Social Science Inquiry II

Week 4: Joint relationships, part II

Molly Offer-Westort

Department of Political Science,
University of Chicago

Winter 2024

Loading packages for this class

```
> library(ggplot2)
```

Miguel, Edward, & Kremer, Michael (2004). Worms: identifying impacts on education and health in the presence of treatment externalities.

An aside on Nobels

Figure: Esther Duflo, Michael Kremer, and Abhijit Banerjee

An aside on Nobels

2019 economics nobel awarded for: "their experimental approach to alleviating global poverty." Why is this a big deal?

- One way to understand human behavior: posit a theory. E.g., propose a utility function.

→ Deduce implications of theory.

→ (maybe) Test implications against data.

→ Update and revise theory.

Experiments

How do we know if policies are working?

- Duflo, Banerjee, Kremer motivated by work of development agencies.
- Interventions are large scale, complicated, and hard to deduce how "effective" they should be based on theory alone.
- A solution: build rigorous evidence for policy efficacy.
- One way to build evidence: experiments.

Joint relationships

- Correlation: X and Y are correlated. E.g., $\uparrow$ age and $\uparrow$ income.
- Causation: W *moves* Y. E.g., when I fertilize the plants, I get a higher yield.
  - Implies a *counterfactual*. . If I *hadn't* moved W, we would see a different outcome for Y.
  - Recall Holland (1986): "No causation without manipulation."

Joint relationships: notation

- Dependent variable: (Y), the thing we want to study variation in; also "outcome", "response".
- Independent variable: (D or W) in causal settings, the thing that we manipulate. Also "treatment." In non-causal settings, something that may have some natural variation in the population associated with the dependent variable.

Potential outcomes: notation

Assume W is binary.

$$Y_i = \begin{cases} Y_i(0) & W_i = 0 \\ Y_i(1) & W_i = 1 \end{cases}$$

Why is $\mathrm{E}[Y_i(1)]$ different from $\mathrm{E}[Y_i|W = 1]$?
(What is the expectation *over*?)

Individual treatment effect

$$\tau_i = Y_i(1) - Y_i(0)$$

We're interested in the treatment effect, the difference in potential outcomes had we applied treatment W, vs. if we had not.

- What we can't see: τ_i
- Fundamental problem of causal inference: we can't see *both* potential outcomes for a given unit *at the same time*.
- Holland (1986)'s assumptions: some structure we can impose on the data to back out counterfactuals. (We usually don't know if assumptions are true.)
- Experiments use the independence assumption. . . this one we have some more control over.

Independence

- Units that receive treatment look, on average, like units that do not.
- In experiments, we assign treatment *randomly*, so we know that there is no inherent difference between subjects that *could have* received treatment and *could have* received control–all subjects *could have* received all versions of treatment, hypothetically. (The subjunctive mood of research design . . .)
- Why is this so special?
- What this give us: $\mathrm{E}\left[Y_i | W_i = 1\right] = \mathrm{E}\left[Y_i(1)\right]$
- We can back out an estimate of the average treatment effect:

$$\tau = \mathrm{E}\left[Y_i(1) - Y_i(0)\right]$$

Back to our study at hand. . .

Reading papers

What to get out of reading a research paper:

- What is the main question of the paper?
- What method do the authors use to address the question? For empirical papers:
  - Data (Where does it come from/how is it generated? What is the sample population? What is being measured?)
  - Research design/strategy
  - Statistical tools
- What is the answer that the authors get to the main question?

How would you answer these questions with the Miguel and Kremer (2004) paper?

The problem

- Is deworming as a large-scale policy effective?
- Why don't we already have good evidence?

Research design

- 75 rural Kenyan primary schools in Busia region phased into treatment in a randomized order.
- Why is treatment assigned at the school level?

Spillovers (or treatment externalities)

- In general, we would like to assume $Y_i(W_i)$ does not change depending on some W_j for $j \neq i$.
- This doesn't always hold. How do helminth infections spread?
- One way to address this: make sure that everyone in one area gets treatment, or everyone gets control.
- Useful for understanding impacts of *policies*. What would outcomes look like if *everyone* got treatment, vs. if *everyone* got control.

