“Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming”

Owen Ozier

ONLINE APPENDIX

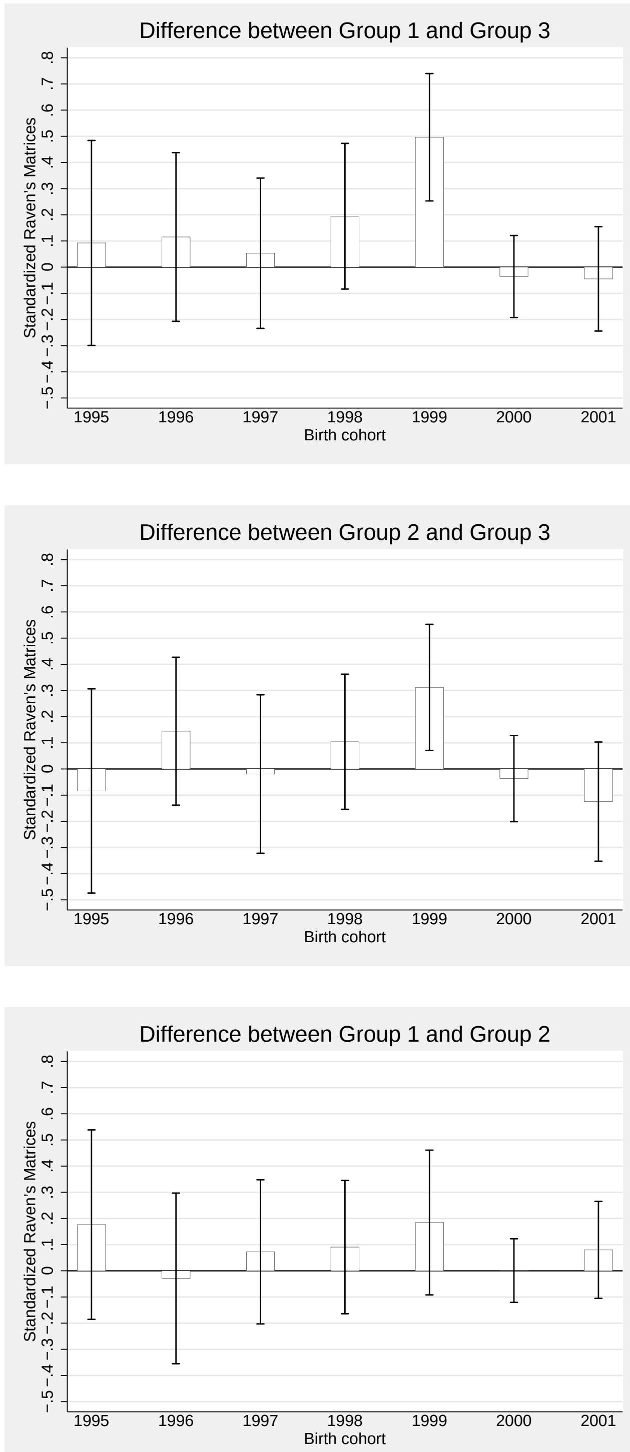
A Appendix

Table A1: Cognitive measures: Principal Components

Principal component:	(1)	(2)	(3)	(4)	(5)	(6)
Verbal Fluency: Foods	0.3612	-0.6743	0.0027	0.2230	0.5550	-0.2390
Verbal Fluency: Animals	0.4443	-0.4238	-0.0030	-0.0594	-0.5293	0.5825
Digit Span Forwards	0.3814	0.2288	0.6677	-0.5286	0.2687	0.0693
Digit Span Backwards	0.3875	0.3937	0.2948	0.7742	-0.0915	-0.0117
Vocabulary: PPVT	0.4762	0.0878	-0.2600	-0.2420	-0.4023	-0.6910
Raven’s Matrices	0.3870	0.3882	-0.6322	-0.0965	0.4115	0.3481
Explained variance:	0.4665	0.6214	0.7464	0.8482	0.9344	1.0000

