Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Worms at work: Long-run impacts of a child health investment

IIEP-WP-2015-16

Sarah Baird
George Washington University

Joan Hamory Hicks
University of California, Berkeley

Michael Kremer
Harvard University
NBER

Edward Miguel
University of California, Berkeley
NBER

November 2015

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

Worms at work: Long-run impacts of a child health investment

Sarah Baird, George Washington University
Joan Hamory Hicks, University of California, Berkeley
Michael Kremer, Harvard University and NBER
Edward Miguel, University of California, Berkeley and NBER*

November 2015

Abstract: This study estimates long-run impacts of a child health investment, exploiting community-wide experimental variation in school-based deworming. The program increased labor supply among men and education among women, with accompanying shifts in labor market specialization. Ten years after deworming treatment, men who were eligible as boys stay enrolled for more years of primary school, work 17% more hours each week, spend more time in entrepreneurship, are more likely to hold manufacturing jobs, and miss one fewer meal per week. Women who were eligible as girls are 25% more likely to have attended secondary school, halving the gender gap. They reallocate time from traditional agriculture into cash crops and entrepreneurship. We estimate an annualized financial internal rate of return to deworming of 32%, and show that mass deworming may generate more in future government revenue than it costs in subsidies.

* We thank Kevin Audi, Pierre Bachas, Chris Blattman, Seth Blumberg, Hana Brown, Lorenzo Casaburi, Lisa Ch Garret Christensen, Evan DeFilippis, Lauren Falcao, Francois Gerard, Eva Arceo Gomez, Felipe Gonzalez, Jonas Hjort, Gerald Ipapa, Maryam Janani, Anne Karing, Jen Kwok, Andrew Fischer Lees, Leah Luben, Jamie McCasland, Owen Ozier, Kristianna Post, Adina Rom, Martin Rotemberg, Jon Schellenberg, Changcheng Song, Sebastian Stumpner, Paula Vinchery, Michael Walker, Paul Wang, Zhaoning Wang and Ethan Yeh for providing excellent research assistance on the KLPS project. We thank Michael Anderson, Kathleen Beegle, Jere Behrman, David Card, Alain de Janvry, Erica Field, Fred Finan, Paul Glewwe, Michael Greenstone, Jim Heckman, Adriana Lleras-Muney, Steve Luby, Isaac Mbiti, Mark Rosenzweig, T. Paul Schultz, Jim Smith, John Strauss, Glen Weyl, Alix Zwane and seminar participants at U.C. Berkeley, USC, Harvard, the J-PAL Africa Conference, the Pacific Conference on Development Economics, UCSF, the Gates Foundation WASH Convening in Berkeley, Yale, University of Oklahoma, Hamilton College, RAND, CGD, the World Bank, Maseno University, the NBER Labor Studies group, BREAD/CEPR Meeting in Paris, American University, University of Chicago, Columbia University, Stanford GSB, Makerere University, the AEA meetings (in San Diego), Notre Dame, University of Washington, Mathematica, the Institute for Fiscal Studies, Hong Kong University of Science and Technology, and the International Health Economics Association conference for helpful suggestions. We gratefully acknowledge our collaborators (International Child Support and Innovations for Poverty Action), and funding from NIH grants R01-TW05612 and R01-HD044475, NSF grants SES-0418110 and SES-0962614, the World Bank, the Social Science Research Council, and the Berkeley Population Center. EM was partially funded by the Oxfam Faculty Chair. All errors remain our own. Corresponding author: Edward Miguel, emiguel@berkeley.edu

1. Introduction

The question of whether – and how much – child health gains affect adult outcomes is of major research interest across disciplines and great public policy importance. The belief that childhood health investments may improve adult living standards currently underlies many school health and nutrition programs in low-income countries.

Existing research suggests several channels through which increasing child health investments could affect long-run earnings. Grossman's (1972) seminal health human capital model interprets health care as an investment that increases future endowments of healthy time. Bleakley (2010) further develops this theory, arguing that how the additional time is allocated will depend on how health improvements affect relative productivity in education and in labor. Pitt, Rosenzweig, and Hassan (2012) – hereafter PRH – further note that time allocation will also depend on how the labor market values increased human capital and improved raw labor capacity, and that this in turn may vary with gender. They present a model in which exogenous health gains in low-income economies tend to reinforce men's comparative advantage in occupations requiring raw labor, while leading women to obtain more education and move into more skill-intensive occupations, and provide evidence consistent with this model.

We examine the case of intestinal worms, which globally affect approximately two billion people according to the World Health Organization (2014). Worms (helminths) are spread when infected individuals deposit fecal matter containing eggs in the local environment. Intense infections lead to lethargy, anemia, and growth stunting (Silva et al. 2003; Stephenson et al.

1993; Guyatt et al. 2001; Stoltzfus et al. 1997) and may also weaken the immunological response to other infections (Kirwan et al. 2010; Kjetland et al. 2006). Chronic parasitic infections in childhood may lead to inflammation and elevated cortisol that produce adverse health consequences later in life (Crimmins and Finch 2005), as well as increased maternal morbidity, low birth weight, and miscarriage (Hotez 2009; Larocque et al. 2006).

There is ongoing debate about whether or not it is appropriate to carry out mass deworming treatment programs in endemic regions. Because treatment is safe and very cheap, but diagnosis is expensive, the WHO recommends periodic mass school-based deworming in high-prevalence areas (World Health Organization 1992). Several other bodies also highlight deworming as a cost-effective investment (Hall and Horton 2008; Disease Control Priorities Project 2008; Givewell 2013; Jameel Poverty Action Lab 2012). In contrast, a recent highly publicized Cochrane review[1] argues that while treatment of those known to be infected may be warranted, there is "quite substantial" evidence that mass deworming program does not improve average nutrition, health, or school performance outcomes (Taylor-Robinson et al. 2015). The Cochrane authors have gone even farther in some of their public statements, calling the idea that deworming might have positive economic benefits "delusional" (Boseley 2015).

Yet, because of its selection criteria focusing on medical-style randomized control trials (RCT's), the Cochrane review includes numerous studies subject to now well-known methodological limitations (Bundy et al. 2009), and excludes rigorous social science evidence.

[1] The Cochrane Reviews are systematic reviews of primary research in human health care and health policy that follow the research norms common in medical trials, and are influential among health policymakers.

For instance, the review excludes Bleakley (2007), which estimates the community-wide impact of deworming in the early 20th century U.S. South using quasi-experimental difference-in-difference methods. That study finds that mass deworming improved literacy and raised long-run adult income by 17%; extrapolating to the higher infection rates in tropical Africa, Bleakley (2010) estimates deworming could boost income there by 24%.[2]

The present paper exploits community-wide experimental variation in a deworming program for children in Kenyan primary schools, combined with a longitudinal data set tracking these children into adulthood, to causally identify the effect of improved child health on later life outcomes. At the time of treatment, program participants had already passed the age window considered most critical for early childhood development, suggesting that the time endowment and time allocation effects emphasized in Bleakley (2010), Grossman (1972) and PRH (2012) may be the most relevant channels of impact. Indeed a survey conducted 1-2 years after treatment found no cognitive gains. However, consistent with (Grossman 1972), treatment led to large gains in school participation, reducing absenteeism by one quarter (Miguel and Kremer 2004). There was also evidence for epidemiological externalities within this primary school-age population: untreated children in treatment schools as well as children living near treatment schools had lower worm infection rates and higher school participation (Miguel and Kremer 2004, 2014), and children less than one year old (who were not eligible for treatment) in treated

[2] A small body of social science research studies the impact of deworming on labor outcomes. In additional to Bleakley (2007, 2010), early work by Schapiro (1919) using a first-difference research design found wage gains of 15-27% on Costa Rican plantations after deworming, while Weisbrod et al. (1973) observe little contemporaneous correlation in the cross-section between worm infections and labor productivity in St. Lucia. We discuss the related literature estimating deworming impacts on educational outcomes below.

communities showed cognitive gains in later tests (Ozier 2014).[3] In the current analysis, we examine health, education, and labor market outcomes a decade later, at which point most subjects were young adults 19 to 26 years of age.

Consistent with PRH, we find important gender distinctions in long-term deworming impacts. Men who were in treatment schools as boys work 3.5 more hours each week (on a base of 20.3 hours), spend more time in entrepreneurship, and are more likely to hold manufacturing jobs with higher wage earnings. Their living standards improve as well, with males in treatment schools eating one more meal per week on average. Women who were in treatment schools spend more time in school as girls, and are more likely to have passed the secondary school entrance exam and to have attended secondary school. They reallocate time from traditional agriculture to entrepreneurship and are also more likely to grow cash crops.

In line with Miguel and Kremer (2004), we also find evidence of positive epidemiological externalities on long-run outcomes across a range of outcomes using a seemingly unrelated regression framework. We report point estimates using the linear approach to estimating externalities used in that paper, but also develop a procedure for bounding the impacts of deworming valid under the more general monotonicity assumption that the direct and epidemiological externality effects on labor market outcomes have the same sign.

---

[3] Miguel and Kremer (2014) and Aiken et al. (2015) discuss coding errors in the original estimation of cross-school externalities in Miguel and Kremer (2004). Miguel and Kremer (2014) and Hicks, Kremer, and Miguel (2015) show that once these errors are corrected, positive cross-school infection and school participation externalities extend out to 3 km or 4 km, rather than to the 6 km in Miguel and Kremer (2004). Clemens and Sandefur (2015) carry out an independent analysis that confirms the robustness of the cross-school externality effects out to 4 km. Note that the critiques raised in Aiken et al. (2015) and Davey et al. (2015) do not apply to the data used in the current study.

Lastly, the estimated impacts of deworming on labor market outcomes, combined with other data, allow us to estimate fiscal impacts. We find that the additional net government revenues generated by increased work hours caused by deworming subsidies may be greater than the direct subsidy cost, suggesting that in the case of deworming, health human capital subsidies are potentially Pareto-improving. At a minimum, this suggests that the expected costs to taxpayers are less than would be suggested by multiplying program costs by 1.2 or 1.4 or some other standard multiplier for the deadweight loss of taxation. We also estimate an annualized financial internal rate of return to deworming subsidies of at least 32%, an extremely high return.

The rest of the paper is organized as follows. Section 2 discusses the Kenyan context, the deworming project, and the data. Section 3 presents the estimation strategy. Section 4 discusses the main results. Section 5 combines the results on the price responsiveness of take-up and long-run impacts to assess the fiscal impacts of deworming subsidies, and computes the internal rate of return. The final section concludes.

2. Data

This section describes the study area, the deworming program, and the survey, including our respondent tracking approach and sample summary statistics.