Random assignment

- 3 groups of 25 schools each
- "The schools were first stratified by administrative subunit (zone) and by their involvement in other nongovernmental assistance programs, and were then listed alphabetically and every third school was assigned to a given project group."
  - What is the stratification doing?
  - Is this assignment random?

Pre-experimental data

TABLE I

1998 AVERAGE PUPIL AND SCHOOL CHARACTERISTICS, PRE-TREATMENT[a]

	Group 1 (25 schools)	Group 2 (25 schools)	Group 3 (25 schools)	Group 1 – Group 3	Group 2 – Group 3
Panel A: Pre-school to Grade 8					
Male	0.53	0.51	0.52	0.01	−0.01
				(0.02)	(0.02)
Proportion girls <13 years, and all boys	0.89	0.89	0.88	0.00	0.01
				(0.01)	(0.01)
Grade progression (= Grade − (Age − 6))	−2.1	−1.9	−2.1	−0.0	0.1
				(0.1)	(0.1)
Year of birth	1986.2	1986.5	1985.8	0.4**	0.8***
				(0.2)	(0.2)
Panel B: Grades 3 to 8					
Attendance recorded in school registers (during the four weeks prior to the pupil survey)	0.973	0.963	0.969	0.003	−0.006
				(0.004)	(0.004)
Access to latrine at home	0.82	0.81	0.82	0.00	−0.01
				(0.03)	(0.03)
Have livestock (cows, goats, pigs, sheep) at home	0.66	0.67	0.66	−0.00	0.01
				(0.03)	(0.03)
Weight-for-age Z-score (low scores denote undernutrition)	−1.39	−1.40	−1.44	0.05	0.04
				(0.05)	(0.05)
Blood in stool (self-reported)	0.26	0.22	0.19	0.07**	0.03
				(0.03)	(0.03)
Sick often (self-reported)	0.10	0.10	0.08	0.02**	0.02**
				(0.01)	(0.01)
Malaria/fever in past week (self-reported)	0.37	0.38	0.40	−0.03	−0.02
				(0.03)	(0.03)
Clean (observed by field workers)	0.60	0.66	0.67	−0.07**	−0.01
				(0.03)	(0.03)
Panel C: School characteristics					
District exam score 1996, grades 5–8[b]	−0.10	0.09	0.01	−0.11	0.08
				(0.12)	(0.12)
Distance to Lake Victoria	10.0	9.9	9.5	0.6	0.5
				(1.9)	(1.9)
Pupil population	392.7	403.8	375.9	16.8	27.9
				(57.6)	(57.6)
School latrines per pupil	0.007	0.006	0.007	0.001	−0.000
				(0.001)	(0.001)
Proportion moderate-heavy infections in zone	0.37	0.37	0.36	0.01	0.01
				(0.03)	(0.03)
Group 1 pupils within 3 km[c]	461.1	408.3	344.5	116.6	63.8
				(120.3)	(120.3)
Group 1 pupils within 3–6 km	844.5	652.0	869.7	−25.1	−217.6
				(140.9)	(140.9)

Data

- Where does it come from/how is it generated?
- What is the sample population?
- What is being measured?

Treatment and phasing

- What does the phasing do for us?
- Sometimes studies need to account for more than just research objectives.

Treatment and phasing

TABLE III

PROPORTION OF PUPILS RECEIVING DEWORMING TREATMENT IN PSDP[a]