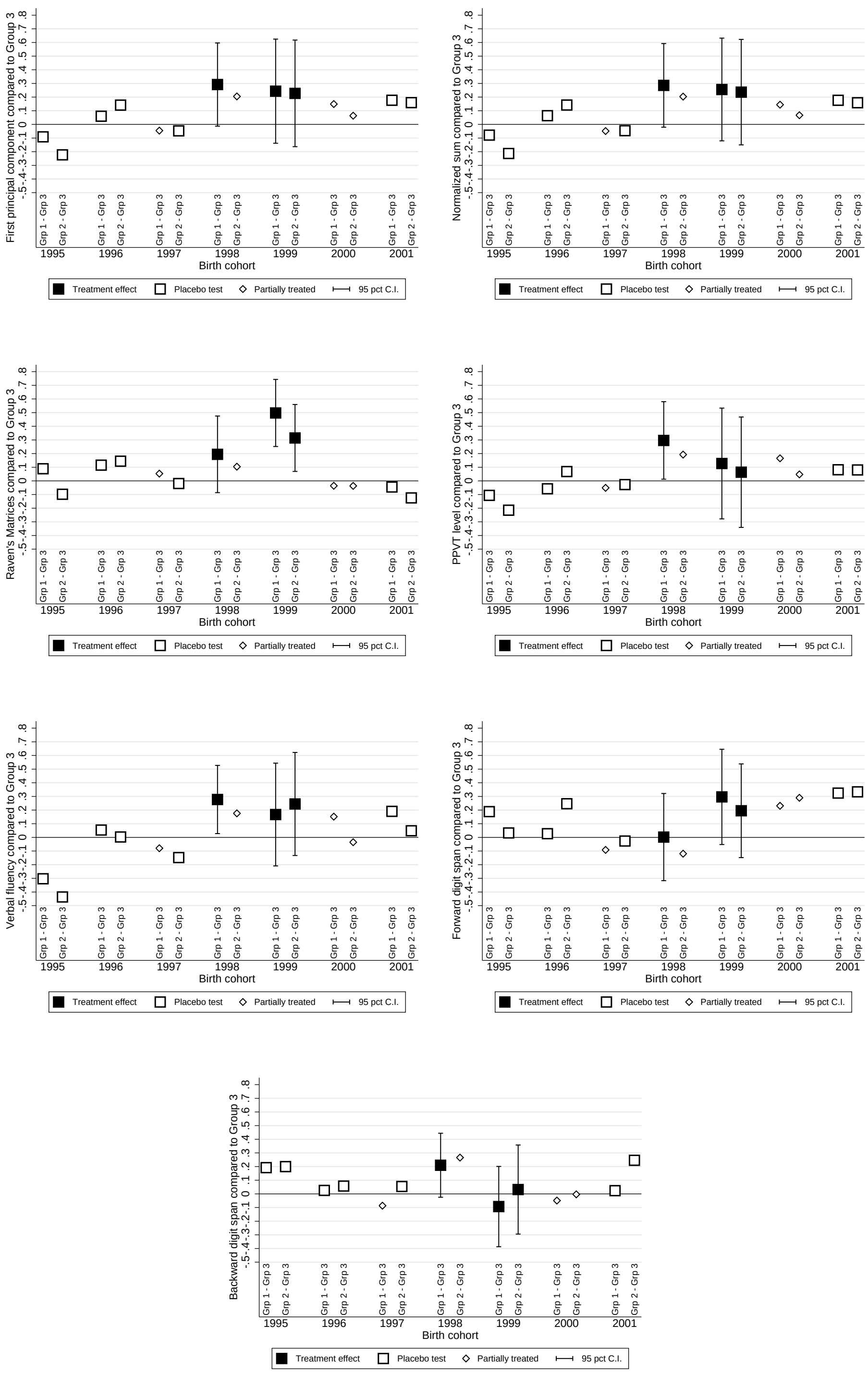
Note that weights for the first principal component are almost equal across the different cognitive measures. The lowest weight is for “Verbal Fluency: Foods,” perhaps the noisiest measure in part because it was the first exercise in the cognitive module. Low R^2 for regressions with this outcome also speak to its relative noisiness.

Figure A1: All pairwise comparisons of Raven's Matrices between groups



Note: Bars represent 95-percent confidence intervals.

Figure A2: Full set of cognitive comparisons against Group 3



Note: Bars represent 95-percent confidence intervals. See notes for Figure 2.