2.1 Study Area and Local Labor Markets

The primary study area is Busia district, a densely-settled farming region in western Kenya adjacent to Lake Victoria that is somewhat poorer than the national average. Outside labor market opportunities for children are meager, and boys and girls both typically attend primary school, with dropout rates rising in grades 7 and 8 (the final two years of primary school). Primary school completion, when children in the study area are typically between 15 to 18 years of age, is a key time of labor market transition. Secondary education in Kenya, like tertiary education in the U.S., depends on exam performance, requires a substantial financial outlay, and often involves moving away from home. In our data, just over half of control group males and just under one third of females continue to secondary school. Occupational and family roles differ markedly by gender, with certain occupations, such as fishing, driving bicycle taxis, and manufacturing, overwhelmingly male, and others, such as small-scale market trading and domestic service, largely female. The model in PRH (2012) suggests that labor market opportunities will affect gender-specific educational and labor responses to health investments.

2.2 The Primary School Deworming Project (PSDP)

In 1998 the non-governmental organization (NGO) International Child Support (ICS) launched the Primary School Deworming Program (PSDP) in two divisions of the district, in 75 primary schools with a total of 32,565 pupils. Parasitological surveys indicated that baseline helminth infection rates were over 90% in these areas. Using modified WHO infection thresholds, over one third of the sample had moderate-heavy infections with at least one helminth (Miguel et al. 2014), a high but not atypical rate in African settings (Brooker et al. 2000; Pullan et al. 2011).

The schools were experimentally divided into three groups (Groups 1, 2, and 3) of 25 schools each: the schools were first stratified by administrative sub-unit (zone), zones were listed alphabetically within each geographic division, and schools were then listed in order of pupil enrollment within each zone, with every third school assigned to a given program group. Appendix section A contains a detailed description of the experimental design, provides further information on the sample (appendix Figure S1), and shows that the three groups were well-balanced along baseline characteristics (appendix Table S1).

Due to the NGO's administrative and financial constraints, the schools were phased into deworming treatment during 1998-2001: Group 1 schools began receiving free deworming and health education in 1998, Group 2 schools in 1999, and Group 3 in 2001. Children in Group 1 and 2 schools were thus assigned 2.41 more years of deworming than Group 3 children on average (appendix Table S2), and these early beneficiaries are the treatment group in the analysis. Take-up rates were approximately 75% in the treatment group and 5% in the control group (Miguel and Kremer 2004). In 2001, the NGO required cost-sharing contributions from parents in a randomly selected half of the Group 1 and Group 2 schools, substantially reducing take-up, and in 2002-2003 it provided free deworming in all schools (Kremer and Miguel 2007).

2.3 Kenya Life Panel Survey (KLPS) Data

The Kenya Life Panel Survey Round 2 (KLPS-2) was collected during 2007-2009, and tracked a representative sample of approximately 7,500 respondents who were enrolled in grades 2-7 in the PSDP schools at baseline. Survey enumerators traveled throughout Kenya and Uganda to interview those who had moved out of local areas. The effective survey tracking rate in KLPS-2 is 82.7% (appendix Table S2), and 84% among those still alive (see appendix sections A and C for further details on survey methodology, tracking rates, and attrition). The effective tracking rate is calculated as a fraction of those found, or not found but searched for during intensive tracking, with weights adjusted appropriately, in a manner analogous to the approach in the U.S. Moving To Opportunity study (Kling, Liebman, and Katz 2007; Orr et al. 2003).

These are high tracking rates for any age group over a decade, and especially for a mobile group of adolescents and young adults. Tracking rates are nearly identical and not significantly different in the treatment and control groups (appendix Table S2).

3. Estimation Strategy

In this section, we define the quantities of interest, describe how to bound them in the presence of potential epidemiological externalities, and then present our econometric strategy.

3.1 Bounding Deworming Treatment Effects in the Presence of Externalities

We need to account for the possibility of externalities in empirically estimating the impact of deworming subsidies. Recall that deworming subsidies were assigned at the school level rather than the individual level. It is therefore worth distinguishing within-school and cross-school externalities. In the potential presence of within-school epidemiological externalities, we cannot separately identify the labor market impact of individual deworming status and of deworming status of others within the school. We can, however, identify the aggregate school-level labor market effect of the deworming subsidy. We, therefore, classify all individuals in schools with a deworming subsidy as “treated” in the empirical analysis.

The remaining issue is cross-school epidemiological externalities. In the remainder of this subsection, we first show that under the relatively weak assumption that the sign of cross-school epidemiological effects on labor market outcomes is not opposite to the sign of direct effects, the difference in outcomes between treatment and control communities is a lower bound on the true total impact of a mass deworming program. For expositional clarity, and to parallel Miguel and Kremer (2004), we start with a discussion of externality effects after one period but generalize them below to longer timeframes. We consider a simple epidemiological model in which worm infection can spread only δ kilometers in a single year, for instance, due to the natural movement of and interaction among the local population. Miguel and Kremer (2004), Miguel and Kremer (2014), and Hicks, Kremer and Miguel (2015) estimate substantial and statistically significant cross-school externalities on worm infections within 3 or 4 km of treatment schools after one year, with less precisely estimated (and not significant) effects from 3 to 6 km.

Consider an outcome Y_{ijt} for individual i in school j at time t, e.g., a labor market outcome. Y_{ijt} is a function of lagged school-level deworming subsidy treatment assignment, $T_{j,t-1} \in \{0,1\}$, and the proportion of other individuals in communities within δ kilometers of that school also received deworming, $P_{j,t-1,\delta} \in [0,1]$. This proportion captures the local “saturation” of the program. This local treatment rate is a function of both the program’s “coverage”, $R_{j,t-1,\delta}$

—i.e., the fraction of pupils in nearby schools assigned to the deworming subsidy treatment, as determined by the research design—and the deworming take-up rate, which is a function of the deworming subsidy level, $Q(S)$. Local treatment saturation is the product of coverage and take-up, $P_{j,t-1,\delta} = R_{j,t-1,\delta}Q(S) + (1 - R_{j,t-1,\delta})Q(0)$, where take-up in the zero subsidy control group is $Q(0)$. Kremer and Miguel (2007) found empirically that control group take-up was very close to zero, implying that $P_{j,t-1,\delta} = R_{j,t-1,\delta}Q(S)$ is a reasonable approximation.[4] For now, we focus on saturation, which is the epidemiologically relevant quantity, but return to the distinction between saturation and coverage in the empirical implementation below.

The first quantity of interest, $\pi_t(1)$, is the expected overall impact of a mass deworming program, namely, the difference in expected outcomes between individuals in treated communities fully exposed to other treatment communities ($P_{j,t-1,\delta} = 1$) versus individuals in untreated communities surrounded by untreated communities:

$$\pi_t(1) \equiv E\left[Y_{ijt} \middle| T_{j,t-1} = 1, P_{j,t-1,\delta} = 1\right] - E\left[Y_{ijt} \middle| T_{j,t-1} = 0, P_{j,t-1,\delta} = 0\right] \quad (1)$$

The second quantity of interest, $\pi_t(p)$, is the impact of a program, such as the one we study, in which the share of nearby population receiving deworming is $P_{j,t-1,\delta} = p, p \in (0,1)$. For each quantity of interest we may also be interested in scaling impact by cost, i.e., $\pi_t(1)/$ (Cost of $P_{j,t-1,\delta} = 1$) and $\pi_t(p)/$ (Cost of $P_{j,t-1,\delta} = p$).

[4] To the extent there was some take-up in control schools, estimates are a lower bound on the impact of deworming.

Define the expected outcome in untreated communities surrounded by other untreated communities (i.e., "pure control" communities uncontaminated by exposure to nearby treatment schools) as $y_{0,t} \equiv E\left[Y_{ijt} | T_{j,t-1} = 0, P_{j,t-1,\delta} = 0\right]$ and define the difference in expected outcomes between treated and untreated communities at a given local treatment saturation proportion p as:

$$\begin{aligned}\lambda_{1t}(p) \equiv E\left[Y_{ijt} |\, T_{j,t-1} = 1, P_{j,t-1,\delta} = p\right] \\ - E\left[Y_{ijt} |\, T_{j,t-1} = 0, P_{j,t-1,\delta} = p\right]\end{aligned} \tag{2}$$

Define the difference in average outcomes between untreated communities at a local treatment proportion p versus pure control communities as:

$$\lambda_{2t}(p) \equiv E\left[Y_{ijt} |\, T_{j,t-1} = 0, P_{j,t-1,\delta} = p\right] - y_{0,t} \tag{3}$$

The sum of these two effects is $\pi_t(p) \equiv \lambda_{1t}(p) + \lambda_{2t}(p)$.

The biological mechanism underlying the spread of worm infections implies that worm load in a particular location at time t is non-decreasing in worm load in that location and neighboring areas within distance $\tilde{\delta}$ at lagged time $t - \tilde{t}$. Both own and neighbors' treatment at time $t - \tilde{t}$ should thus reduce own worm load at t. This is captured in our first assumption (where to make the notion of monotonicity concrete, the first inequality establishes that the direct effect of treatment on Y is positive, without loss of generality):

Assumption 1 (Monotonic externality effects): Suppose for all p, $E\left[Y_{ijt} \middle| T_{j,t-1} = 1, P_{j,t-\tilde{t},\tilde{\delta}} = p\right] \geq E\left[Y_{ijt} \middle| T_{j,t-1} = 0, P_{j,t-\tilde{t},\tilde{\delta}} = p\right]$, then for any two levels of local treatment saturation $p'' > p'$, $E\left[Y_{ijt}\middle|T_{j,t-1} = \mu, P_{j,t-\tilde{t},\tilde{\delta}} = p''\right] \geq E\left[Y_{ijt}\middle|T_{j,t-1} = \mu, P_{j,t-\tilde{t},\tilde{\delta}} = p'\right]$ for all $\mu \in \{0,1\}$.

In a setting with real-world saturation level p, analysis that does not account for cross-community spillover effects focuses on estimating $\lambda_{1t}(p)$. Assumption 1 implies that $\lambda_{1t}(p)$ is a lower bound on both quantities of interest, $\pi_t(1)$ and $\pi_t(p)$.

Proposition 1 (Bounding the treatment effect): Suppose for all p, $E\left[Y_{ijt}\middle|T_{j,t-1} = 1, P_{j,t-1,\delta} = p\right] \geq E\left[Y_{ijt}\middle|T_{j,t-1} = 0, P_{j,t-1,\delta} = p\right]$, then $\pi_t(1) \geq \pi_t(p) \geq \lambda_{1t}(p)$ for all $p \in (0,1)$.