	Group 1		Group 2		Group 3	
	Girls <13 years, and all boys	Girls ≥ 13 years	Girls <13 years, and all boys	Girls ≥ 13 years	Girls <13 years, and all boys	Girls ≥ 13 years
	Treatment		*Comparison*		*Comparison*	
Any medical treatment in 1998 (For grades 1–8 in early 1998)	0.78	0.19	0	0	0	0
Round 1 (March–April 1998), Albendazole	0.69	0.11	0	0	0	0
Round 1 (March–April 1998), Praziquantel[b]	0.64	0.34	0	0	0	0
Round 2 (Oct.–Nov. 1998), Albendazole	0.56	0.07	0	0	0	0
	Treatment		*Treatment*		*Comparison*	
Any medical treatment in 1999 (For grades 1–7 in early 1998)	0.59	0.07	0.55	0.10	0.01	0
Round 1 (March–June 1999), Albendazole	0.44	0.06	0.35	0.06	0.01	0
Round 1 (March–June 1999), Praziquantel[b]	0.47	0.06	0.38	0.06	0.01	0
Round 2 (Oct.–Nov. 1999), Albendazole	0.53	0.06	0.51	0.08	0.01	0
Any medical treatment in 1999 (For grades 1–7 in early 1998), among pupils enrolled in 1999	0.73	0.10	0.71	0.13	0.02	0
Round 1 (March–June 1999), Albendazole	0.55	0.08	0.46	0.08	0.01	0
Round 1 (March–June 1999), Praziquantel[b]	0.53	0.07	0.45	0.07	0.01	0
Round 2 (Oct.–Nov. 1999), Albendazole	0.65	0.09	0.66	0.11	0.01	0

[a]Data for grades 1–8. Since month of birth information is missing for most pupils, precise assignment of treatment eligibility status for girls born during the "threshold" year is often impossible; all girls who turn 13 during a given year are counted as 12 year olds (eligible for deworming treatment) throughout for consistency.

[b]Praziquantel figures in Table III refer only to children in schools meeting the schistosomiasis treament threshold (30 percent prevalence) in that year.

Estimation strategy

$$Y_{ijt} = a + \beta_1 \times T_{1it} + \beta_2 \times T_{2it} + X'_{ijt}\delta + \sum_d \left(\gamma_d \times N^T dit\right) + \sum_\delta (\phi_d \times N_{dit}) + u_i + e_{ijt}$$

- Y_{ijt} is individual health/education outcome for school i, student j, and year $t \in 1, 2$.
- T_{1it} and T_{2it} are indicators for treatment in years 1 and 2
- X'_{ijt} are school/pupil characteristics
- N_{dit} total number of pupils in primary schools at distance d from school i in year t
- N^T_{dit} number of these pupils in schools randomly assigned to deworming treatment

Within-school externalities

- The coefficients on T_{1it} and T_{2it}, β_1 and β_2 capture *direct effects* on the treated, as well as *indirect effects* on untreated in treatment schools.
- Why is it challenging to decompose these components?
- Proposed solution: make assumptions about who got treatment, compare untreated in Group 1 in 1998 to those who *would be* untreated in Group 2 in 1999 (why might this work?)
- Result: evidence of large good externalities; infections lower among Group 1 untreated
- Why might externalities be so large within schools?