Table A2: Cognitive measure correlations

	Fluency: Foods	Fluency: Animals	Digit Span Forwards	Digit Span Backwards	Raven's Matrices	Vocab: PPVT
Foods	1.0000					
Animals	0.5007	1.0000				
Digit Span Forwards	0.2400	0.3389	1.0000			
Digit Span Backwards	0.2323	0.3183	0.3778	1.0000		
Raven's Matrices	0.2218	0.3014	0.2742	0.3477	1.0000	
PPVT	0.3490	0.5204	0.3989	0.3899	0.5083	1.0000

Table A3: Cognitive performance (first principal component, normalized) as a function of observables

	All			Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Grade	0.451***	0.382***	.	0.459***	0.407***	.	0.449***	0.355***	.
	(0.011)	(0.007)		(0.015)	(0.009)		(0.016)	(0.01)	
Age	-0.089***	.	0.261***	-0.069***	.	0.292***	-0.118***	.	0.226***
	(0.011)		(0.009)	(0.015)		(0.012)	(0.015)		(0.012)
Constant	-0.872***	-1.607***	-3.021***	-1.071***	-1.652***	-3.369***	-0.606***	-1.554***	-2.624***
	(0.095)	(0.032)	(0.103)	(0.134)	(0.043)	(0.146)	(0.133)	(0.048)	(0.145)
Observations	2583	2583	2585	1372	1372	1373	1203	1203	1204
R^2	0.555	0.543	0.254	0.582	0.576	0.287	0.532	0.51	0.218

Table A4: Cognitive performance (normalized) as a function of observables

	Outcome											
	Vocabulary: PPVT		Verbal fluency: Foods		Verbal fluency: Animals		Memory: Digit Span Forwards		Memory: Digit Span Backwards		Reasoning: Raven's Matrices	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Grade	0.372***	.	0.196***	.	0.279***	.	0.219***	.	0.222***	.	0.247***	.
	(0.007)		(0.009)		(0.008)		(0.009)		(0.009)		(0.009)	
Age	.	0.261***	.	0.143***	.	0.212***	.	0.118***	.	0.139***	.	0.17***
		(0.009)		(0.01)		(0.009)		(0.01)		(0.01)		(0.009)
Constant	-1.565***	-3.012***	-0.81***	-1.642***	-1.169***	-2.444***	-0.918***	-1.363***	-0.936***	-1.608***	-1.034***	-1.960***
	(0.032)	(0.101)	(0.042)	(0.112)	(0.039)	(0.107)	(0.042)	(0.115)	(0.043)	(0.115)	(0.041)	(0.111)
Observations	2661	2665	2664	2667	2664	2667	2633	2635	2591	2593	2663	2667
R^2	0.519	0.255	0.145	0.078	0.292	0.168	0.179	0.052	0.184	0.072	0.227	0.107

A.1 Health versus cognitive mechanism

In the "School Assistance Program" (SAP) program that Glewwe et al. (2009) studied, the timing of resources delivered to schools varied by study arm. The key variants in their analysis were "SAP group 1," in which textbooks were distributed to schools in 1996, and the "comparison," or "SAP group 4," in which grants were delivered to schools in 2000. High-performing school-age children exhibited better academic performance in the short term as a consequence of the textbook distribution, as Glewwe et al. (2009) show in their Table 8. Thus, since SAP beneficiaries experienced cognitive benefits, I can test whether impacts of this kind affect younger siblings and neighbors in this setting. If SAP had spillover benefits on younger cohorts, it increases the plausibility of a cognitive channel for the deworming externalities I examine. If it did not, a health channel is more likely. (This test is, of course, not definitive, as the cognitive mechanisms potentially at work for the SAP and deworming interventions' direct beneficiaries are surely different.)

The SAP schools have some intersection with the deworming program schools; roughly one quarter of the sample surveyed for this paper attends a school that is included in either "SAP group 1" or "SAP group 4." Thus, in this dataset, I can exploit the fact that from 1996 through 1999, the "SAP group 1" schools were differentially advantaged in relation to the "SAP group 4" schools. To test the cognitive spillovers that would be most analogous to the present analysis of deworming spillovers, I examine cohorts which were no more than one year old by that time, following the specification below:

$$Y_i = \beta_5^{SAP} \cdot SAPcohort_i \cdot SAP1_i + \gamma_{G1} \cdot SAP1_i + \gamma_{cohort} \cdot SAPcohort_i + \sum_{A,S,Y} \gamma_{ASY} D_{Age_i=a} \cdot D_{Sex_i=S} \cdot D_{Year_i=Y} + \epsilon_{5i} \ . \quad (5)$$

In Equation 5, $SAP1_i$ is an indicator for whether the individual is in an "SAP group 1" school, as opposed to an "SAP group 4" school; the sample for this estimation is restricted to those two groups. The indicator $SAPcohort_i$ designates whether the individual is born in a cohort that would have received differential spillovers from SAP (more on this below). The coefficient of interest, β_5^{SAP}, is on the interaction of these two indicators. As before, fixed effects for the interaction of age, sex, and data collection year are also included.