Proof: We proceed in two steps. We first show that $\pi_t(p'') \geq \pi_t(p')$ for all $p'' > p'$. Note that $\pi_t(p'') - \pi_t(p') = \left(E\left[Y_{ijt} \mid T_{j,t-1} = 1, P_{j,t-1,\delta} = p''\right] - y_{0,t}\right) - \left(E\left[Y_{ijt} \mid T_{j,t-1} = 1, P_{j,t-1,\delta} = p'\right] - y_{0,t}\right) = E\left[Y_{ijt} \mid T_{j,t-1} = 1, P_{j,t-1,\delta} = p''\right] - E\left[Y_{ijt} \mid T_{j,t-1} = 1, P_{j,t-1,\delta} = p'\right]$. This is greater than or equal to zero by the monotonicity assumption, implying that $\pi_t(1) \geq \pi_t(p)$ for all $p < 1$. We next show that $\pi_t(p) \equiv \lambda_{1t}(p) + \lambda_{2t}(p) \geq \lambda_{1t}(p)$. For all $p > 0$, Assumption 1 implies that $\lambda_{2t}(p) \equiv E\left[Y_{ijt} \mid T_{j,t-1} = 0, P_{j,t-1,\delta} = p\right] - E\left[Y_{ijt} \mid T_{j,t-1} = 0, P_{j,t-1,\delta} = 0\right] \geq 0$. The result follows. □

It is possible to tie this result more closely to the empirical analysis by taking into account the fact that local saturation rates actually differ across communities. Allow $P_{j,t-1,\delta}$ to be distributed across communities as $P_{j,t-1,\delta} \sim F$, with density f. Then in practice the average difference in outcomes across treated and untreated communities is:

$$\int_{P=0}^{P=1} \lambda_{1t}(P) f(P) dP.$$

Since the result in Proposition 1 holds for all $p \in (0,1)$, it holds for this above expression, which is effectively a weighted average across different saturation proportions p in this set.

The above discussion abstracts away from other covariates. As we discuss below, their inclusion in a regression analysis is important given the nature of the experimental design and stratified sampling, and also potentially improves statistical precision. One covariate that we include in the empirical analysis is the local density of all primary school pupils (in all schools, treatment and control). We show in Table S2 of the appendix and in Miguel and Kremer (2004) that the local numbers of all primary school pupils and of treatment school pupils are unrelated to treatment school assignment, although there is a statistically significant but small difference in the treatment saturation proportion; the fact that this proportion is slightly lower in treatment schools implies that the treatment school versus control school difference is, once again, likely to be a lower bound on true impacts. Drug take-up rates in treatment schools are also not significantly correlated with the local density of either treatment schools or of all schools (Miguel and Kremer 2004, Appendix Table A.II). Taken together, these patterns imply that the coefficient estimate on the treatment school indicator is unlikely to be biased, or that any potential bias would again lead us to understate deworming impacts.

Note that the bound above will still be valid, albeit looser, if the geographic spread of epidemiological externalities over time means that even "pure control" (i.e., $T = 0$ and $P = 0$) schools are subject to some spillover from the program. Those whose infection intensity falls due to cross-school spillovers could themselves generate positive spillovers for other nearby schools, which would then lead to less local re-infection with worms, and so on.

Denote worm prevalence at location j at time t by ω_{jt}. Given the geographic spread of worm infections by δ kilometers per year, ω_{jt} will be a non-decreasing function of worm prevalence at time $t - \tilde{t}$ at all locations within radius $\delta\tilde{t}$. Thus given the results in Miguel and Kremer (2004), worm infection prevalence after the decade-long gap between treatment and the follow-up survey in our study will potentially be reduced by worm treatment within a distance of at least 30 km (=10 years x 3 km per year) and perhaps beyond. And while, of course, these effects may fade over time, no school in our study area of roughly 15 km by 40 km can be considered a "pure control" in the presence of these externalities.

It is straightforward to generalize the bounding result above to the empirically relevant case of an extended follow-up period. Denote the time period of the original deworming program as $t = 0$, and subsequent years take on values of $t = 1, 2, 3, \dots t^*$, where t^* is the period of the follow-up survey. While in the short-run (as in Miguel and Kremer 2004) the cross-school local treatment saturation measure due to the deworming program ($P_{j,0,\delta}$) is likely to fairly accurately capture the magnitude of the externality impacts, over time the infection "feedback" effects generated in all directions among nearby schools would lead us to understate the magnitude of the true cross-school externalities. Determining the magnitude of all these externality effects is beyond the scope of this paper, as the spatial and temporary variation in our data do not allow us to precisely estimate the wide range of potentially relevant parameters, but in Appendix B we prove that the bounding result still holds in this case.

As noted, Miguel and Kremer (2004) report cross-school externalities up to 3 km from the school, and at 3-6 km. There was a statistical program coding error in the construction of the cross-school externality term in Miguel and Kremer (2004) limiting the analysis to the 12 closest schools. Correcting the coding error does not substantively alter the estimated effects of externalities between 0-3 or 0-4 km, since there were never more than 12 schools within 4 km, but does lead to less precisely estimated overall effects between 3-6 km from a school; Miguel and Kremer (2014) and Hicks, Kremer, and Miguel (2015) contain a complete discussion of the updated empirical results. We consider cross-school externalities up to 6 km in the analysis in this paper for two reasons. First, we do so since spillover effects are likely to diffuse spatially over time, as discussed above. Second, we consider externality effects out to 6 km because an F-test in a seemingly unrelated regression (SUR) framework rejects the hypothesis that the externality effects are zero in the 3-6 km range for the outcomes we consider (P-value < 0.001), indicating that their inclusion is appropriate (see appendix B2 for details). The main results are largely unchanged using alternative specifications for the cross-school externality effect, including dropping these terms from the analysis entirely, as we discuss below.

3.2 Estimation

The econometric approach relies on the PSDP's prospective experimental design, namely, that the program exogenously provided individuals in treatment (Group 1 and 2) schools two to three additional years of deworming. We focus on intention-to-treat estimates, since compliance rates are high, and previous research showed that untreated individuals within treatment communities experienced gains (Miguel and Kremer 2004), complicating estimation of treatment effects on the treated within schools. Since PRH suggest potentially different labor market effects of health investments on males and females in low-income "brawn-based economies", occupations are sharply differentiated by gender in our data, and roughly twice as many women in our sample have children compared to the men, we follow the tradition in the labor market literature of examining prime-age women and men separately (Altonji and Blank 1999; Bertrand 2011).

The dependent variable is outcome Y_{ij}, for individual i in school j, in the KLPS-2 survey:

$$Y_{ij} = \alpha + \lambda_1 T_j + \lambda_2 P_j + X'_{ij,0}\beta + \varepsilon_{ij} \tag{4}$$

The outcome is a function of the assigned deworming program treatment status of the individual's primary school (T_j); the treatment saturation proportion among neighboring schools within 6 km during the original treatment phase of the PSDP (P_j)[5]; a vector $X_{ij,0}$ of baseline individual and school controls; and a disturbance term ε_{ij}, which is clustered at the school level.

[5] One issue with employing local saturation rates as an explanatory variable in practice is that they are a function of the local treatment decisions of households in the relevant local area, leading to possible endogeneity concerns, for instance, if take-up is higher in areas where people have unobservably better labor market prospects. To address these concerns we construct the local saturation measure P_j as a function of the local coverage rate R_j of treatment school pupils within 6 km of school j, which is exogenously determined by the experimental design, times the average take-up rate of deworming drugs in the entire sample at the full subsidy level. This implies that variation in the local saturation variable is driven entirely by the experimental design, with the average take-up rate serving as a useful "rescaling" to allow for a more meaningful interpretation of the magnitude of estimated effects.

The $X_{ij,0}$ controls include school geographic and demographic characteristics used in the PSDP "list randomization", the student gender and grade characteristics used for stratification in drawing the KLPS sample (Bruhn and McKenzie 2009), a pre-program average school test score to capture academic quality, the 2001 cost-sharing school indicator (described below), the total number of primary school pupils within 6 km of the school, and survey month and wave controls. Estimates are weighted to make the results representative of the full PSDP sample originally in grades 2-7, taking into account the sampling for KLPS and the tracking strategy.

The main coefficient of interest is λ_1, which captures gains accruing to individuals in treatment schools relative to the control; since deworming was assigned by school rather than at the individual level, some of the gains in treatment schools are likely due to within-school externalities. This is an attractive coefficient to focus on since it is a lower bound on the overall effect of deworming (Proposition 1). Another coefficient of some interest is λ_2, which captures the spillover effects for nearby schools, following Miguel and Kremer (2004). As explained further in that paper, since reinfection rates are very high in the area, the magnitude of externality effects may be either larger or smaller than the effect of own-school treatment. We have analyzed other specifications, including interactions between treatment and local saturation, and non-linearities in saturation (appendix B), but cannot reject that T_j and P_j are additively separable and enter in linearly.

The direct treatment effect estimates and externality effects are locally relevant to the infection rates and treatment saturation rates in the setting we study, and while we do not find evidence of interaction effects or non-linear externalities, it remains possible that such effects would emerge at treatment levels outside the support of values that we observe. One case of potential interest is one in which treatment coverage rates are even higher than those observed in our setting, for instance, if all local schools were assigned to treatment (rather than approximately two-thirds, as in our case). In this case, it is possible to place bounds on the cost-effectiveness of deworming using our data under the highly conservative assumption that there are no additional benefits from boosting deworming treatment saturation, i.e., in the notation above that $\pi(p) = \pi(p')$ and $\lambda_2(p) = \lambda_2(p')$ for all $p' > p$.

For concreteness, consider the case in which all estimates are based on local treatment saturation rates in the neighborhood of $p < 1$ and program coverage $R < 1$. Due to externalities, program benefits are experienced both in the schools assigned to treatment and the control schools, and can be represented as $R\pi(p) + (1 - R)\lambda_2(p) = R\lambda_1(p) + \lambda_2(p)$. Then under an assumption of constant marginal per capita treatment costs (which again is likely to be conservative given the fixed costs of setting up a treatment program), the cost of expanding local program coverage to all schools in the area ($R = 1$) is $1/R$ times the cost of covering proportion R of the population. In our case, this is implemented by multiplying the baseline costs of deworming treatment by $1/(2/3) = 1.5$, while the total benefits are assumed to remain unchanged. We present bounds using this approach in section 5 below.[6]

---

[6] Of course, if $\pi(p) = \pi(p')$ and $\lambda_2(p) = \lambda_2(p')$ for all $p' > p$, policymakers have the option of replicating a program like that implemented in this study, in which case the relevant cost-effectiveness calculations would be based on the costs and benefits at coverage and saturation levels found in our data.

4. Results

After briefly discussing long-run health effects, we present impacts on education, labor outcomes and living standards, by gender. Results are broadly consistent with the PRH model.