Geographic component

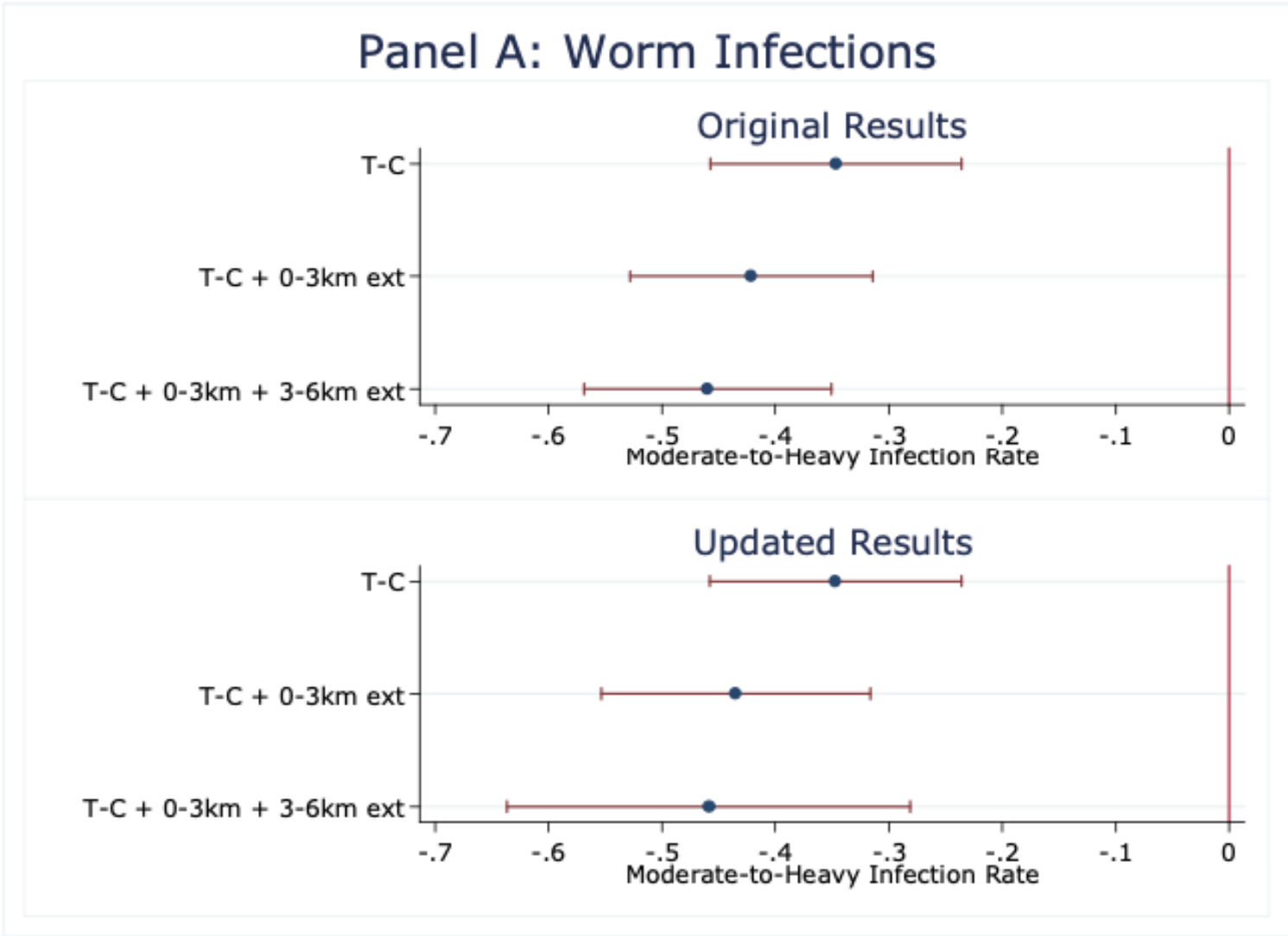
Since local population density may affect disease transmission, and since children who live or attend school near treatment schools could have lower environmental exposure to helminths, which would lead to less reinfection and lower worm burdens, worm burden may depend on both the total number of primary school pupils (N_{dit}) and the number of those pupils in schools randomly assigned to deworming treatment (N_{dit}^{T}) within a certain distance from school i in year t of the program.