I also face a decision about which cohorts to designate as having (plausibly) received differential spillovers from SAP, since the advantaged position of "SAP group 1" schools lasted from 1996 through 1999. To be analogous to the analysis of deworming spillovers by the first year of life, I can either consider those who were born either between 1995 and 1998 (no more than *just under one year old* when textbooks arrived, but also no more than that age when the grants arrived in the comparison schools), or between 1994 and 1997 (no more than *two years old*). I show results for both.

In Table A5, for those children in schools that benefitted from SAP's textbook and grant distribution (SAP groups 1 and 4), I test whether the arrival of textbooks by the first year of a child's life improves outcomes today. Though the subsample of schools leaves this test with only modest power, I find no consistent (or statistically significant) evidence of this pattern.

Thus, health remains the most plausible channel for deworming spillovers.

Table A5: Testing for spillovers on younger sibings from textbook distribution

Outcome:	Raven's Matrices		Principal Component	
Age encoding:	1994-1997	1995-1998	1994-1997	1995-1998
	[1]	[2]	[3]	[4]
Textbooks arrive early	−0.105	−0.038	0.145	0.058
	(0.144)	(0.132)	(0.149)	(0.138)
Observations	663	663	663	663

The table above presents tests of cognitive responses to the intervention described by Glewwe, Moulin, and Kremer (2009). This table is discussed in Appendix Section A.1, and shows the β_5^{SAP} coefficient from estimation of Equation 5. The outcome in columns 1 and 2 is standardized Raven's Matrices; in columns 3 and 4, it is the first principal component of all cognitive outcome measures. The even-numbered columns differ from the odd-numbered columns in the way a "treated" cohort is defined. The textbook program began in "SAP group 1" schools in 1996; the comparison schools, "SAP group 4," received grants in 2000. Cohorts just under 1 year old at the start of the textbook intervention were born in 1995; those just over one year old were born in 1994. Likewise, since the comparison group was treated with grants in 2000, those just under one year old were were born in 1999; those just over were born in 1998. Thus, the odd-numbered columns show the coefficients on the interaction between being in SAP group 1 and being born in the years 1994-1997, cohorts who may have benefited from intervention spillovers in SAP group 1 but not yet in SAP group 4; the even-numbered columns show analogous results but with the indicator for cohorts born in years 1995-1998. Standard errors, in parentheses, are clustered at the school-cohort level.

A.2 Tests of threats to validity

A.2.1 Demographics

Changes in the composition of cohorts in this study that are due to variation in school-based deworming treatment by community could potentially confound the analysis. Changes of this sort could arise if deworming changed mortality rates, leaving disproportionately healthy children as survivors. Note, however, that a mortality mechanism does not have empirical support from studies that have examined it directly; see Awasthi, Peto, Read, Richards, Pande, Bundy and DEVTA Team (2013). One could also imagine that if adults adjusted their fertility patterns in response to school-based mass deworming—in either direction—such adjustment might change the interpretation of estimated effects. Bleakley and Lange (2009), for example, document decreases in fertility in response to the Rockefeller Sanitary Commission deworming work in the US South. A simple approach to mortality and fertility is to test whether respondents exposed to spillovers from school-based deworming from birth have more or fewer siblings than those who were exposed only later. An analogous approach is to test whether the actual quantity of age-eligible respondents in each school systematically varies as a function of deworming exposure. Tests of these hypotheses are shown in Appendix Table A6. I find no evidence of either pattern using either approach.

A.2.2 Migration

If school-based deworming induced out-migration differentially among the families of those treated earlier or later, the set of children I find in 2009 and 2010 might be differentially selected by treatment arm, possibly resulting in biased coefficient estimates. The ideal test would look for families who had lived in the study area at the time of the deworming intervention, and would check whether they had moved.

I construct a test for this by taking advantage of the first round of the Kenya Life Panel Survey (KLPS1), which took place from 2003 to 2005. Although it followed the older children who had originally been in school at the time of deworming, rather than these children's families, the youngest respondents in KLPS1 were still in primary school and (generally) living with their parents at the time of the follow-up, so their location tells us their parents' location. For example, those who were in their second year of primary school (Standard 2) in 1998 would have been in Standard 7, Standard 8, or the first Form of secondary school from 2003 to 2005, if they had never repeated a grade. Most students repeat at least one grade, however, meaning most of the youngest respondents were still in primary school: of the pupils who had been in Standard 2 in 1998, and were interviewed as late as 2005, more than 95 percent were not yet in secondary school; of 1998 Standard 2 pupils interviewed in 2003 and 2004, more than 99 percent were not yet in secondary school. The KLPS1 respondents were randomized into two waves of surveying, so I can also focus a test on the first wave (from 2003 to 2004) to be sure not to confuse a departure for secondary schooling or marriage for the migration of the respondent's parents.

