The Very Long-Run Effect of Large-Scale Deworming in China

Chao Liu (Northwestern University)
Gordon G. Liu (Peking University)

January 15, 2023

Motivation

- Recent studies about early-life health and lifetime outcomes
  - Use of randomized controlled trials
  - Availability of comprehensive administrative data
- Little on the long-term effect of large-scale public health interventions in developing areas
  - Long-run vs. short-run
  - Large-scale vs. RCT
- Debate on the effectiveness of mass deworming programs
  - Positive: Miguel and Kremer, 2004; Baird et al., 2016; Ozier, 2018
  - Limited: Taylor-Robinson et al., 2015; Croker and Atun, 2019

This Paper

- **Shock:** deworming campaign in China in the late 1950s
  - A large-scale program targeted at schistosomiasis
  - Achieved great success sufficiently long ago
  - Mobilized in a top-down manner
- **Empirical Approach:** cohort difference-in-differences
- **Data:**
  - Hand-collected disease data
  - census data in 1990
  - survey data in 2010
- **Outcomes:** multiple adult outcomes at different stages of the life course
  - Education: treated cohorts and their offspring
  - Employment: after 30 and 50 years
  - Economic Status: after 50 years

Preview of the Results

Long-run effect: evidence from census

- Significantly positive effect on educational attainment
- Rural females benefited more
- Little effect on employment status in one's thirties

Longer-run effect: evidence from China Family Panel Studies 2010 (CFPS)

- Better economic status for rural males after over 50 years
- Labor market success in one's fifties
- Individuals from a low socioeconomic background benefited more
- Significant positive effect on the second generation's schooling

Related Literature

- Effect of deworming programs
  - Experiment (Miguel and Kremer, 2004; Baird et al., 2016; Ozier, 2018)
  - Natural experiment (Bleakley, 2007)
  - **Early-life intervention**
  - **Over a longer term, at different periods, and intergenerational effects**
- Early-life health and adult outcomes
  - Negative shocks in developing areas (Chen and Zhou, 2007; Almond et al., 2007; Barreca 2010, Bleakley 2010, Cutler et al. 2010, Lucas 2010)
  - Positive shocks in developed areas (Bütikofer and Salvanes, 2020; Grönqvist et al., 2020; Bhalotra et al., 2022)
  - **A positive shock in an extremely underdeveloped economy**

Outline

- **Data**
- **Background**
- **Results**
  - Long-run effect
  - Longer-run effect
- **Conclusion**

Data

Data

1. China Schistosomiasis Atlas (Qian 1988)
   - pre-control prevalence: the probability of an infection in a population
   - 255 endemic counties among 858 counties in 12 endemic provinces
2. 1 percent sample of the 1990 China Population Census ▸ Summary Statistics
   - rural *hukou*, born between 1946 and 1966, in endemic provinces
   - migration between counties was scant in rural areas
   - rare cases of people changing their *hukou* type.
3. China Family Panel Studies 2010 (CFPS) ▸ Summary Statistics
   - drop individuals who moved out of their birthplaces
   - rural *hukou* when the respondent was age 3 years
   - 23 endemic counties of total 112 counties

Background

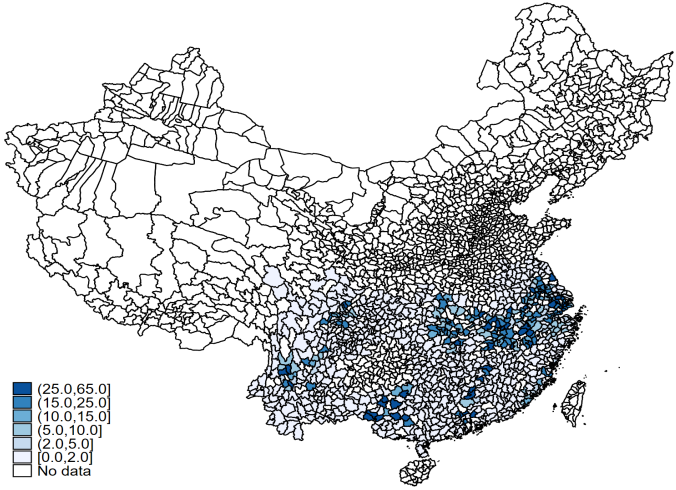
Schistosomiasis

- An acute and chronic parasitic disease
- Infection: larval forms of the parasite penetrate the skin through infested water
- Typical symptoms: abdominal pain, diarrhea, bloody stool, and blood in the urine
- Child patients: poor growth and learning difficulties
- Long-time adult patients: kidney failure, liver damage, and bladder cancer (WHO 2017)

The Spread in China

- An ancient medicine book (circa 400 BC) recorded clinical symptoms
- China was the largest endemic area of S. japonicum infection in the early 20th century
- Spread across 12 provinces in the south of China or along the Yangtze River
- The estimated number of patients was 10 million in the early 1950s
- More than 100 million people were at risk of infection (Mao and Shao 1982).