4. 1 Long-run health impacts

While treatment dramatically reduced moderate-heavy infections in the short-run (Table 1, row 1), adult helminth lifespans are typically between one and four years (Hotez et al. 2006), so the direct effects of treatment will no longer be present a decade later in the data used in this analysis. Any long-run effects would instead be due to effects on other diseases through an immunological channel, or to the effects of changes in schooling or labor outcomes.

Although we find no long-term effects on physical growth or body mass index, there is some evidence of persistent health gains in terms of self-reported health and reduced miscarriage. Respondent reports that their health was "very good" rose by 4.0 percentage points (SE 1.8, $P < 0.05$), on a base of 67.3% in the control group. We cannot reject equal effects for both genders, but gains are slightly larger for women. Furthermore, deworming reduced miscarriage rates among treatment group females by 2.8 percentage points (SE 1.3, $P < 0.05$) on a base of 3.9 percent in a probit analysis (where each pregnancy is the unit of observation). The lack of miscarriage impact among the partners of men in the treatment group suggests a health, rather than a living standards, channel for the impacts estimated among sample women.

4.2 Education impacts

The medium-run follow up (Miguel and Kremer 2004) found increased primary school participation among both boys and girls, consistent with the idea that health investment increased the endowment of healthy time (Grossman 1972), and that for children, this increased time went into schooling rather than working. The long-run follow up data show that treatment continued to boost boys' primary school enrollment, but that average academic performance did not improve, with high rates of grade repetition and no significant differences between the treatment and control groups in rates of passing the secondary school exam or enrolling in secondary school (Table 2). We do not have data on whether increased primary-school enrollment improved non-cognitive skills, a possible channel for labor market impacts (Heckman, Stixrud, and Urzua 2006). Recall that in the models in Bleakley (2010) and Pitt, Rosenzweig, and Hassan (2012), deworming would not increase secondary schooling if attractive work opportunities emerged around the time of primary school completion (roughly ages 15 to 18) and if health investments raised the marginal return to work as much as the discounted return to secondary schooling.

In contrast, our primary specification suggests that deworming leads to marked academic gains for girls, increasing the rate at which girls passed the secondary school entrance exam by 9.6 percentage points ($P < 0.05$) on a base of 41%. This increase of roughly 25% reduces the existing gender gap in exam performance by half. Consistent with the model in PRH (2012), in which positive health shocks disproportionately induce women to allocate more time to human capital acquisition, treatment also halved the gender gap in secondary school entry, increasing girls' secondary enrollment by 0.325 years, or over a third (appendix Table S3), and increasing overall years of school enrollment for women by 0.354 years (SE 0.179, P < 0.05) (Table 2).

4.3 Impact on labor hours and occupation

Average weekly hours worked in the control group are quite low, at 20.3 for men and 16.3 for women (although many women in our sample are engaged in home production or child-rearing activities, and time spent on these activities was not systematically collected in KLPS-2). Among men, deworming increased time spent working by 17%, or 3.49 hours per week (SE 1.42, P < 0.05, Table 3, Panel A). In contrast, estimated effects on non-household work hours among women are small. It is worth noting that three quarters of both the treatment and control groups are no longer in school by the time of the survey (Table 2). In this subpopulation, treated women worked 2.14 more hours per week, and although we cannot reject the hypothesis of no effect for women, we also cannot reject the hypothesis of equal treatment effects by gender.

Deworming changes how work hours are allocated across sectors and occupations, with important distinctions by gender (Table 3, Panel B). Taking the genders together, hours in non-agricultural self-employment increase by 45% (P<0.01), and results are shown by gender in Figure 1 (Panels A and B). There are no statistically significant changes in hours worked in agriculture or wage employment.

Breaking results down by gender, point estimates suggest that deworming leads men to increase total work hours (Table 3, Panel B), and we cannot reject the hypothesis of equal percentage increases across sectors. In contrast, women increase time in non-agricultural self-employment ("entrepreneurship") by 1.86 hours (SE 0.81, $P < 0.05$) on a base of 2.7 hours, nearly 70%, and reduce hours worked in agriculture by 1.27 hours (SE 0.56, $P < 0.05$). This shift from agricultural work into entrepreneurship could potentially be interpreted as consistent with PRH, although the evidence is not dispositive. 77% of self-employed women work in retail, which seems less physically-intensive than agriculture, and there is evidence that retail profits are tied to math skills (Kremer et al. 2013). However, there is no significant difference in education levels between women in agricultural employment and those in entrepreneurship.

Deworming treatment also leads to shifts in occupational choice (Table 3, Panel C). Treatment respondents are three times more likely to work in manufacturing (coefficient 0.0110, $P < 0.05$) from a low base of 0.005. On the flip side, casual labor – which typically does not require regular work hours – falls significantly ($P < 0.05$). Manufacturing jobs require more hours per week than other occupations: they average 53 hours per week, compared to 42 hours for all wage earning jobs, 34 hours for self-employment and 11 hours for agriculture. Workers in manufacturing tend to miss relatively few work days due to poor health, at just 1.1 days in the last month (in the control group), compared to 1.5 days among all wage earners. Manufacturing jobs are highly paid, with average earnings almost double those in casual labor (Table S17). Deworming also leads to an increase in cash crop cultivation for the entire sample (Table 3, Panel C), with a gain of 1.04 percentage points ($P < 0.05$) on a low base of 0.45 percent.

Estimates of occupational effects by gender are less precise, but there are significant increases in manufacturing among men and in growing cash crops among women. The particularly large effect of deworming on physically-demanding and well-paid manufacturing employment among men is consistent with the PRH model. There is suggestive evidence of a shift into high work hour occupations for men but not women (see appendix C).

The increase in secondary education, entrepreneurship, and cash crop cultivation among women may reflect a desire to engage in higher productivity activities within existing family and social constraints, which may complicate moves into manufacturing or other lucrative male-dominated jobs. More speculatively, these may pay off in the form of higher future earnings, even if not yet apparent in our data.

4.4 Impact on living standards

Living standards can be assessed using data on either consumption or earnings. We do not have data on overall consumption, but do have data on the number of meals consumed. Treatment respondents eat 0.095 more meals per day (SE 0.029, $P < 0.01$, Table 4, Panel A). The increase in meals eaten is larger for men, at 0.125 meals/day ($P < 0.01$) than for women (0.051 meals), implying that treatment males miss just under one fewer meal each week than control males.

Total earnings are the sum of earnings in wage labor, in non-agricultural self-employment, and in agriculture, each weighted by the proportions working in each sector. In principle, these proportions could differ by treatment group, but there are no significant differences by treatment status (see appendix Table S5, odd numbered columns). We consider each source of income below.

Those working in wage employment likely have the best measured data. The distribution of log wage earnings is shifted to right for both men (Figure 1, Panel C) and women (Panel D) in the treatment group relative to control. Log earnings (Table 4, Panel B) are 26.9 log points (SE 8.5, $P < 0.01$) greater. The estimated differences in earnings are larger than those of hours, consistent with the hypotheses that treatment leads men to shift into jobs that require more work hours and that pay better. Log wages computed as earnings per hour worked (among those who work at least 10 hours per week) are 19.7 log points (SE 10.2, $P < 0.10$) greater in the treatment group. Wage earnings differences between treatment and control are also positive among the larger number of respondents who had ever earned wages since 2007, with an average difference of 22.5 log points ($P < 0.01$) during the most recent earnings period.

The data on self-employment profits are likely measured with more noise. In the full sample, monthly profits are 22% larger in the treatment group, but the difference is not significant (Table 4, Panel C), in part due to large standard errors created by a few male outliers reporting extremely high profits. In a version of the profit data that trims the top 5% of observations, the difference is 28% ($P < 0.10$).

With no changes in the proportion of respondents in different sectors, and estimated increases in earnings of more than 20% among wage earners – and similar (if less precisely estimated) profit increases among the self-employed – treatment will have increased overall earnings unless agricultural earnings declined. Unfortunately, we lack sufficient data on agricultural earnings to perform a direct test. However, several patterns suggest that it is unlikely agricultural earnings declined, and highly unlikely that they declined sufficiently to outweigh the gains in other sectors. Recall that cash crop cultivation increased, and that hours worked in agriculture did not change. Most importantly, if agricultural productivity had declined, one might expect that food consumption among those working in agriculture would decline, but there is in fact an increase of 0.065 meals (SE 0.033) in this group. There is no evidence that the quality of agricultural labor fell in the treatment group (appendix C).

In general, while weighted earnings by sector can always be summed to generate total earnings, the treatment versus control differences within particular sectors reflect a combination of treatment and selection effects. There are significantly different patterns of selection into wage employment and non-agricultural self-employment by treatment status (Table S5). However, the similar rates at which treatment and control individuals work as wage laborers and the similar selection patterns along observable dimensions (Tables S5, S14-S15) suggest that the wage differences might plausibly be interpreted as primarily due to treatment effects.

4.5 Heterogeneous Treatment Effects and Alternative Specifications

While statistical power is limited, we do not find strong evidence of heterogeneous treatment effects on education, labor market or living standards outcomes by baseline school grade, local treatment saturation, or the presence of schistosomiasis (as proxied for by distance to Lake Victoria, see appendix section C.4 and Tables S6-S13).

Estimated deworming impacts are largely robust to whether or not we account for the cross-school spillovers at all, and to accounting for cross-school externalities at different distances (appendix Tables S6-S9, column 5). Appendix Figure S5 shows that effects typically remain statistically significant across alternative specifications of the externality effects for key outcome measures (although for the "passed primary exam" outcome for females, P-values range from 0.02 to 0.26). The externality results are similar if we focus on the number of local pupils, rather than the proportion, in treatment schools (appendix Tables S6-S9, column 2).

4.6 Accounting for multiple inference

To further assess robustness, we next account for multiple inference, and then examine two additional sources of variation in exposure to deworming.

Appendix Tables S18-S21 present the false discovery rate adjusted q-values (analogues to the standard P-value) that limit the expected proportion of rejections within a set of hypotheses that are Type I errors (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008). Key results are robust to this adjustment: taking both genders together, the deworming impact on meals eaten and labor earnings is statistically significant at the 1% level (q-value < 0.01), on total hours worked in non-agricultural self-employment, trimmed self-employed profits, and manufacturing employment is significant at the 5% level, and the reduction in casual labor jobs and the increase in cash crops are significant at the 10% level. There is less power with the gender subsamples but most key results continue to hold at the 10% level (appendix section C.5).

4.7 Variation in cost-sharing

Because the temporary 2001 deworming treatment cost-sharing program substantially reduced take-up, it provides an additional, orthogonal source of variation in treatment, albeit with less statistical power. Reassuringly, the estimated effect of cost-sharing has the opposite sign of the main deworming treatment effect for 24 of the 28 outcomes presented in Tables 1-4 (excluding the first outcome in Table 1, which was measured before cost-sharing was introduced), and this pattern seems extremely unlikely to occur by chance. In addition, stacking the data and using seemingly unrelated regression (SUR) estimation across outcomes, we reject the hypothesis that the cost-sharing coefficients are zero (P<0.001); see appendix section C.