- the γ_d capture deworming treatment externalities across schools

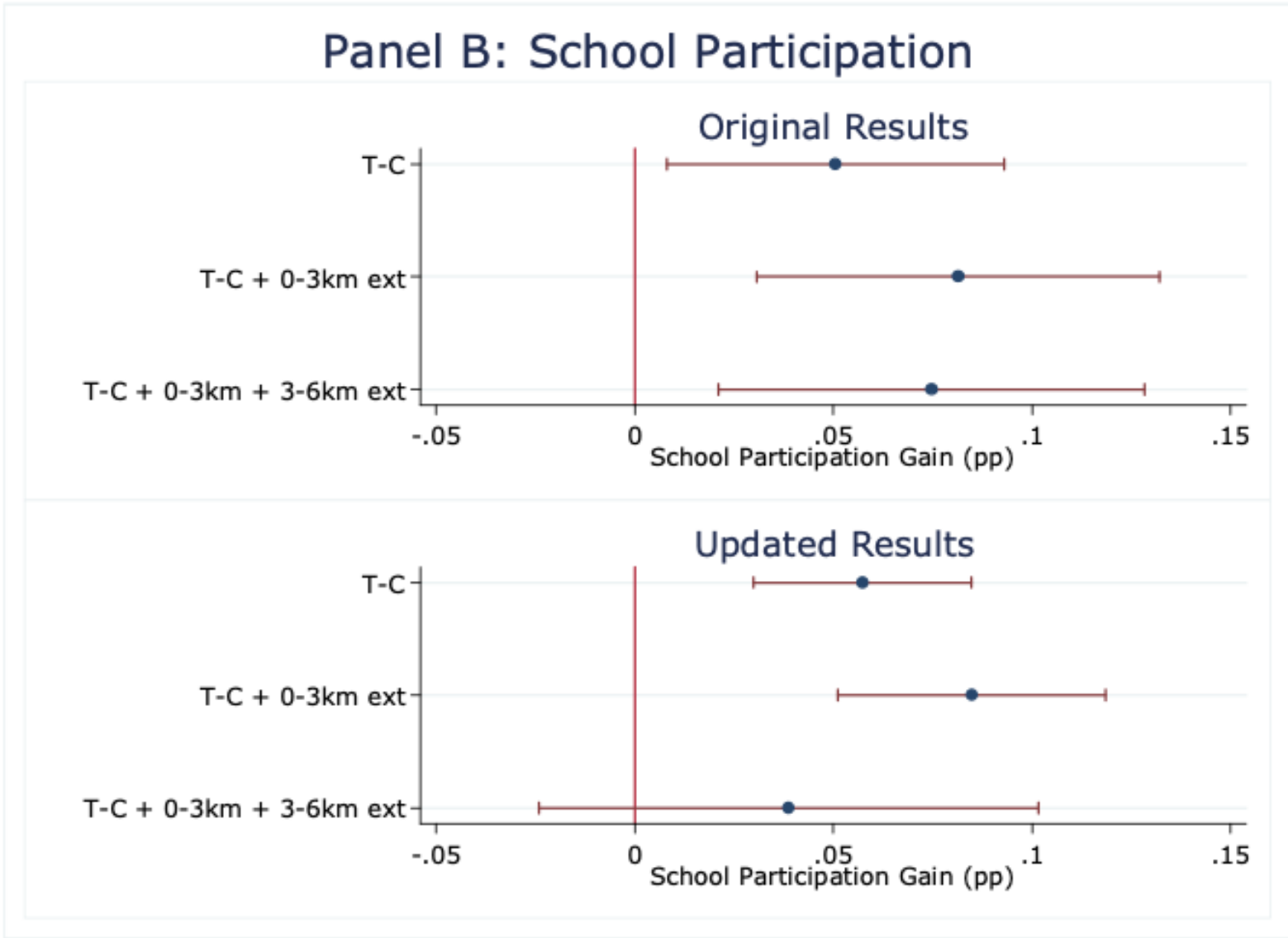
What is the answer that the authors get to the main question?