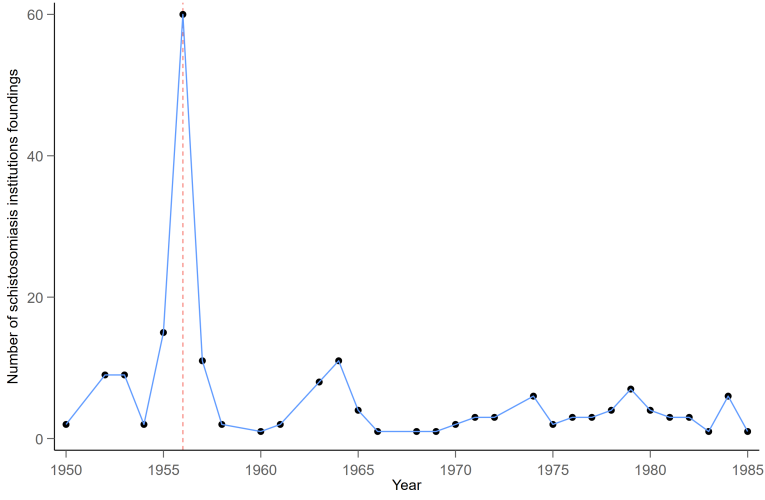
Pre-Control Prevalence



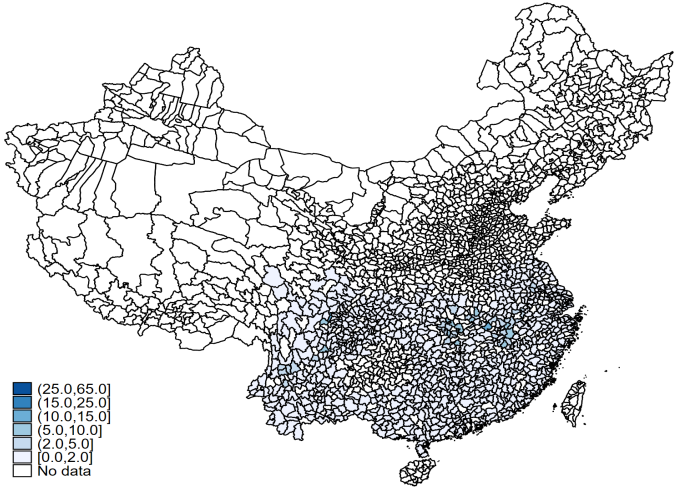
Schistosomiasis Control in China

- Set up a leading group to take charge of nationwide schistosomiasis control in 1955
- Issued an instruction on the eradication of schistosomiasis in 1956
- A disease control campaign was carried out in 12 endemic provinces
- Key tools: snail control through environmental modification and molluscicide
- Special anti-schistosomiasis health stations at three levels
- A great decline in schistosomiasis infections immediately (Yuan 1989)

Medical Institutions Specialized in Schistosomiasis



Post-Control Prevalence



Results From Census: Long-Run Effect

Empirical Strategy

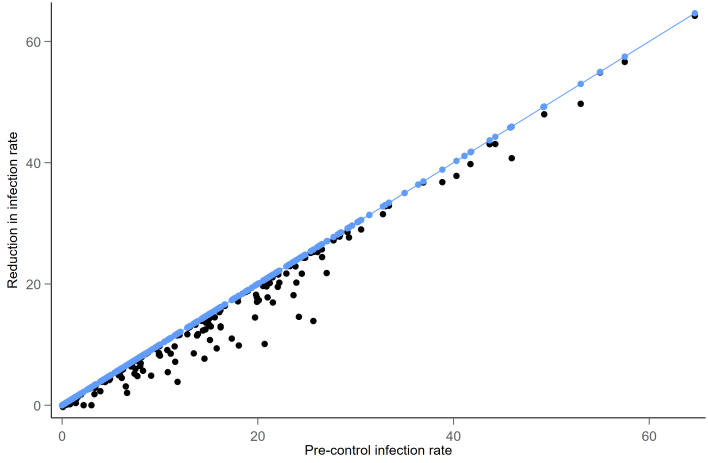
$$y_{ijc} = \alpha + \beta\,(sch_j \times post_{ic}) + X_{ijc}\Theta + \delta_j + \delta_c + \varepsilon_{ijc}$$