4.8 Cross-school treatment externalities

Cross-school externalities provide a third source of exogenous variation in exposure to deworming. Several of the externality effect estimates in Tables 1-4 are significant and large in magnitude, including for miscarriage, manufacturing employment, and meals eaten (P < 0.05). Under the null hypothesis of no epidemiological externalities, there should be no correlation with the direct treatment effect. In 24 of the 29 specifications in Tables 1-4, the sign of the treatment effect and the cross-school externality effect are the same, which is extremely unlikely to occur by chance; an alternative test estimates a correlation of 0.655 between the t-statistics for the direct effect and the externality effect across outcomes (P-value < 0.002); and using SUR, we reject the hypothesis that the 0-6 km cross-school externality effects are zero (P<0.001); see appendix B. The existence of cross-school externalities provides additional evidence on the robustness of the deworming impacts, and some reassurance that estimated effects are not simply due to some form of reporting bias in the treatment schools.

5. The Rate of Return and Fiscal Impacts of Deworming Subsidies

The estimated impacts of deworming on labor market outcomes, combined with other data, allow us to estimate the internal rate of return and fiscal impacts of deworming subsidies.

We observe only a snapshot of labor market outcomes at the time of the follow-up survey, rather than the whole path of future hours and earnings, and thus the calculations in this section are by necessity somewhat speculative. We adopt what we consider to be a reasonably conservative approach in bounding the effect of lifetime income. In particular, we base our calculations on differences in hours worked between the treatment and control groups. This is likely to be conservative for a number of reasons: 1) estimated differences in earnings among wage workers are larger than differences in hours (Table 4, Panel B); 2) among women, treatment is associated with greater educational attainment and higher test scores, and it seems plausible that this could lead to higher future earnings; and 3) there is increased entrepreneurial activity, particularly among women, and it seems plausible that some of this consists of investments which could pay off in increased earnings later.

For projections about the future path of earnings and thus government revenues, we examine the following expression:

$$S_2 Q(S_2) - S_1 Q(S_1) < \sum_{\gamma} N_{\gamma} \left[\tau(S_1) \sum_{t=0}^{t=50} r^t w_t \left(\lambda_{1,\gamma} + \frac{p\lambda_{2,\gamma}}{R} \right) - K \sum_{t=0}^{t=50} r^t \Delta \bar{E}_{\gamma t}(S_1, S_2) \right] \quad (5)$$

The left hand side is the fiscal cost to the government of increasing a deworming subsidy from S_1 to S_2, which in turn may affect deworming take-up Q; take-up is non-decreasing in the subsidy. To compute this, we use information on take-up at different price levels from Kremer and Miguel (2007), and current estimates of per pupil mass deworming treatment costs (provided by the NGO Deworm The World) of \$0.59 per year. The total direct deworming cost then is the 2.41 years of average deworming in the treatment group times this figure, or M = \$1.42 per person treated and \$1.07 per pupil in a deworming treatment school, given average take-up of 75%. Under partial deworming subsidies, as implemented in the 2001 cost-sharing program, individuals paid an average of \$0.27 for the medicines, so the direct cost to the government would be \$1.15 for each fully dewormed individual over 2.41 years. In Table 5, Panel A, we compare these subsidy levels with the default case of no subsidies, $S_1 = 0$.

The right hand side captures the implications for government revenue of increasing the subsidy from S_1 to S_2. N_{γ} is the fraction of individuals in the sample of type $\gamma \in \Gamma$, which we operationalize as gender, following the empirical analysis. The first term in the square brackets captures the increase in tax revenue generated by any increase in work hours: $\tau(S_1)$ is the prevailing tax rate; r is the per period interest rate; w_t is the wage rate in year t; $\lambda_{1,\gamma}$ is the estimated deworming impact on work hours in treatment schools for gender γ; $\lambda_{2,\gamma}$ is the estimated externality effect; and p and R denote the program's saturation and coverage, as above. These gains are captured over an individual's working life, which we take to be 50 years.

The second term in the square brackets accounts for the fact that improved child health may lead the government to accrue additional educational expenditures, for instance, if secondary schooling rates increase for type γ, which we find for females. Let K capture the cost of an additional unit of schooling, and $\Delta\bar{E}_{\gamma t}(S_1, S_2)$ denote the average increase in schooling for type γ when the deworming subsidy increases from S_1 to S_2. To compute the right hand side of eqn. (5), we use a combination of estimates from this paper and other Kenyan data. The hours worked estimates (Table 3) indicate that treatment group males work 3.49 more hours per week ($\lambda_{1,male} = 3.49$), whereas the treatment effect estimate for women is near zero ($\lambda_{1,female} = 0.32$). The point estimate of the increase in work hours due to epidemiological externalities is 10.20 hours/week for an increase in treatment saturation from 0 to 100%, and we combine this information with each school's local density of treated pupils to determine $p\lambda_{2,\gamma}$.[7] Since this externality estimate is not significant at conventional levels, we focus on the case of no epidemiological externality ($\lambda_{2,\gamma} = 0$) in panel B, and present results in Panel C assuming the externality has the estimated magnitude for completeness. We examine the impact of a program that treated two thirds of local schools, as in the PSDP, and scale up externality gains by the inverse of the coverage rate ($1/R$) since the control group also benefits from externalities.

[7] Results are similar when externalities are disaggregated by gender (not shown).

At the time of writing, the Government of Kenya pays 11.85% interest on its sovereign debt and inflation is approximately 2%, so we set the real cost of capital r to 9.85%.[8] We assume that the sample population begins working ten years after they first began receiving deworming and retires after 40 years of work.[9] From year 10 post-treatment onwards, we combine estimated $\lambda_{1,\gamma}$ and $\lambda_{2,\gamma}$ values from Tables 3-4 above with the pattern of lifecycle earnings reported in the most recent publicly available data, the 1998/1999 Kenya Integrated Labour Force Survey, and assume recent Kenyan economic growth trends continue. This forward projection of earnings is necessary given the limitations of existing data, and implies that the calculations that follow are somewhat speculative. We also assume the initial starting wage w is \$0.18 per hour, which is a weighted average of wages by sector in our data and the mean Kenyan agricultural wage in Suri (2011), with weights corresponding to control group mean hours per sector (Table 4).[10] Kenyan taxes (mainly on consumption) absorb roughly 16.5% of GDP so we set the tax rate under no subsidy to 16.5%.[11]

We estimated deworming impacts on school enrollment by gender and year (appendix Table S3), and also gathered detailed information on current teacher salaries and class sizes from the Ministry of Education, allowing us to estimate per capita schooling costs K for both primary

---

[8] See http://www.centralbank.go.ke/securities/bonds/manualresults.aspx and World Bank Development Indicators.
[9] This ten year gap roughly corresponds to the time elapsed from the start of PSDP until the KLPS2 survey (2007-09). By ignoring the time before KLPS2 data was collected, it underestimates gains due to greater work hours prior to the survey. Yet it misses any reduction in work hours due to substitution of school for work. However, existing estimates of child labor productivity suggest these foregone earnings are likely to be small (Udry 1996).
[10] Suri (2011) mean agricultural wage is \$0.16, and the control group mean of \$0.23 (Table 4, Panel B) for those working for wages. Self-employed wages are calculated by dividing control group monthly profits (Table 4, Panel C) by 4.5 times the hours worked per week among those working in self-employment, for a wage of \$0.14.
[11] From World Development Indicators, government expenditures are roughly 19.5% of GDP, and from http://blogs.worldbank.org/africacan/three-myths-about-aid-to-kenya about 15% of government expenditure is financed from donors, thus 0.195*0.85=0.165.

and secondary schooling. Because the PSDP program did not increase the number of teachers or classrooms in primary schools, and there is no reason to believe the Kenyan government adjusted these factors in response to the program (based on our observations as well as on discussions with local officials), any costs of increased classroom congestion at the primary level due to deworming would have been incurred by students in these schools and thus is already captured in the labor market outcomes in our data. We therefore focus on measuring the fiscal costs to the government of increased secondary school enrollment, since these costs would be incurred either by the government (by paying for additional teachers) or by secondary school students. Teacher salaries constitute the bulk of recurrent government education spending, at over 90% of secondary school spending (Otieno and Colclough 2009), and most other expenses are traditionally covered by tuition and local parent fees. We factor in the costs that the government would need to incur in order to maintain the secondary school pupil-teacher ratio, using our estimated per student secondary school teacher cost of \$116.85 per year (Table 5, Panel A).

Assuming no externality gains, $\sum_{\gamma}\sum_{t=0}^{t=50} r^t w_t \lambda_{1,\gamma}$ = \$142.43, implying that individuals gain an average of \$119 in take-home pay and the NPV of government revenue increases by \$23 per person (Table 5, Panel B). The additional public educational costs incurred are estimated to be approximately \$10.71, so the net increase in government revenue is \$12.90, far greater than the \$1.07 subsidy. If deworming also generates positive externalities, the earnings gains are much larger, with a per capita net increase in government revenue of \$102.97 (Panel C).

A policy relevant case is one in which the coverage (R) of the population assigned to deworming increased from the roughly two thirds in our study sample up to all local primary schools, as in a national mass treatment program. In that case, the relative cost-effectiveness of the program could depend on the degree to which total program treatment effects depend on local treatment saturation, i.e., on the shapes of both $\pi(p)$ and $\lambda_2(p)$, something we cannot directly estimate (the 10-90 range for saturation rate P_j in our data is 0.427 to 0.599). However, we can bound the cost-effectiveness of a program that covered the entire population under the conservative assumption that there are no additional net benefits from boosting the treatment rate. The cost per treatment school student (under full subsidies) would rise by 50% from $1.07 to $1.60 while the NPV net increase in government revenue would remain unchanged at $12.90, implying that a program treating all schools would also be highly cost effective.

In terms of other extensions, our model assumes a linear income/consumption tax but the result is robust to a range of alternative assumptions on taxation, including the possibility of a lower tax rate in our predominantly rural sample; see appendix section C for further discussion.

A standard approach to assessing the desirability of a program is to calculate the social internal rate of return (IRR), which solves for the interest rate that equates the NPV of the full social cost and all earning gains, whether taxed or untaxed: in the above notation, $MQ(S) = \sum_{\gamma}\left[\sum_{t=0}^{t=50} r^t w_t \left(\lambda_{1,\gamma} + \frac{p\lambda_{2,\gamma}}{R}\right)\right]$. The annualized social IRR with no health spillovers ($\lambda_{2,\gamma} = 0$) is very high at 32.2%, and with health spillovers is a massive 51.6%.