In Appendix Table A7, I show four versions of this test, using either Standard 2 or both Standards 2 and 3, and using either the first wave of KLPS1 or both the first and second waves. The outcome variable is whether the KLPS1 interview took place outside Busia District. The mean is around six percent, but it does not vary by treatment arm in any specification, and the standard errors are relatively small (1.4 percentage points in Column 1, for example). In this way, I establish (with some precision) that there was not differential out-migration of families by original treatment arm.

A.2.3 School academic characteristics

Although Figure 1 shows no evidence of systematic differences between the schools in the original three treatment arms in terms of outcomes of interest, it is possible that some part of the variation in cognitive abilities measured in this study is due to pre-existing differences in levels of academic ability that vary at the community level.

I first test whether there were such differences using school-level average scores on the Standard 8 primary school leaving examination, the Kenya Certificate of Primary Education (KCPE). Tests of differences by PSDP treatment arm are shown in Appendix Table A8. The last column shows that, in the five years prior to the start of PSDP, there were no differences on average between the three arms. The first five columns show that none of fifteen pairwise tests of differences is significant at the 5 percent level, though Groups 2 and 3 differed slightly (significant at the 10 percent level) in a single year, 1994.

I then test whether these small variations are responsible for the patterns attributed to deworming spillovers in this paper, by replicating the analysis shown in Table 2, but including five separate controls for KCPE scores in the five years before KCPE. Results are shown in Appendix Table A9. This robustness check involves discarding more than ten percent of observations, since I am missing some schools' KCPE records from the mid-1990s. Despite this reduction in sample size (and thus power), the magnitudes and patterns of statistical significance remain largely unchanged.

Table A6: Testing for fertility or mortality responses to deworming

	Sibling N	Respondent N
	[1]	[2]
Deworming before 1 (main specification)	0.056	0.062
	(0.055)	(0.799)
Deworming before birth (alternative for fertility)	0.061	0.862
	(0.049)	(0.832)
Observations	15630	1740

The table above presents tests of a fertility response to school-based deworming in the community. In the first column, the outcome variable is the number of younger siblings reported by the respondent. In the second column, observations have been aggregated at the level of the $\{$ data collection year $\times$ birth year $\times$ gender $\times$ migration indicator $\times$ school $\}$. Thus, in the second column, the outcome is simply the count of observations in these bins. The first row presents the same specification as elsewhere in the paper, showing an indicator for school-based deworming arriving in the community in the respondent's year of birth or earlier; the second row presents an alternative specification, using an indicator for whether deworming arrived before the respondent was born. In any of the four cells, a significant coefficient could indicate a change in fertility in response to mass school-based deworming starting in that community. In the first column, standard errors are clustered at the school-cohort level. In the second column, because observations are already aggregated, standard errors are simply heteroskedasticity-robust.

Table A7: Testing for parent migration response to deworming

	KLPS Waves 1 and 2		KLPS Wave 1 only	
1998 Class:	*Std2 or Std3*	*Std2*	*Std2 or Std3*	*Std2*
	[1]	[2]	[3]	[4]
Group 2	-0.006	-0.001	-0.02	-0.006
	(0.014)	(0.018)	(0.019)	(0.028)
Group 3	0.0004	-0.003	0.005	-0.005
	(0.014)	(0.018)	(0.02)	(0.027)
Constant	0.064***	0.056***	0.068***	0.065***
	(0.01)	(0.013)	(0.014)	(0.019)
Joint F p-value	0.887	0.985	0.392	0.974
Group 2 = Group 3	0.662	0.923	0.202	0.979
Observations	1871	924	924	454
R^2	0.0001	0.00003	0.002	0.0001

The table above presents tests of a migration response to school-based deworming. The outcome variable is constructed from the first round of the Kenya Life Panel Survey, 2003-2005, and is an indicator for whether the respondent lived outside Busia at the time of the interview. The first two columns aggregate the two rounds of KLPS surveying; the third and fourth column restrict attention to the first round. The odd-numbered columns look at those who were enrolled in either Standard 2 or Standard 3 in 1998; the even-numbered columns look only at those who were in Standard 2 in 1998. Schools were grouped for deworming timing; this table shows coefficients on the indicators for being in a Group 2 or Group 3 school in 1998; Group 1, which received deworming first, is the excluded group. The "Joint F p-value" row indicates the p-value from the test that the two coefficients shown are different from zero. The "Group 2 = Group 3" row indicates the p-value from the test that the coefficients on Group 2 and Group 3 are equal. Heteroskedasticity-robust standard errors are in parentheses.