- sch_j: the pre-control schistosomiasis prevalence of county j
- $post_{ic}$: equals 1 if the individual i of cohort c was born after 1957
- δ_j, δ_c: county fixed effects and cohort fixed effects

This research design necessarily assumes that

- there were regional variations in schistosomiasis exposure
- areas where schistosomiasis was more endemic benefited more

Pre-Control Infection Rate and Reduction in Infection Rate



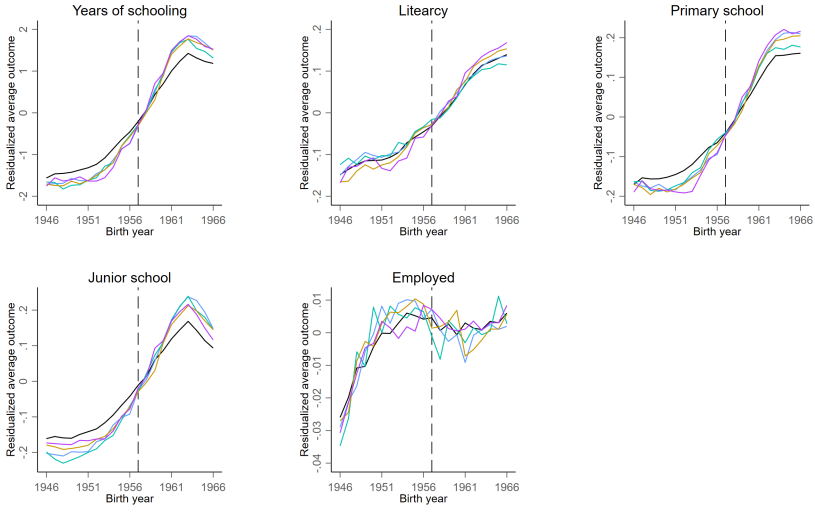
If a county eradicated schistosomiasis, it is located on the diagonal line

Baseline Results

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Female		
sch × post	1.855	0.099	0.228	0.202	-0.002
	(0.348)	(0.045)	(0.044)	(0.056)	(0.012)
Observations	651,812	651,812	651,812	651,812	651,812
			Panel B. Male		
sch × post	1.137	-0.032	0.120	0.170	-0.002
	(0.246)	(0.024)	(0.038)	(0.036)	(0.003)
Observations	676,545	676,545	676,545	676,545	676,545
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes

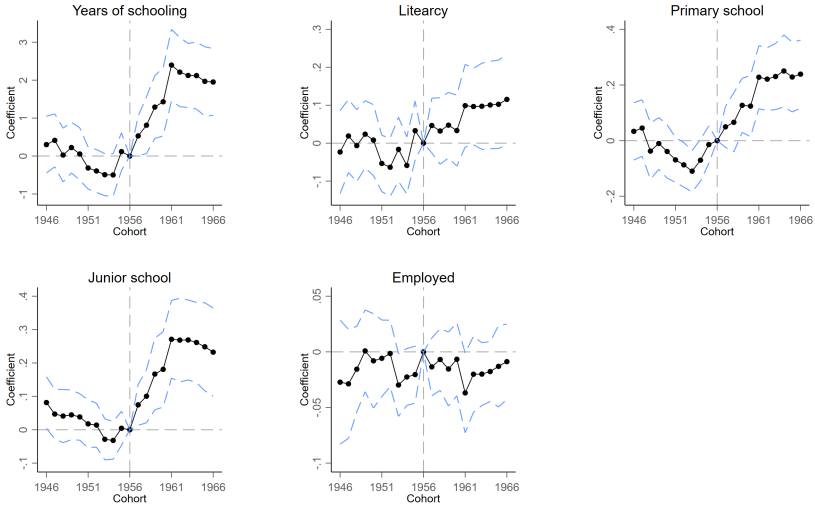
10 pp ↓ schistosomiasis infection rate → 0.19 years ↑ expected schooling for females

Parallel Trends Assumption (Females)



▸ Males

Cohort Analysis (Females)