These fiscal and IRR calculations are speculative for several reasons, including the projection of future earnings, as noted above. This exercise also ignores broader general equilibrium effects of a mass national deworming program on wage levels and the capital stock; these macroeconomic effects could theoretically either increase or decrease the effects we present in this section, although they seem unlikely to overturn the main patterns (appendix C contains a discussion). They are also relatively imprecisely estimated: we bootstrapped standard errors (with 1000 runs), and find that net revenue gains are less than zero 24% of the time for the case of no health spillovers. So while estimates indicate that the expected net revenue effects of deworming are large, there remains considerable uncertainty around these estimates.

Yet these calculations are also conservative in several dimensions. For one, note that even in cases where the net revenue effects are not positive, the gains in the labor market due to deworming help partially offset the original expenditure outlay on deworming subsidies, substantially reducing their net fiscal cost. The fiscal and internal rate of return exercises above also only rely on income and ignore any welfare gains through other channels. It is plausible that those who had better health and nutrition as a result of deworming benefited from an increased endowment of healthy hours, and experienced direct utility gains from simply feeling better, and the same could be said for the inherent utility benefits of increased schooling. Finally, we do not incorporate recent evidence that positive deworming externalities extend beyond those in our sample to other age groups: Ozier (2014) finds that living in a deworming treatment community early in life (age 0 to 2) leads to improved cognitive and academic performance ten years later.

Older individuals in the area also plausibly benefited from the health spillovers of treatment but we lack data to quantify any such gains.

6. Conclusion

Previous work (Miguel and Kremer 2004) found that a primary-school deworming program increased school participation. This paper shows that some education and labor market outcomes improve one decade after receiving deworming. These gains could have important positive welfare impacts for households living near subsistence, like many in our Kenyan sample. We estimate that the annualized financial internal rate of return (IRR) to deworming is extremely high at 32.2%. Our best estimate is that deworming subsidies will generate more in future government revenue than they cost in up-front expenditures.[12]

The high rate of return to deworming in our Kenyan context is consistent with findings in the 20th century U.S. South (Bleakley 2007, 2010), and recent evidence on positive long-run educational impacts in East Africa in Ozier (2014) and Croke (2014). Of course, there is uncertainty around our estimates and returns could differ in other environments, but even given some uncertainty, or substantial weight on priors that the returns to deworming are smaller, this growing body of evidence suggests that the expected financial rate of return would likely exceed conventional hurdles for public health investment (Ahuja et al. 2015).

[12] Some have argued that certain other public health investments could also have this property, including tobacco cessation (Lightwood and Glantz 2013) and reduced drunk driving (Ditsuwan et al. 2013).

The results also have implications for several related literatures. Many studies argue that early childhood health gains *in utero* or before age three have the largest impacts (Almond and Currie 2010) and some have argued that interventions outside a narrow window of child development will not have major effects. Our evidence suggests that health interventions among school-aged children, which are too late in life to affect cognition or height, can have long-run impacts on labor outcomes by affecting the amount of time people spend in school or work.

While there is a literature on differences in work hours across wealthy countries (Prescott 2004), the determinants of labor hours in poor countries are less studied. Work hours are quite low in some low-income settings (Fafchamps 1993), including among our control group. The findings here suggest that poor child health may be one factor behind this low adult labor supply.

Finally, our analysis does not account for potential negative externalities from deworming through drug resistance. Geerts and Gryseels (2000, 2001) highlight mass deworming policy approaches that could minimize the development of resistance, and while there is limited current evidence on drug resistance related to human deworming, it has been documented in livestock (Albonico, Engels, and Savioli 2004). Despite their concerns, Geerts and Gryseels (2001) do still conclude that community-based mass deworming treatment makes sense in high morbidity settings, such as our Kenyan study area, and we agree it is unlikely that resistance would be large enough to overturn the case for subsidies. Worm prevalence is likely to decline over time with economic development, as more people have sanitation facilities, wear shoes, and take other actions to avoid infection, and it is therefore unlikely to be optimal to hold back on treating the sick today in order to "save" the drug for later. Moreover, if there is a need to cut back on drug administration to reduce the risk that resistance will develop, cutting back on veterinary use in high-income countries is likely to be a more appropriate initial response.

References

Ahuja, Amrita, Sarah Baird, Joan Hamory Hicks, Michael Kremer, Edward Miguel, and Shawn Powers. 2015. "When should governments subsidize health? The case of mass deworming." *World Bank Economic Review* 29 (supplement):S9-S24.

Aiken, Alexander M., Calum Davey, James R. Hargreaves, and Richard J. Hayes. 2015. "Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication." *International Journal of Epidemiology*.

Albonico, M., D. Engels, and L. Savioli. 2004. "Monitoring drug efficacy and early detection of drug resistance in human soil-transmitted nematodes: a pressing public health agenda for helminth control." *International Journal for Parasitology* 34:1205-1210.

Almond, Douglas, and Janet Currie. 2010. "Human Capital Development before Age Five." *NBER Working Paper #15827*.

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Race and gender in the labor market." In *Handbook of Labor Economics*, edited by O. Ashenfelter & D. Card, 3143-3259.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedaian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484):1481-1495.

Benjamini, Y., A. Krieger, and D. Yekutieli. 2006. "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3):491-507.

Bertrand, Marianne. 2011. "New Perspectives on Gende." In *Handbook of Labor Economics*, edited by O. Ashenfelter & D. Card. Elsevier.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics* 122 (1):73-117.

Bleakley, Hoyt. 2010. "Health, Human Capital, and Development." *Annual Review of Economics* 2: 283-310.

Boseley, Sarah. 2015. "New research debunks merits of global deworming programs." *The Guardian*.

Brooker, Simon, M. Rowlands, L. Haller, L. Savioli, and D.A. Bundy. 2000. "Towards an Atlas of Helminth Infection in Sub-Saharan Africa: The Use of Geographical Information Systems (GIS)." *Parasitology Today* 16 (7):303-307.

Bruhn, M., and D. McKenzie. 2009. "In pursuit of balance: Randomization in practice in development field experiments." *American Economic Journal: Applied Economic* 1 (4).

Bundy, DAP., M. Kremer, H. Bleakley, MCH. Jukes, and E. Miguel. 2009. "Deworming and Development: Asking the Right Questions, Asking the Questions Right." *PLoS Negl Trop Dis* 3 (1):e362. doi:10.1371/journal.pntd.0000362.

Clemens, M., and J. Sandefur. 2015. "Mapping the Worm Wars: What the Public Should Take Away from the Scientific Debate about Mass Deworming." *Center for Global Development Blog*.

Crimmins, Eileen M., and Caleb E. Finch. 2005. "Infection, inflammation, height, and longevity." *Proceedings of the National Academy of Sciences* 103 (2):498-503.

Croke, K. T. 2014. "The Long Run Effects of Early Childhood Deworming on Literacy and Numeracy: Evidence from Uganda." *Unpublished Manuscript*.

Davey, C. , AM. Aiken, RJ. Hayes, and JR. Hargreaves. 2015. "Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a statistical replication of a cluster quasi-randomized stepped-wedge trial." *International Journal of Epidemiology*.

Disease Control Priorities Project. 2008. Deworming Children Brings Huge Health and Development Gains in Low-Income Countries

Ditsuwan, V., J.L. Veerman, M. Bertram, and T. Vos. 2013. "Cost-Effectiveness of Interventions for Reducing Road Traffic Injuries Related to Driving under the Influence of Alcohol." *Value in Health* 16:23-30.

Fafchamps, Marcel. 1993. "Sequential Labor Decisions Under Uncertainty: an Estimable Household Model of West African Farmers." *Econometrica* 61 (5):1173-1197.

Geerts, S., and B. Gryseels. 2000. "Drug Resistance in Human Helminths: Current Situation and Lessons from Livestoc." *Clinical Microbiology Reviews* 13 (2):207-222.

Geerts, S., and B. Gryseels. 2001. "Anthelmintic resistance in human helminths: a review." *Tropical Medicine and International Health* 6 (11):915-921.

Givewell. 2013. Combination deworming (mass drug administration targeting both schistosomiasis and soil-transmitted helminths).

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2):223-255.

Guyatt, H.L., S. Brooker, C.M. Kihamia, A. Hall, and D.A. Bundy. 2001. "Evaluation of efficacy of school-based anthelmintic treatments against anaemia in children in the United Republic of Tanzania." *Bulletin of the World Health Organization* 79 (8):695-703.

Hall, Andrew, and Sue Horton. 2008. Best Practice Paper: New Advice from CCo8: Deworming. Copenhagen: Copenhagen Consensus Center.

Heckman, J., J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics.* 24 (3):411-482.

Hicks, JH., M. Kremer, and E. Miguel. 2015. "Commentary: Deworming externalities and schooling impacts in Kenya: a comment on Aiken et al. (2015) and Davey et al. (2015)." *International Journal of Epidemiology*.

Hotez, P.J. 2009. "Empowering Women and Improving Female Reproductive Health through Control of Neglected Tropical Diseases." *PLoS Neglected Tropical Diseases* 3 (11):e559.