Table A8: Mean KCPE differences by deworming group

KCPE Year	1993 [1]	1994 [2]	1995 [3]	1996 [4]	1997 [5]	Average [6]
Group 2	0.02	0.012	0.017	0.022	-0.01	0.011
	(0.014)	(0.016)	(0.019)	(0.019)	(0.014)	(0.013)
Group 3	0.007	-0.025	0.016	0.011	-0.006	-0.0006
	(0.016)	(0.018)	(0.018)	(0.017)	(0.014)	(0.014)
Constant	0.466***	0.485***	0.464***	0.478***	0.479***	0.475***
	(0.01)	(0.01)	(0.012)	(0.012)	(0.009)	(0.009)
Joint F p-value	0.347	0.153	0.57	0.519	0.768	0.666
Group 2 = Group 3	0.409	0.056	0.935	0.571	0.833	0.436
Observations	71	71	70	73	72	67
R^2	0.025	0.063	0.016	0.02	0.007	0.012

The table above presents tests of differential KCPE averages across schools. The KCPE is the Kenya Certificate of Primary Education, the primary school leaving examination taken at the end of Standard 8 (8th grade). The data available to me are missing some years' KCPE averages at some schools, so the number of observations varies from column to column. The outcome variable in each column is the school-level mean KCPE score (as a percent out of 700) in one of the years prior to the deworming program, except in Column 6, where the outcome is the average of the five KCPE means used in the first five columns. Schools were grouped for deworming timing; this table shows coefficients on the indicators for being in a Group 2 or Group 3 school in 1998; Group 1, which received deworming first, is the excluded group; mean KCPE percentage for Group 1 is given in the "Constant" row. The "Joint F p-value" row indicates the p-value from the test that the two coefficients shown are different from zero. The "Group 2 = Group 3" row indicates the p-value from the test that the coefficients on Group 2 and Group 3 are equal. Heteroskedasticity-robust standard errors are in parentheses.

Table A9: Main effects: robustness to five years of KCPE controls

Outcome	**Effect**
Raven's Matrices	0.194**
	(0.083)
PPVT Level	0.135*
	(0.078)
Verbal fluency	0.191**
	(0.093)
Memory: digit span forwards	0.114
	(0.103)
Memory: digit span backwards	0.098
	(0.080)
All cognitive: First principal component	0.210**
	(0.095)
All cognitive: Normalized sum	0.210**
	(0.096)
Height (cm)	−0.064
	(0.321)
Height-for-age z-score	−0.012
	(0.047)
Stunting (HAZ<-2)	0.006
	(0.017)

In the table above, the excluded group comprises the cohorts whose communities received school-based deworming during their second year of life or later. Each coefficient comes from a separate regression of the specified outcome on indicators for the age at deworming. Standard errors are clustered at the school-cohort level; gender×age×data collection year fixed effects are included. All cognitive outcomes are standardized (variance=1). Only non-migrants are included in this analysis. The only difference between this table and Table 2 is the inclusion of five separate years of school-level KCPE means as control variables. Because I have incomplete KCPE data, however, this means omitting a number of schools for which I do not have complete KCPE data. The number of observations used for the regression in the first row, for example, is 2,422 in Table 2, but is only 2,171 here.

A.3 Other additional tables

Table A10: Main anthropometric effects: sample with cognitive data

Outcome	**Effect**
Height (cm)	0.320
	(0.745)
Height-for-age z-score	0.049
	(0.108)
Stunting (HAZ<-2)	-0.003
	(0.036)

The table above is analogous to the bottom panel of Table 2, but restricting the analysis to the subsample for whom cognitive data were collected. Standard errors are clustered at the school-cohort level; gender×age×data collection year fixed effects are included. Only non-migrants are included in this analysis.