▸ Males

Threats: Concurrent Events

$$y_{ijc} = \alpha + \beta\,(sch_j \times post_{ic}) + \sum_c \gamma_c\,(GF_j \times cohort_{ic}) + \sum_c \tau_c\,(CR_j \times cohort_{ic}) + \sum_c \lambda_c\,(SD_j \times cohort_{ic}) + X_{ijc}\Theta + \delta_j + \delta_c + \varepsilon_{ijc}$$

- GF_j: the Great Famine
- CR_j: the Cultural Revolution
- SD_j: the Send-Down Movement

Controlling for Concurrent Events

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Female		
sch × post	1.461	0.111	0.213	0.105	-0.010
	(0.384)	(0.049)	(0.050)	(0.063)	(0.015)
Observations	517,919	517,919	517,919	517,919	517,919
			Panel B. Male		
sch × post	1.246	0.004	0.157	0.164	-0.000
	(0.232)	(0.026)	(0.039)	(0.037)	(0.004)
Observations	539,500	539,500	539,500	539,500	539,500
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes
Control concurrent events	Yes	Yes	Yes	Yes	Yes

Threats: Measurement Error

- low-endemic counties report a higher infection rate to receive more fiscal support
- high-endemic counties report a lower infection rate to hide the large disease burden

$$\Rightarrow \text{OLS underestimates}$$

- IV: the transmission of schistosomiasis relied on infected water

$$D_{ijc} = sch_j \times post_{ic} = \alpha + \rho\left(water_j \times post_{ic}\right) + X_{ijc}\Theta + \delta_j + \delta_c + \varepsilon^1_{ijc}$$

$$y_{ijc} = \alpha + \beta\hat{D}_{ijc} + X_{ijc}\Theta + \delta_j + \delta_c + \varepsilon^2_{ijc}$$

- $water_j$: the share of water area in county j

Exclusion Restriction

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
	Panel A. Female in infected areas				
water × post	2.151	0.155	0.262	0.202	-0.002
	(0.533)	(0.062)	(0.068)	(0.083)	(0.011)
Observations	551,273	551,273	551,273	551,273	551,273
	Panel B. Female in uninfected areas				
water × post	-0.295	0.009	-0.020	-0.052	0.064
	(0.531)	(0.118)	(0.111)	(0.126)	(0.052)
Observations	265,065	265,065	265,065	265,065	265,065
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes

▸ Males

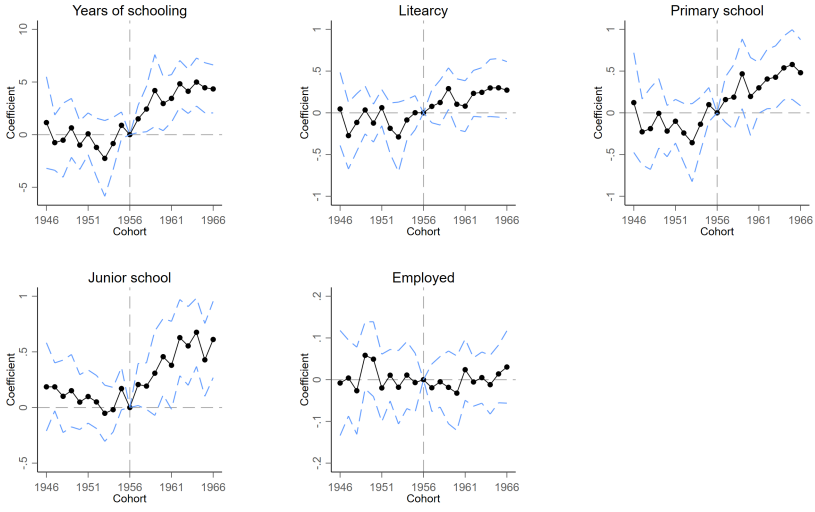
Schistosomiasis is the only channel through which water areas can affect the outcomes

IV Results

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Female		
sch × post	4.293	0.309	0.522	0.403	-0.004
	(1.180)	(0.109)	(0.124)	(0.191)	(0.022)
F statistic	21.564	21.564	21.564	21.564	21.564
Observations	551,273	551,273	551,273	551,273	551,273
			Panel B. Male		
sch × post	3.389	0.035	0.394	0.460	0.007
	(0.792)	(0.088)	(0.108)	(0.104)	(0.006)
F statistic	23.895	23.895	23.895	23.895	23.895
Observations	573,097	573,097	573,097	573,097	573,097
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes

30 pp $\downarrow$ schistosomiasis infection rate $\rightarrow$ 1 year $\uparrow$ expected schooling for males

IV Cohort Analysis (Females)