Results

Figure B1. Original vs. updated "overall effect", with 95% confidence intervals

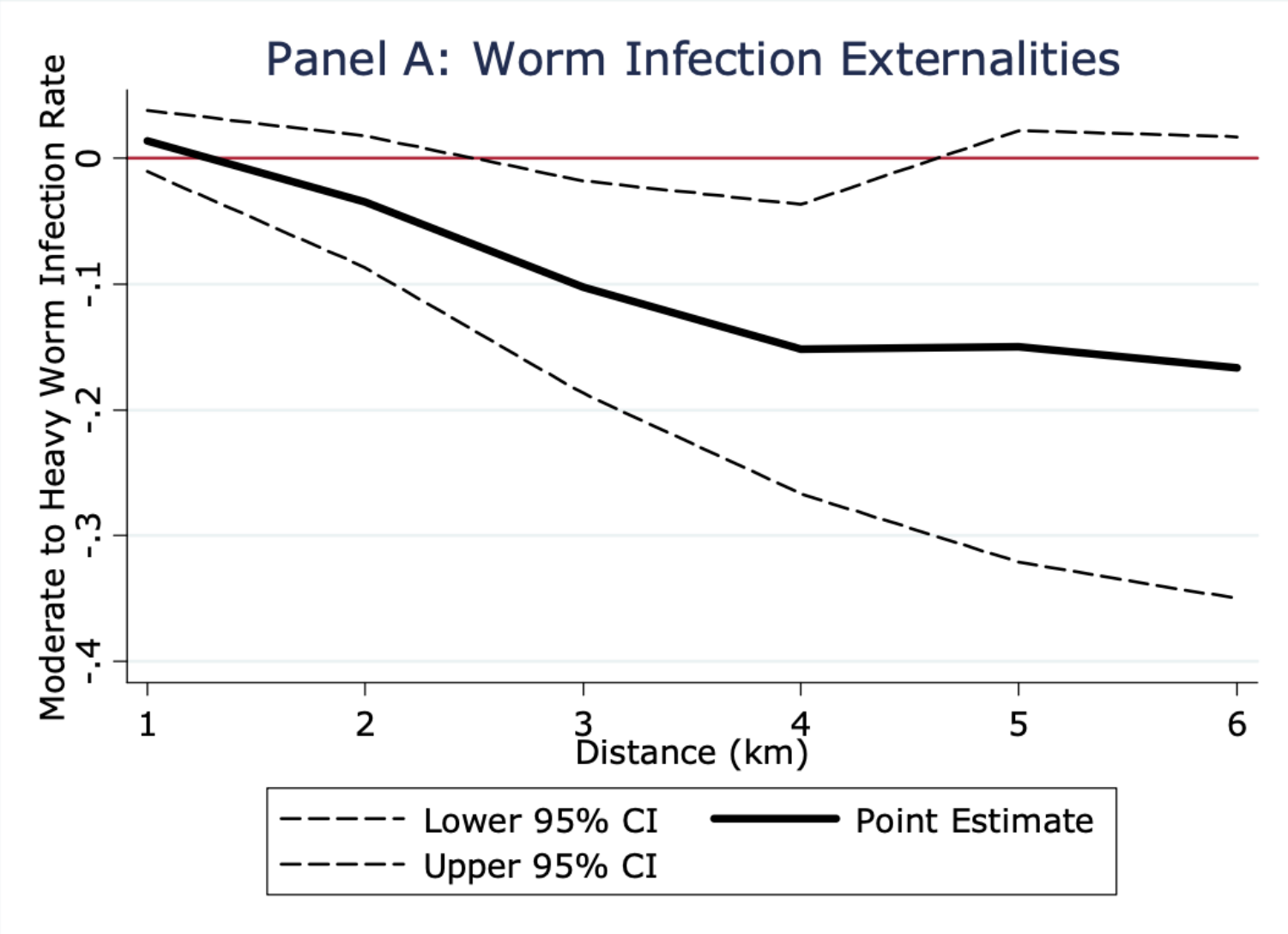


Results

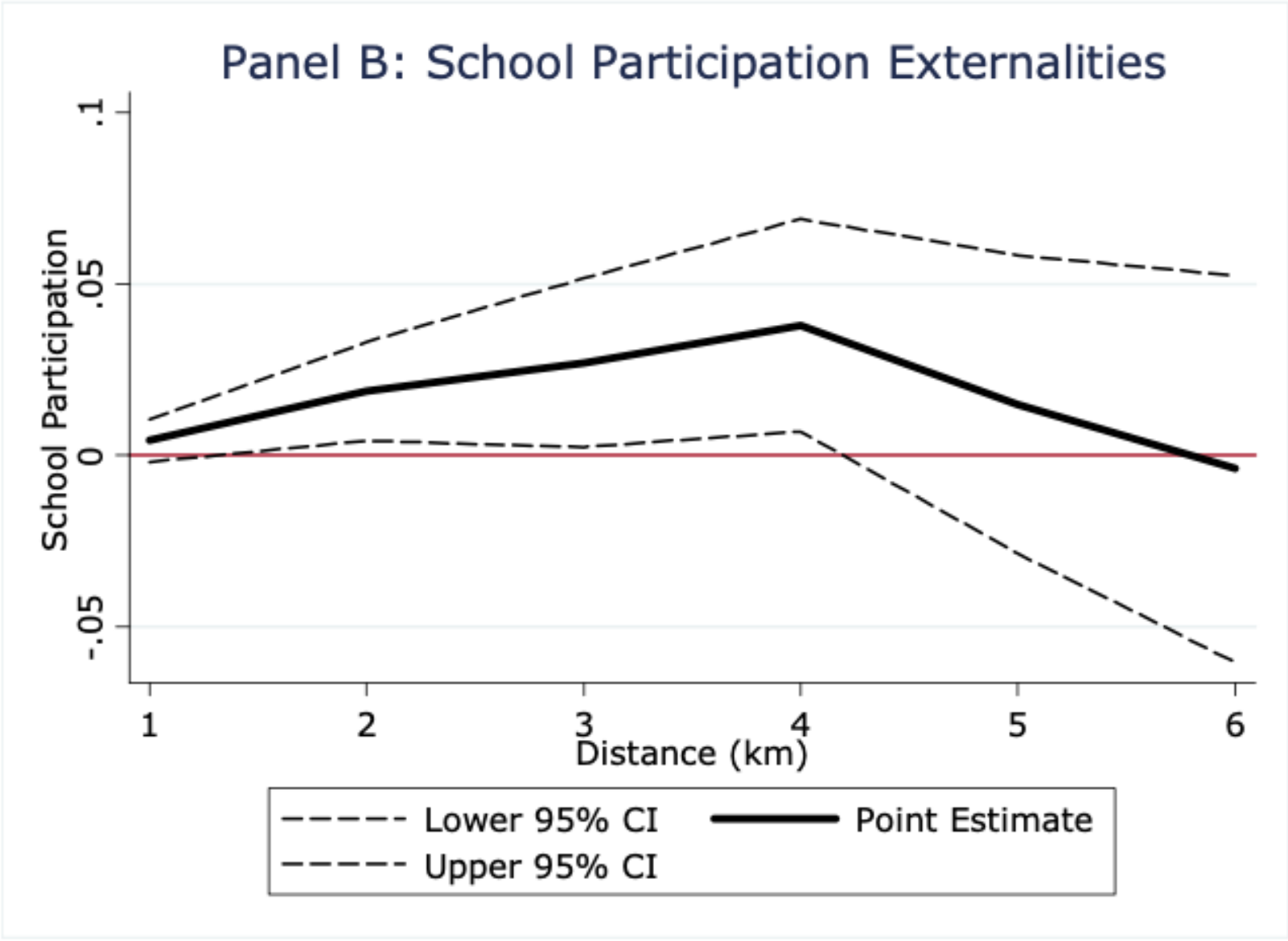

Panel B: School Participation
Original Results
T-C
T-C + 0-3km ext
T-C + 0-3km + 3-6km ext
-.05
0
.05
.1
.15
School Participation Gain (pp)
Updated Results
T-C
T-C + 0-3km ext
T-C + 0-3km + 3-6km ext
-.05
0
.05
.1
.15
School Participation Gain (pp)

