cognitive function: Evidence from China. *The Journal of the Economics of Ageing, 12*, 183–191. https://doi.org/10.1016/j.jeoa.2018.05.001
Lawler, E. C. (2020). Giving teens a boost? *American Journal of Health Economics, 6*(2), 251–287. https://doi.org/10.1086/707834
Lei, X., Smith, J. P., Sun, X., & Zhao, Y. (2014). Gender differences in cognition in China and reasons for change over time: Evidence from CHARLS. *J Econ Ageing, 4*, 46–55. https://doi.org/10.1016/j.jeoa.2013.11.001
Liang, X., & Liu, Y. (2019). My experience as director of the National Immunization Program (historical review). In X. Liang (Ed.), *Immunization program in China* (pp. 1–14). Singapore: Springer Singapore.
McArdle, J. J., Smith, J. P., & Willis, R. (2011). *Cognition and economic outcomes in the health and retirement survey*. Massachusetts: National Bureau of Economic Research, Explorations in the Economics of Aging Cambridge.
Meng, X., & Gregory, R. G. (2002). The impact of interrupted education on subsequent educational attainment: A cost of the Chinese cultural revolution. *Economic Development and Cultural Change, 50*(4), 935–959.
Molina, T. (2016). Reporting heterogeneity and health disparities across gender and education levels: Evidence from four countries. *Demography, 53*(2), 295–323. https://doi.org/10.1007/s13524-016-0456-z
Nandi, A., Deolalikar, A. B., Bloom, D. E., & Laxminarayan, R. (2019). Haemophilus influenzae type b vaccination and anthropometric, cognitive, and schooling outcomes among Indian children. *Annals of the New York Academy of Sciences, 1449*(1), 70–82. https://doi.org/10.1111/nyas.14127
Nandi, A., Shet, A., Behrman, J. R., Black, M. M., Bloom, D. E., & Laxminarayan, R. (2019). Anthropometric, cognitive, and schooling benefits of measles vaccination: Longitudinal cohort analysis in Ethiopia, India, and Vietnam. *Vaccine, 37*(31), 4336–4343. https://doi.org/10.1016/j.vaccine.2019.06.025
Nandi, A., Kumar, S., Shet, A., Bloom, D. E., & Laxminarayan, R. (2020). Childhood vaccinations and adult schooling attainment: Long-term evidence from India's universal immunisation Programme. *Social Science & Medicine, 250*, Article 112885. https://doi.org/10.1016/j.socscimed.2020.112885
Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics, 37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711
Ozawa, S., Mirelman, A., Stack, M. L., Walker, D. G., & Levine, O. S. (2012). Cost-effectiveness and economic benefits of vaccines in low- and middle-income countries: A systematic review. *Vaccine, 31*(1), 96–108. https://doi.org/10.1016/j.vaccine.2012.10.103
Pan, X., & Chee, K. H. (2020). The power of weak ties in preserving cognitive function: A longitudinal study of older Chinese adults. *Aging and Mental Health, 24*(7), 1046–1053. https://doi.org/10.1080/13607863.2019.1597015
Pepper, S. (2000). *Radicalism and education reform in 20th-century China: The search for an ideal development model*. Cambridge University Press.
Petersen, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., … Jack, C. R., Jr. (2009). Mild cognitive impairment: Ten years later. *Archives of Neurology, 66*(12), 1447–1455. https://doi.org/10.1001/archneurol.2009.266
Roth, A., Garly, M. L., Jensen, H., Nielsen, J., & Aaby, P. (2006). Bacillus Calmette-Guerin vaccination and infant mortality. *Expert Review of Vaccines, 5*(2), 277–293.
Saadatian-Elahi, M., Aaby, P., Shann, F., Netea, M. G., Levy, O., Louis, J., … Warren, W. (2016). Heterologous vaccine effects. *Vaccine, 34*(34), 3923–3930. https://doi.org/10.1016/j.vaccine.2016.06.020
Sidel, V. W., & Sidel, R. (1975). The health care delivery system of the People's Republic of China. *World Development, 3*(7–8), 539–549.
Smith, J. P. (2009). Reconstructing childhood health histories. *Demography, 46*(2), 387–403. https://doi.org/10.1353/dem.0.0058
Smith, J. P., Strauss, J., & Zhao, Y. (2014). Healthy aging in China. *Journal of the Economics of Ageing, 4*, 37–43. https://doi.org/10.1016/j.jeoa.2014.08.006
Sundquist, J. (2021). *How to use spatial instruments*. Mimeo.
The Lancet Child & Adolescent Health (Ed.). (2019). Vaccine hesitancy. *The Lancet Child & Adolescent Health., 3*(5), 281.
Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods, 18*(2), 137.
Viinikainen, J., Bryson, A., Böckerman, P., Elovainio, M., Hutri-Kähönen, N., Juonala, M., … Pulkki-Råback, L. (2020). Do childhood infections affect labour market outcomes in adulthood and, if so, how? *Economics & Human Biology, 37*, Article 100857. https://doi.org/10.1016/j.ehb.2020.100857
WHO. (2005). IABS Scientific workshop on neurovirulence tests for live virus vaccines: Final report. https://cdn.who.int/media/docs/default-source/biologicals/vaccine-quality/iabs-scientific-workshop-on-neurovirulence-tests-for-live-virus-vaccines5dc26d99-a7a1-4dfc-b9f5-ce9d4e686467.pdf?sfvrsn=f1e0c7f2_1&download=true.
WHO. (2019). Immunization coverage accessed 31.03.2020. https://www.who.int/news-room/fact-sheets/detail/immunization-coverage
World Bank. (1984). *The health sector in China*. Washington, DC: Population, Health and Nutrition Department, World Bank.
Yu, W., Lee, L. A., Liu, Y., Scherpbier, R. W., Wen, N., Zhang, G., … Wang, H. (2018). Vaccine-preventable disease control in the People's Republic of China: 1949-2016. *Vaccine, 36*(52), 8131–8137. https://doi.org/10.1016/j.vaccine.2018.10.005
Zhang, L., Yang, J., Liao, Z., Zhao, X., Hu, X., Zhu, W., & Zhang, Z. (2019). Association between diabetes and cognitive function among people over 45 years old in China: A cross-sectional study. *International Journal of Environmental Resources and Public Health, 16*(7), 1294. https://doi.org/10.3390/ijerph16071294
Zhang, Z., Li, L. W., Hongwei, X., & Liu, J. (2019). Does widowhood affect cognitive function among Chinese older adults? *SSM Population Health, 7*, Article 100329.
Zheng, Y., Rodewald, L., Yang, J., Qin, Y., Pang, M., Feng, L., & Yu, H. (2018). The landscape of vaccines in China: History, classification, supply, and price. *BMC Infectious Diseases, 18*(1), 502. https://doi.org/10.1186/s12879-018-3422-0