▸ Males

Selective Mortality

- Mortality rate is low: 1% of the malaria mortality rate (Naghavi et al., 2017)
- Deworming reduced mortality, saving the least healthy people (Li and Wei, 2017)
- Health is positively correlated with educational attainment and labor participation
- Least healthy people survived and lowered the average level of the treatment group
- Explains why we find a weak effect on literacy rate

Critical Period

$$y_{ijc} = \alpha + \beta_1 \left(sch_j \times cohort_{ic}^1\right) + \beta_2 \left(sch_j \times cohort_{ic}^2\right) + X_{ijc}\Theta + \delta_j + \delta_c + \varepsilon_{ijc}$$

- $cohort_{ic}^1 = 1$ if individual i was born 1957-1966 $\rightarrow$ treated in utero
- $cohort_{ic}^2 = 1$ if individual i was born 1952-1956 $\rightarrow$ treated in early childhood
- reference group: if individual i was born 1942-1951 $\rightarrow$ treated in school

Fetal Disease Environment is Crucial

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high
		Panel A. Female		
sch × cohort1	1.757	0.125	0.213	0.166
	(0.419)	(0.053)	(0.051)	(0.062)
sch × cohort2	-0.272	0.018	-0.036	-0.061
	(0.213)	(0.033)	(0.029)	(0.022)
Observations	723,278	723,278	723,278	723,278
		Panel B. Male		
sch × cohort1	1.059	-0.030	0.124	0.135
	(0.291)	(0.034)	(0.043)	(0.041)
sch × cohort2	0.026	0.006	0.024	-0.038
	(0.275)	(0.022)	(0.029)	(0.044)
Observations	751,627	751,627	751,627	751,627
County FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes

Robustness Checks

- Add a linear county-specific time trend
- Mean reversion: add $\bar{y}_{j,46-56} \times cohort_{ic}$
- Heterogeneous treatment effects DID_M (De Chaisemartin and d'Haultfoeuille, 2020)
- Urban sample
- Fake intervention in 1950
- Randomly assigned infection rates

▸ Table Females ▸ Table Males ▸ Figure Females ▸ Figure Males

Results From CFPS: Longer-Run Effect

Using CFPS Data

Three goals:

- replicate the above results in a smaller but finer sample
- evaluate longer-run effects on adult outcomes
- explore intergenerational effects

Three refinement:

- drop individuals who moved out of their birthplaces
- the measure of early-life exposure to schistosomiasis is more precise
- control for parents' education levels (Beach et al., 2022)

Replication Using CFPS

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Female		
sch × post	6.137	0.606	0.942	0.510	0.753
	(2.327)	(0.312)	(0.266)	(0.238)	(0.277)
Observations	3,677	3,677	3,677	3,677	3,677
			Panel B. Male		
sch × post	4.761	0.405	0.417	0.501	-0.163
	(2.460)	(0.340)	(0.302)	(0.276)	(0.344)
Observations	3,804	3,804	3,804	3,804	3,804
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes

Still a large positive effect on educational attainment

Employment Status in Fifties

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Female		
sch × post	6.137	0.606	0.942	0.510	0.753
	(2.327)	(0.312)	(0.266)	(0.238)	(0.277)
Observations	3,677	3,677	3,677	3,677	3,677
			Panel B. Male		
sch × post	4.761	0.405	0.417	0.501	-0.163
	(2.460)	(0.340)	(0.302)	(0.276)	(0.344)
Observations	3,804	3,804	3,804	3,804	3,804
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes

A long lasting effect on women's ability to work over 50 years later

Adult Economic Status

	(1) Household income	(2) Expenditure	(3) Consumption	(4) Net worth
		Panel A. Female		
sch × post	0.969	0.498	0.360	0.116
	(0.452)	(0.382)	(0.358)	(0.608)
Observations	2,951	2,951	2,951	2,951
		Panel B. Male		
sch × post	1.059	0.877	1.044	1.382
	(0.411)	(0.338)	(0.389)	(0.641)
Observations	3,095	3,095	3,095	3,095
County FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

10 pp ↓ schistosomiasis infection rate → 10.6% ↑ household income per capita for males

Suggestive Evidence on Channels: Labor Market

	(1) Working hours	(2) Occupational Prestige
	Panel A. Female	
sch × post	6.169	-3.690
	(2.260)	(10.917)
Observations	3,677	1,803
	Panel B. Male	
sch × post	1.133	19.246
	(3.872)	(9.227)
Observations	3,804	2,315
County FEs	Yes	Yes
Cohort FEs	Yes	Yes
Individual controls	Yes	Yes