Results

Figure B2. Average externality impacts at various distances

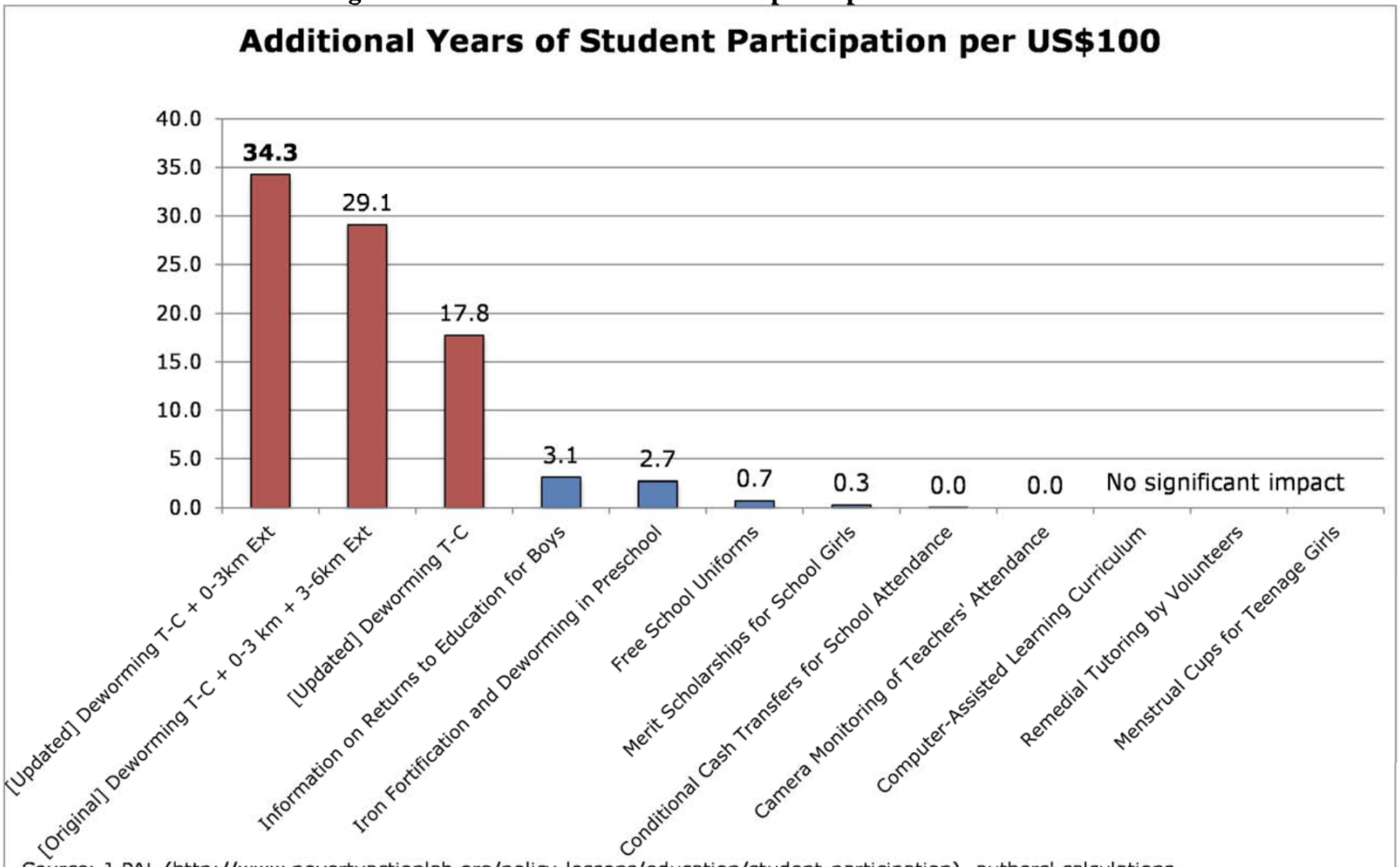


Results


Panel B: School Participation Externalities
School Participation
.1
.05
0
-.05
1
2
3
4
5
6
Distance (km)
Lower 95% CI
Upper 95% CI
Point Estimate

Results

Figure B3: Cost-effectiveness of school participation interventions



Source: J-PAL (http://www.povertyactionlab.org/policy-lessons/education/student-participation), authors' calculations.
Notes: Some values are adjusted for inflation but the deworming costs are not. The cost of deworming is US$ 0.49, as in Miguel and Kremer (2004). This is likely an overestimate, since the per-child cost in the 2009 Kenya National School-Based Deworming Program was US$0.36 (http://www.dewormtheworld.org/our-work/kenya-national-school-based-deworming-program for more details).

References I

Holland, P. W. (1986). Statistics and causal inference. Journal of the American statistical Association, 81(396):945–960.

Miguel, E. and Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. Econometrica, 72(1):159–217.