Hotez, Peter J., Donald A. P. Bundy, Kathleen Beegle, Simon Brooker, Lesley Drake, Nilanthi de Silva, Antonio Montresor, Dirk Engels, Matthew Jukes, Lester Chitsulo, Jeffrey Chow, Ramanan Laxminarayan, Catherine Michaud, Jeff Bethony, Rodrigo Correa-Oliveira, Xiao Shuhua, Alan Fenwick, and Lorenzo Savioli. 2006. *Helminth Infections:*

Soil-transmitted Helminth Infections and Schistosomiasis. 2nd Edition ed, *Disease Control Priorities in Developing Countries* Washington, DC: World Bank.
Jameel Poverty Action Lab. 2012. Deworming: A Best Buy for Development. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
Kirwan, P., Andrew L. Jackson, Samuel O. Asaolu, Sile F. Molloy, Titlayo C. Abiona, Marian C. Bruce, Lisa Ranford-Cartwright, Sandra M. O'Neill, and Celia V. Holland. 2010. "Impact of repeated four-monthly anthelmintic treatment on Plasmodium infection in preschool children: a double-blind placebo-controlled randomized trial." *BMC Infectious Diseases* 10 (277).
Kjetland, E.F., P.D. Ndhlovu, Gomo E., T. Mduluza, N. Midzi, L. Gwanzura, P.R. Mason, L. Sandvik, H. Friis, and S.G. Gundersen. 2006. "Association between genital schistosomiasis and HIV in rural Zimbabwean women." *AIDS* 20 (4):593-600.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1):83-119.
Kremer, Michael, J. Lee, J. Robinson, and O. Rostapshova. 2013. "Behavioral Biases and Firm Behavior: Evidence from Kenyan Retail Shops." *American Economic Review Paper and Proceedings* 103:362-368.
Kremer, Michael, and Edward Miguel. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics* 112 (3):1007-1065.
Larocque, R., M. Casapia, E. Gotuzzo, J.D. MacLean, J.C. Soto, Rahme E., and T.W. Gyorkos. 2006. "A double-blind randomized controlled trial of antenatal mebendazole to reduce low birthweight in a hookworm-endemic area of Peru." *Tropical Medical International Health* 11 (10):1485-95.
Lightwood, J., and S.A. Glantz. 2013. "The Effect of the California Tobacco Control Program on Smoking Prevalence, Cigarette Consumption, and Healthcare Costs: 1989-2008." *PLoS One* 8 (2).
Miguel, Edward, and M. Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1):159-217.
Miguel, Edward, and M. Kremer. 2014. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Guide to Replication of Miguel and Kremer (2004)." *Center for Effective Global Action Working Paper Series #39*.
Miguel, Edward, Michael Kremer, Joan Hamory Hicks, and Carolyne Nekesa. 2014. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Data User's Guide." *Center for Effective Global Action Working Paper Series #40*.
Orr, L., J. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, LF. Katz, JB. Liebman, and JR. Kling. 2003. Moving to Opportunity: Interim Impacts Evaluation. Washington, D.C.: U.S. Dept. of HUD.
Otieno, W. , and C. Colclough. 2009. "Financing Education in Kenya: Expenditures, Outcomes and the Role of International Aid." *RECOUP Working Paper 25*.
Ozier, Owen. 2014. "Exploiting Externalities to Estimate the Long-term Benefits of Early Childhood Deworming " *World Bank Working Paper Series #7052*.
Pitt, Mark M., M.R. Rosenzweig, and N. Hassan. 2012. "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy." *American Economic Review* 102 (7):3531–3560.
Prescott, Edward. 2004. "Why do Americans work so much more than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review* 28 (1):2-13.

Pullan, R.L., P.W. Gething, J.L. Smith, C.S. Mwandawiro, Sturrock H.J., Gitonga C.W., Hay S.I., and Brooker S. 2011. "Spatial modelling of soil-transmitted helminth infections in Kenya: a disease control planning tool." *PLoS Neglected Tropical Diseases* 5 (2):e958.

Schapiro, Louis. 1919. "The Physical and Economic Benefits of Treatment for Hookworm Disease." *Journal of the American Medical Association* 73 (20):1507-09.

Silva, N.R. de, S Brooker, P.J. Hotez, A. Montresor, D. Engels, and L. Savioli. 2003. "Soil-transmitted helminth infections: updating the global picture." *Trends in Parasitology* (12):547-51.

Stephenson, L.S., M.C. Latham, E.J. Adams, S.N. Kinoti, and A. Pertet. 1993. "Physical fitness, growth and appetite of Kenyan school boys with hookworm, Trichuris trichiura and Ascaris lumbricoides infections are improved four months after a single dose of albendazole." *The Journal of nutrition* 123 (6):1036-1046.

Stoltzfus, R.J., H.M. Chwaya, Tielsch J.M., Schulze K.J., Albonico M., and Savioli L. 1997. "Epidemiology of iron deficiency anemia in Zanzibari schoolchildren: the importance of hookworms." *American Journal of Clinical Nutrition* 65 (1):153-9.

Suri, Tavneet. 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79 (1):159-209.

Taylor-Robinson, DC., N. Maayan, K. Soares-Weiser, S. Donegan, and P. Garner. 2015. "Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin, and school performance." *The Cochrane Database of Systematic Reviews* 7:1-157.

Weisbrod, B.A., R. L. Andreano, R. E. Balwin, and A. C. Kelley. 1973. *Disease and economic development: The impact of parasitic diseases in St. Lucia*. Madison: Univ. of Wisconsin Press.

World Health Organization. 1992. Model Describing Information. Drugs used in Parasitic Diseases. Geneva: WHO.

World Health Organization. 2014. Soil-transmitted helminth infections. In *WHO Fact sheet N°366*.

Table 1: Deworming impacts on health

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
	All	Male	Female	All	All	Male	Female
Moderate-heavy worm infections in 2001	-0.166***	-0.191***	-0.144***	-0.074	0.327	0.319	0.337
	(0.026)	(0.028)	(0.032)	(0.223)	(0.469)	(0.466)	(0.473)
					2,297	*1,216*	*1,081*
Self-reported health "very good" indicator at KLPS-2	0.040**	0.023	0.051**	0.128	0.673	0.713	0.629
	(0.018)	(0.025)	(0.025)	(0.115)	(0.469)	(0.452)	(0.483)
					5,070	*2,585*	*2,485*
Height at KLPS-2	-0.109	0.072	-0.301	-1.891	167.3	171.7	162.3
	(0.271)	(0.382)	(0.387)	(1.667)	(8.0)	(6.5)	(6.5)
					5,072	*2,585*	*2,487*
Body mass index (BMI) at KLPS-2	0.022	-0.012	0.058	0.317	27.22	26.50	28.03
	(0.045)	(0.060)	(0.066)	(0.269)	(1.31)	(1.02)	(1.11)
					5,072	*2,585*	*2,497*
Miscarriage indicator (obs. at pregnancy level) at KLPS-2	-0.015*	0.000	-0.028**	-0.078**	0.030	0.015	0.039
(for females – themselves; for males – their partners)	(0.008)	(0.004)	(0.013)	(0.037)	(0.171)	(0.123)	(0.194)
					5,022	*1,622*	*3,238*

Notes: The sample includes all individuals surveyed in KLPS-2 (2007-2009), except for the moderate-heavy worm infection data, which is from the 2001 PSDP parasitological survey. Each entry is from a separate OLS regression except the miscarriage outcome, which are marginal probit specifications in which each observation is a pregnancy. All observations are weighted to maintain initial population proportions, except for the 2001 moderate-heavy worm infection results. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence. The coefficient on the deworming treatment indicator term is λ_1 in equation 1. The cross-school externality term is the "saturation rate" – the number of treatment group (Group 1,2) pupils within 6 km divided by the total number of primary school pupils within 6 km, multiplied by the average deworming take-up rate in the sample – demeaned, and the coefficient on the externality term is λ_2 in equation 1. All regressions except for the first include controls for baseline 1998 primary school population, geographic zone of the school, survey wave and month of interview, a female indicator variable, baseline 1998 school grade fixed effects, the average school test score on the 1996 Busia District mock exams, total primary school pupils within 6 km, and the cost-sharing school indicator. The first row includes controls for baseline 1998 primary school population, geographic zone of the school, a female indicator variable, baseline 1998 school grade fixed effects, the average school test score on the 1996 Busia District mock exams, and total primary school pupils within 6 km. Self-reported health "very good" takes on a value of one if the answer to the question "Would you describe your general health as somewhat good, very good, or not good?" is "very good", and zero otherwise.

Table 2: Deworming impacts on education

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
	All	Male	Female	All	All	Male	Female
Total years enrolled in school, 1998-2007	0.294**	0.150	0.354**	1.015	6.69	7.05	6.29
	(0.145)	(0.166)	(0.179)	(0.839)	(2.97)	(2.93)	(2.96)
					5,037	*2,567*	*2,470*
Total years enrolled in primary school, 1998-2007	0.155**	0.238**	0.026	0.784	4.38	4.43	4.32
	(0.075)	(0.102)	(0.098)	(0.485)	(2.48)	(2.42)	(2.55)
					5,037	*2,567*	*2,470*
Repetition of at least one grade (1998-2007) indicator	0.063***	0.072***	0.053*	0.099	0.672	0.669	0.676
	(0.018)	(0.025)	(0.030)	(0.123)	(0.470)	(0.471)	(0.468)
					5,084	*2,595*	*2,489*
Grades of schooling attained by 2007	0.150	-0.030	0.261	0.323	8.72	9.06	8.34
	(0.143)	(0.148)	(0.171)	(0.842)	(2.21)	(2.28)	(2.07)
					5,084	*2,595*	*2,489*
Attended secondary school indicator	0.030	-0.035	0.090**	-0.032	0.421	0.504	0.329
	(0.035)	(0.038)	(0.038)	(0.217)	(0.494)	(0.500)	(0.470)
					5,084	*2,595*	*2,489*
Passed secondary school entrance exam during 1998-2007 indicator	0.050	0.004	0.096**	0.220	0.505	0.590	0.409
	(0.031)	(0.030)	(0.040)	(0.161)	(0.500)	(0.492)	(0.492)
					4,974	*2,541*	*2,433*
Out-of-school (at 2007-09 survey) indicator	-0.006	0.022	-0.029	0.185	0.75	0.70	0.80
	(0.022)	(0.030)	(0.026)	(0.142)	(0.43)	(0.46)	(0.40)
					5,058	*2,582*	*2,476*

Notes: For details on the regressions, see the notes for Table 1. Each entry is from a separate OLS regression.

Table 3: Deworming impacts on labor hours and occupational choice

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Hours worked	All	Male	Female	All	All	Male	Female
Hours worked in all sectors in last week, full sample	1.58	3.49**	0.32	10.20	18.4	20.3	16.3
	(1.04)	(1.42)	(1.36)	(7.80)	(23.1)	(24.6)	(21.1)
					5,084	*2,595*	*2,489*
Hours worked in all sectors in last week, out-of-school sample	2.93**	4.55**	2.14	14.61	22.0	25.9	18.3
	(1.29)	(1.95)	(1.49)	(9.16)	(24.8)	(26.5)	(22.4)
					3,873	*1,869*	*2,004*
Panel B: Sectoral time allocation							
Hours worked in non-agricultural self-employment in last week, full sample	1.51***	1.35*	1.86**	6.00*	3.3	3.8	2.7
	(0.55)	(0.73)	(0.81)	(3.23)	(12.8)	(13.7)	(11.7)
					5,084	*2,595*	*2,489*
Hours worked in agriculture in last week, full sample	-0.07	1.03*	-1.27**	-0.55	8.3	7.8	8.8
	(0.42)	(0.55)	(0.56)	(3.41)	(11.4)	(11.6)	(11.2)
					5,084	*2,595*	*2,489*
Hours worked in wage earning in last week, full sample	0.14	1.11	-0.27	4.74	6.9	8.8	4.8
	(0.84)	(1.32)	(1.08)	(5.07)	(18.5)	(20.0)	(16.5)
					5,084	*2,595*	*2,489*
Panel C: Occupational choice (full sample)							
Manufacturing job indicator	0.0110***	0.0192**	0.0050	0.0531**	0.0049	0.0068	0.0027
	(0.0040)	(0.0077)	(0.0035)	(0.0250)	(0.0698)	(0.0824)	(0.0522)
					5,084	*2,595*	*2,489*
Construction/casual labor job indicator	-0.0053**	-0.0031	-0.0073	-0.0196	0.0048	0.0040	0.0057
	(0.0026)	(0.0030)	(0.0045)	(0.0154)	(0.0691)	(0.0628)	(0.0756)
					5,084	*2,595*	*2,489*
Domestic service job indicator	-0.0050	0.0016	-0.0134	-0.0097	0.0192	0.0067	0.0331
	(0.0061)	(0.0038)	(0.0129)	(0.0322)	(0.1372)	(0.0813)	(0.1791)
					5,084	*2,595*	*2,489*
Grows cash crop indicator	0.0104**	0.0032	0.0187**	-0.0171	0.0045	0.0048	0.0042
	(0.0051)	(0.0044)	(0.0090)	(0.0228)	(0.0671)	(0.0692)	(0.0648)
					5,018	*2,565*	*2,453*

Notes: For details on the regressions, see the notes for Table 1. Each entry is from a separate OLS regression. Agricultural work in Panel B includes both farming and pastoral activities.