Suggestive Evidence on Channels: Health

	(1) Height	(2) BMI	(3) Self-reported health	(4) Interviewer -reported health
			Panel A. Female	
sch × post	-0.656	-3.787	-0.041	0.338
	(4.972)	(2.303)	(0.370)	(0.303)
Observations	3,464	3,431	3,677	3,677
			Panel B. Male	
sch × post	-0.568	1.765	0.278	0.686
	(3.061)	(3.130)	(0.217)	(0.346)
Observations	3,748	3,737	3,804	3,804
County FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

Intergenerational Effects

Three generations:

- G0: G1's parents
- G1: born between 1940-1970 (including treated and untreated)
- G2: G1's children

Two intergenerational effects:

- Between G0 and G1
  - divide individuals into high and low socioeconomic backgrounds
- Between G1 and G2
  - first child (at least age 25 years)
  - two possible channels: better birth outcomes and better family living standard

Heterogeneous Effects By Family Background

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high
	Panel A. Female whose father is illiterate			
sch × post	6.212	0.720	0.906	0.695
	(3.123)	(0.346)	(0.349)	(0.304)
Observations	2,485	2,485	2,485	2,485
	Panel B. Female whose father is literate			
sch × post	0.252	0.074	0.162	-0.128
	(3.497)	(0.375)	(0.452)	(0.353)
Observations	1,672	1,672	1,672	1,672
County FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

▸ Males

Effect on Children's Education

Dependent variable: Generation 2	(1) Schooling (years)	(2) Primary school	(3) Junior high	(4) Senior high	(5) College
		Panel A. Generation 1 Female			
sch × post	6.620	0.261	0.182	0.865	0.349
	(2.432)	(0.222)	(0.312)	(0.396)	(0.273)
Observations	2,090	2,090	2,090	2,090	2,090
		Panel B. Generation 1 Male			
sch × post	0.595	-0.258	-0.622	0.305	0.203
	(5.932)	(0.366)	(0.681)	(0.511)	(0.447)
Observations	1,980	1,980	1,980	1,980	1,980
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes

Conclusion

Evaluate long-term effects of a large-scale deworming program in China

- Improve human capital accumulation
- Improve economic status in adulthood

As many people suffer from schistosomiasis today as they did 50 years ago (Sokolow et al., 2018)

- Offer new justifications for mass deworming efforts in an underdeveloped economy
- Earlier deworming interventions may be more effective

Thank You!

Feedback and comments are much appreciated:

chao.liu1@kellogg.northwestern.edu
gordonliu@nsd.pku.edu.cn

Summary Statistics of the Census Sample

	Female	Male
Age	32.735	32.804
	(6.157)	(6.143)
Ethnicity (Han)	0.919	0.920
	(0.273)	(0.271)
Years of schooling	4.796	6.867
	(3.567)	(3.154)
Literacy	0.744	0.927
	(0.436)	(0.260)
Primary school	0.565	0.797
	(0.496)	(0.402)
Junior high school	0.228	0.423
	(0.420)	(0.494)
Employed	0.933	0.993
	(0.249)	(0.083)
Observations	651,812	676,545