Table A11: Duplicate of main effects (Table 2) for reference

Outcome	**Effect**
Raven's Matrices	0.211***
	(0.079)
PPVT Level	0.169*
	(0.097)
Verbal fluency	0.200**
	(0.091)
Memory: digit span forwards	0.128
	(0.096)
Memory: digit span backwards	0.022
	(0.089)
All cognitive: PC1	0.215**
	(0.099)
All cognitive: Normalized sum	0.215**
	(0.098)
Height (cm)	0.210
	(0.298)
Height-for-age z-score	0.030
	(0.044)
Stunting (HAZ<-2)	0.001
	(0.015)

Table A12: Main effects, but restricting sample to 1998 and 1999 birth cohorts only

Outcome	**Effect**
Raven's Matrices	0.279***
	(0.094)
PPVT Level	0.190
	(0.123)
Verbal fluency	0.232**
	(0.108)
Memory: digit span forwards	0.100
	(0.120)
Memory: digit span backwards	0.098
	(0.102)
All cognitive: PC1	0.255**
	(0.123)
All cognitive: Normalized sum	0.257**
	(0.122)
Height (cm)	0.342
	(0.346)
Height-for-age z-score	0.051
	(0.051)
Stunting (HAZ<-2)	−0.010
	(0.018)

Table A13: Placebo/falsification test of cognitive effects: spillovers on migrants

Outcome	**Effect**
Raven's Matrices	0.087
	(0.208)
PPVT Level	−0.068
	(0.266)
Verbal fluency	0.023
	(0.269)
Memory: digit span forwards	−0.414
	(0.400)
Memory: digit span backwards	−0.428
	(0.276)
All cognitive: First principal component	−0.281
	(0.303)
All cognitive: Normalized sum	−0.310
	(0.307)

In the table above, the excluded group comprises the cohorts whose communities experienced school-based deworming during their second year of life or later. Each coefficient comes from a separate regression of the specified outcome on indicators for the age at deworming. Standard errors are clustered at the school-cohort level; gender×age×data collection year fixed effects are included. All cognitive outcomes are standardized (variance=1). For the purposes of this "placebo" or "falsification" test, only migrants are included in this analysis. Depending on when their families moved to the areas in which I find them, the early childhood health environment of these respondents is less likely to have been driven by the timing of deworming in their current home area. This contrasts with Table 2 and much of the rest of the paper, in which analysis is restricted to non-migrants. Two weaknesses of this test are that there are relatively few migrants for whom data were collected, and that even for those whose families migrated in, because we do not know the precise date of arrival, we cannot be sure they migrated in late enough for the deworming timing not to influence their early childhood health environment, only that it is less likely to have been an influence. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A14: Subpopulations analysis: variations on Table 3

Panel A: Duplicate of Table 3 for reference

Subpopulation: / Outcome:	[1] Full sample	[2] With older siblings[a]	[3] Without older siblings[a]	[4] Female siblings[b]	[5] Male siblings[b]	[6] Female[c]	[7] Male[c]
Raven's Matrices	0.211***	0.440***	0.238**	0.842***	0.103	0.220*	0.198
	(0.079)	(0.162)	(0.117)	(0.268)	(0.195)	(0.114)	(0.126)
All cognitive: First PC	0.215**	0.408**	0.186	0.771***	0.261	0.249**	0.180
	(0.099)	(0.158)	(0.134)	(0.254)	(0.237)	(0.121)	(0.136)
All cognitive: Normalized sum	0.215**	0.395**	0.190	0.752***	0.272	0.246**	0.182
	(0.098)	(0.158)	(0.136)	(0.255)	(0.235)	(0.122)	(0.137)
Observations	2365	533	894	235	225	1113	1252

Panel B: Restricting sample to 1998 and 1999 birth cohorts only

Subpopulation: / Outcome:	Full sample	With older siblings[a]	Without older siblings[a]	Female siblings[b]	Male siblings[b]	Female[c]	Male[c]
Raven's Matrices	0.279***	0.766***	0.305**	1.341***	0.055	0.345**	0.217
	(0.094)	(0.186)	(0.132)	(0.231)	(0.210)	(0.136)	(0.152)
All cognitive: First PC	0.255**	0.567***	0.210	0.924***	0.254	0.277*	0.231
	(0.123)	(0.186)	(0.152)	(0.283)	(0.291)	(0.144)	(0.171)
All cognitive: Normalized sum	0.257**	0.577***	0.213	0.941***	0.280	0.274*	0.238
	(0.122)	(0.182)	(0.153)	(0.280)	(0.286)	(0.144)	(0.171)
Observations	685	138	268	65	54	311	374

Tables A12 and A14 show that the analysis is robust to restricting the sample to only the 1998 and 1999 birth cohorts, where within-cohort experimental variation includes both individuals who were not yet one year old, and who were more than one year old, when school-based deworming began in their communities.