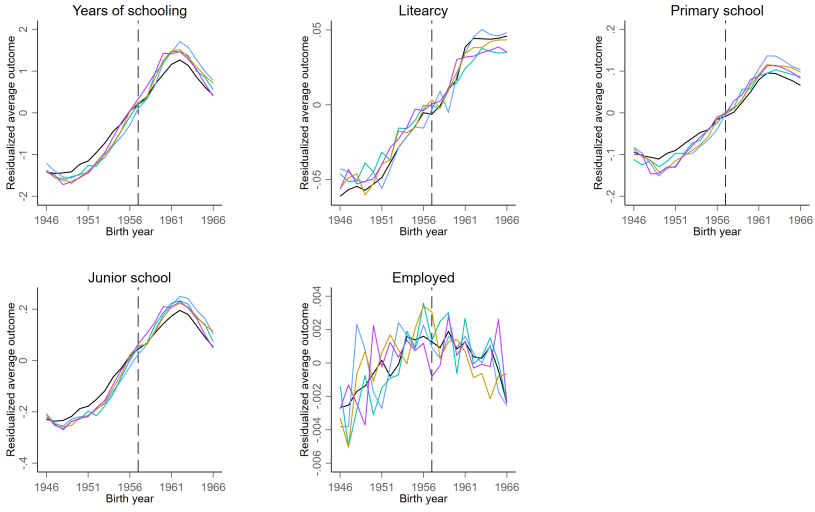
Back

Summary Statistics of the CFPS Sample

	Female	Male
Age	51.955 (8.385)	52.145 (8.575)
Ethnicity (Han)	0.943 (0.232)	0.941 (0.236)
Years of schooling	4.211 (4.253)	6.733 (4.100)
Literacy	0.468 (0.499)	0.739 (0.439)
Primary school	0.481 (0.500)	0.743 (0.437)
Junior high school	0.247 (0.431)	0.461 (0.498)
Employed	0.507 (0.500)	0.636 (0.481)
per capita Household income	7,674 (6,321)	7,703 (6,377)
per capita Expenditure	6,805 (5,264)	6,819 (5,199)
per capita Consumption	5,602 (4,259)	5,691 (4,287)
per capita Net worth	51,000 (69,510)	51,900 (69,764)
Observations	4,850	4,889

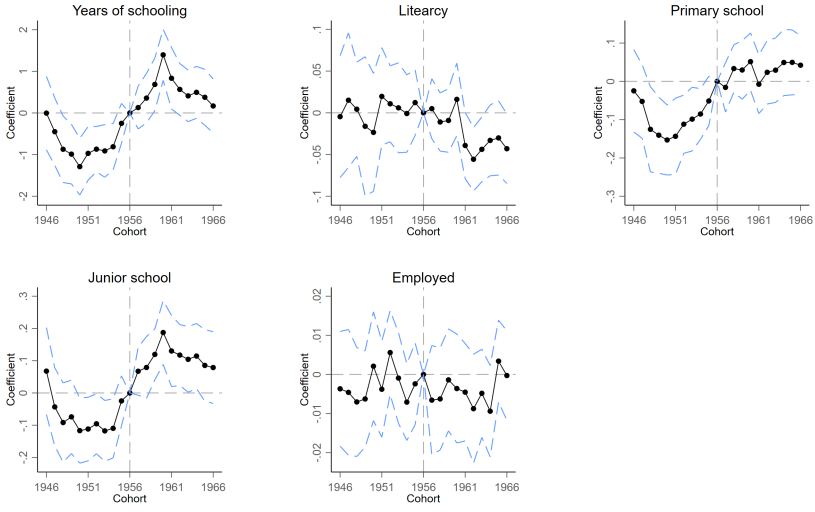
Back

Parallel Trends Assumption (Males)



Back

Cohort Analysis (Males)



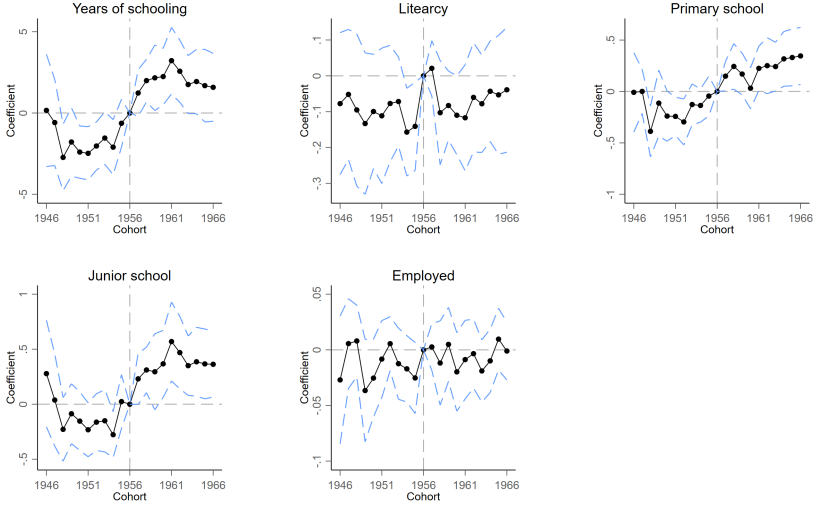
Back

Exclusion Restriction

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high	(5) Employed
			Panel A. Male in infected areas		
water × post	1.714	0.017	0.199	0.233	0.003
	(0.498)	(0.045)	(0.066)	(0.062)	(0.003)
Observations	573,097	573,097	573,097	573,097	573,097
			Panel B. Male in uninfected areas		
water × post	-1.406	-0.065	-0.104	-0.196	-0.004
	(1.118)	(0.112)	(0.114)	(0.118)	(0.013)
Observations	268,284	268,284	268,284	268,284	268,284
County FEs	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes

Back

IV Cohort Analysis (Males)



Back

Robustness Checks (Females)

Dependent variable	(1) Linear trends	(2) Mean reversion	(3) DID_m	(4) Urban sample	(5) Fake event
Schooling (years)	1.324	1.864	3.584	-0.279	-0.272
	(0.274)	(0.346)	(0.595)	(0.434)	(0.220)
Literacy	0.061	0.113	0.036	-0.031	0.041
	(0.030)	(0.034)	(0.060)	(0.030)	(0.037)
Primary school	0.144	0.210	0.387	-0.025	-0.050
	(0.034)	(0.043)	(0.073)	(0.045)	(0.032)
Junior high school	0.150	0.229	0.551	-0.033	-0.066
	(0.035)	(0.046)	(0.098)	(0.071)	(0.019)
Employed	-0.005	0.001	-0.025	-0.025	0.017
	(0.012)	(0.011)	(0.020)	(0.033)	(0.021)
Observations	651,812	651,812	350,149	82,317	376,802

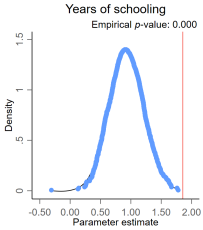
Back

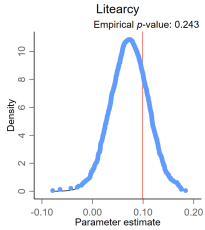
Robustness Checks (Males)

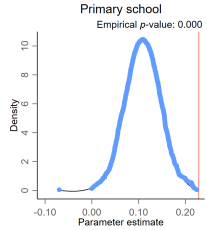
Dependent variable	(1) Linear trends	(2) Mean reversion	(3) DID_m	(4) Urban sample	(5) Fake event
Schooling (years)	1.100	1.022	2.786	0.063	-0.187
	(0.278)	(0.209)	(0.401)	(0.477)	(0.267)
Literacy	-0.011	-0.017	-0.049	0.049	0.024
	(0.013)	(0.010)	(0.034)	(0.022)	(0.021)
Primary school	0.068	0.064	0.260	0.036	0.007
	(0.031)	(0.022)	(0.034)	(0.029)	(0.031)
Junior high school	0.187	0.165	0.445	-0.043	-0.077
	(0.043)	(0.036)	(0.067)	(0.081)	(0.042)
Employed	-0.007	-0.002	-0.002	-0.004	0.003
	(0.005)	(0.003)	(0.004)	(0.024)	(0.003)
Observations	676,545	676,545	361,523	117,743	394,620

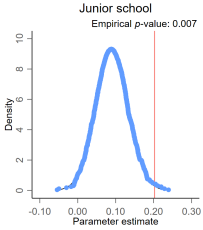
Back

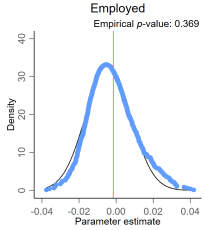
Placebo Tests (Females)





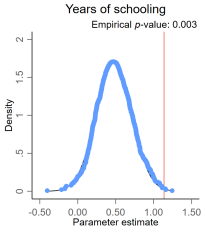


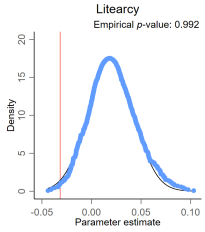


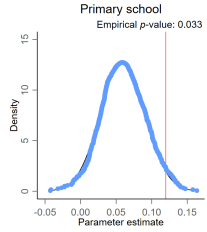


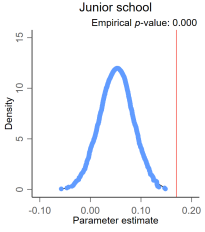
Back

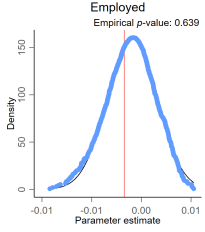
Placebo Tests (Males)











Back

Heterogeneous Effects By Family Background

	(1) Schooling (years)	(2) Literacy	(3) Primary school	(4) Junior high
	Panel A. Female whose father is illiterate			
sch × post	4.980	0.398	0.377	0.436
	(2.159)	(0.351)	(0.269)	(0.329)
Observations	2,556	2,556	2,556	2,556
	Panel B. Female whose father is literate			
sch × post	0.951	0.042	0.058	0.166
	(4.122)	(0.395)	(0.465)	(0.403)
Observations	1,643	1,643	1,643	1,643
County FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

Back