A.4 Appendices on details of surveying

A.4.1 Precise wording of survey questions

The "child survey" was administered to children aged 8-14 in 2009, and aged 9-15 in 2010. It was administed at school, meaning that many household characteristics could not be directly observed.[40] Key questions:

1.	What is your date of birth? (or age)?	*Ulizaliwa mwaka gani? Mwezi/tarehe gani? (umri / miaka mingapi?)*
2.	Have you ever lived in a different place from the one you live in now?	*Umewahi kuishi mahali pengine mbali na mahali unapoishi sasa?*
3.	Do you have any siblings who are currently living and share the same mother and father as you?	*Una dada na ndugu / kaka wowote ambao umezaliwa nao kwa baba na mama mmoja walio hai?*
3.B.	Sex	*Jinsia*
3.C.	Older/younger/twin?	*Mkubwa/mdogo/pacha?*
3.D.	Still in school?	*Bado anasoma?*
3.E.	Did the sibling ever attend this primary school?	*Aliwahi hudhuria shule ya msingi hii?*

Note that these questions do not require respondents to reliably provide siblings' ages, but do require respondents to report whether siblings are older or younger.

A.4.2 Sampling respondents for cognitive module

In both 2009 and 2010, a random sample of respondents was chosen for cognitive tests, since these modules took roughly ten times as long as anthropometrics. In both years, random uniform draws were made centrally using Stata 10.1 to sample ID numbers for cognitive tests. The draws were made multiple times in each year in order to prevent predictability in the relationship between ID numbers and sampling for cognitive testing across schools. The procedure differed slightly in 2009 and 2010. In 2009, roughly one in twenty

[40] Because of the brevity and low risk of this work, and because the cognitive testing closely resembled normal school activity, the protocol entailed giving teachers copies of a "parent information sheet" to send home with children before the team's visits to schools to inform parents of the team's planned activities, and to provide them with the opportunity to opt out. This protocol, and the details of the sheet, were approved by KEMRI and UC Berkeley CPHS.

respondents was sampled for cognitive data. In 2010, roughly one in four respondents was sampled for cognitive data, conditional on being a nonmigrant. In both years, ID numbers were serially assigned to respondents at the time of completing the child survey; the key questions in that survey are described in Appendix A.4.1. Following that survey by a first field team member, children were measured for height and weight by a second group from the field team; at this second stage, if the ID tag given to the child by the first team matched a list of those numbers randomly sampled for cognitive testing that day, the second group referred the child to a third enumerator for completion of the cognitive tests. Thus because the enrollment team was obliged to enroll participants and assign them ID numbers serially, and did not know, as they gave ID numbers to children, either which random ID number sampling list was used that day or, within it, what numbers would be chosen for cognitive testing (because this list of randomly selected numbers was not in that team's possession), the procedure achieved what some call "allocation concealment" vis-à-vis this sampling process.

A.4.3 Details on number of respondents per "arm"

Evans and Popova (2015) have pointed out that the inclusion of a few additional details in research papers, perhaps particularly those with complex designs, can ease not only the interpretation of the paper, but also the task of determining whether and how to include those papers in meta-analyses. With that in mind, below, I report a few key numbers with respect to sample sizes.

Data collection year:	2009			2010		
Schools visited:	37			36		
PSDP Group	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
Schools visited:	9	9	19	16	14	6
N Interviewed:	3092	1965	4594	4600	5003	2055
N Non-migrant:	2394	1393	3256	3117	3612	1386
N Non-migrant, cognitive:	97	62	138	826	872	376

Cell sizes for main results, Table 2				
Outcome	$A_{id} < 1$ ("treated")	$A_{id} \approx 1$	$A_{id} > 1$ ("comparison")	Total
Raven's Matrices	1171	367	884	2422
PPVT Level	1172	366	882	2420
Verbal fluency	1172	367	884	2423
Memory: digit span forwards	1169	363	873	2405
Memory: digit span backwards	1150	356	864	2370
All cognitive: First principal component	1149	356	860	2365
All cognitive: Normalized sum	1149	356	860	2365
Height (cm)	6230	2204	6719	15153
Height-for-age z-score	6230	2204	6719	15153
Stunting (HAZ<-2)	6230	2204	6719	15153