Table 4: Deworming impacts on living standards and labor earnings

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Consumption	All	Male	Female	All	All	Male	Female
Number of meals eaten yesterday, full sample	0.095***	0.125***	0.051	0.415***	2.16	2.10	2.23
	(0.029)	(0.041)	(0.043)	(0.124)	(0.64)	(0.65)	(0.62)
					5,083	*2,595*	*2,488*
Number of meals eaten yesterday, out-of-school sample	0.102***	0.158***	0.037	0.542***	2.16	2.08	2.25
	(0.029)	(0.046)	(0.044)	(0.168)	(0.64)	(0.66)	(0.62)
					3,872	*1,869*	*2,003*
Panel B: Wage earnings (among wage earners)							
Ln(Total labor earnings), past month	0.269***	0.244**	0.165	1.141	7.79	7.92	7.46
	(0.085)	(0.109)	(0.175)	(0.869)	(0.88)	(0.87)	(0.81)
					710	*542*	*168*
Ln(Wage = Total labor earnings / hours), past month, if ≥10 hours per week of work	0.197*	0.181	0.225	0.378	2.68	2.88	2.21
	(0.102)	(0.128)	(0.194)	(0.898)	(0.91)	(0.89)	(0.81)
					601	*448*	*153*
Ln(Total labor earnings), most recent month worked since 2007	0.225***	0.221**	0.178*	0.941	7.83	7.97	7.54
	(0.070)	(0.097)	(0.104)	(0.597)	(0.91)	(0.89)	(0.89)
					1,175	*819*	*356*
Panel C: Non-agricultural self-employment outcomes (among non-agricultural self-employed)							
Total self-employed profits (self-reported) past month	384	111	250	-77	1,766	2,135	1,265
	(308)	(465)	(265)	(1,646)	(2,619)	(3,235)	(1,261)
					585	*313*	*272*
Total self-employed profits past month, top 5% trimmed	341*	259	80	440	1,221	1,184	1,265
	(177)	(309)	(219)	(1.256)	(1,151)	(1,056)	(1,261)
					553	*284*	*269*
Total employees hired (excluding self)	0.416	0.245	0.603	-0.886	0.188	0.253	0.097
	(0.361)	(0.403)	(1.275)	(2.547)	(0.624)	(0.614)	(0.630)
					633	*343*	*290*

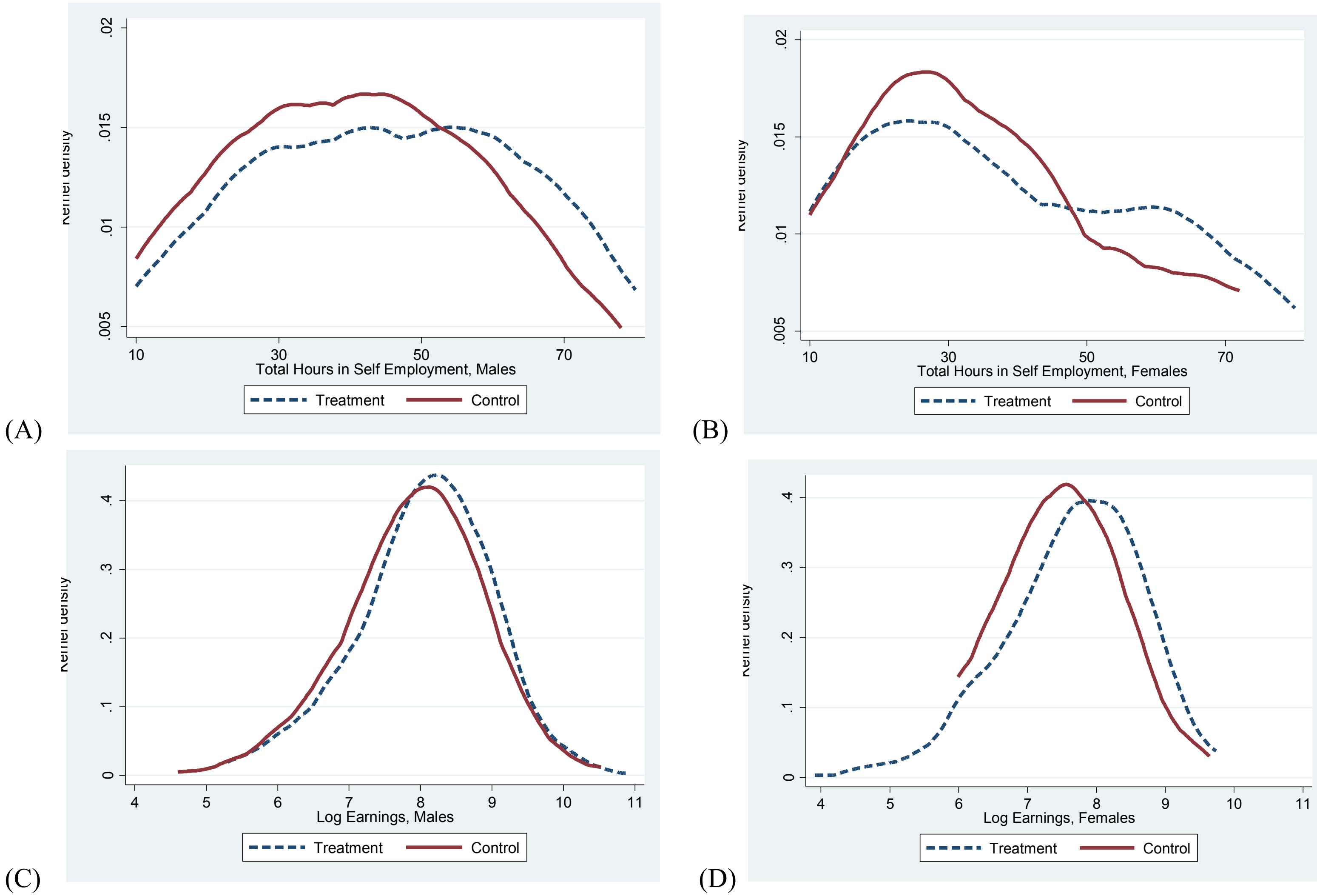
Notes: For details on the regressions, see the notes for Table 1. Each entry is from a separate OLS regression, except for "total employees hired" in Panel C, which utilizes a negative binomial regression. Real earnings measures account for the higher prices found in the urban areas of Nairobi and Mombasa. We collected price surveys in both rural western Kenya and in urban Nairobi during KLPS-2, and base the urban price deflator on these data; results are unchanged without this price adjustment. The wage, earnings and profits results in Panels B and C are among those who reported wage employment or non-agricultural self-employment, respectively. When computing wages, we exclude those with fewer than 10 hours per week to address division bias from noise in estimation of number of hours worked. "Total employees hired" is among those who are self-employed.

Table 5: Fiscal Impacts of Deworming Subsidies

Panel A: Calibration Parameters	No Subsidy	Partial Subsidy	Full Subsidy	*Notes*
Size of Subsidy: S	$0.00	$1.15	$1.42	From Deworm the World; Kremer and Miguel (2007)
Take-up rate: *Q(S)*	5%	19%	75%	From Kremer and Miguel (2007)
Average per-person cost: *SQ(S)*	$0.00	$0.22	$1.07	= Subsidy x take-up rate
Mean per person increase in work hours/week: λ_1	0.00	0.44	1.75	Men: increase of 3.49 hours/week; women: no change (Table 3). Partial subsidy multiplied by Q(S)/Q(full)
Mean increase in work hours/week from externality: $p\lambda_2$	0.00	1.76	5.21	10.20 (Table 3) x Coverage of treatment school students within 6 km (R, 68.1%) x [Q(S) for full subsidy, Q(S)/Q(full) for partial subsidy]
Mean increase in schooling costs	0.00	2.71	10.71	NPV additional secondary schooling costs per pupil-year ($116.85) x direct increase in secondary schooling
Mean increase in schooling costs from externality	0.00	3.40	13.42	NPV additional secondary schooling costs per pupil-year ($116.85) x externality increase in secondary schooling
Panel B: No health spillovers				
Annual increase in per-person earnings	$0.00	$3.91	$15.44	λ_1 x starting wage x 52
NPV increase in per-person earnings (relative to no subsidy)	-	$36.08	$142.43	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	$3.27	$12.90	NPV earnings x 16.5% tax rate – Direct schooling costs
Panel C: With health spillovers				
Annual increase in per-person earnings	$0.00	$26.77	$83.11	(λ_1 + (p/R) λ_2) x starting wage x 52
NPV increase in per-person earnings (relative to no subsidy)	-	$246.99	$766.81	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	$34.84	$102.97	NPV earnings x 16.5% tax rate – (Direct+externality schooling costs)

Notes: The deworming cost is US$0.59 per year, and the average number of years treated was 2.41 years. Figures in Panels B and C are relative to the "no subsidy" case. We use a starting hourly wage rate (w) of $0.18, a weighted average of wages by sector with weights corresponding to control group mean hours per sector (Table 4). We use Suri's (2011) mean wage of $0.16 as the agricultural wage, and the control group mean of $0.23 (Table 4, Panel A) for those working for wages. Self-employed wages are calculated by dividing control group monthly profits (Table 4, Panel B) by 4.5 times the hours worked per week among those working in self-employment, for a wage of $0.14. The public finance data is from the Kenyan Central Bank website and the World Bank Development Indicators. The NPV of per-person lifetime earnings in the no subsidy and no health spillovers case is $1,509.96. We assume that earnings start 10 years after deworming treatment and continue for 40 years. Life cycle earnings profiles for Kenya are created using data from the 1998/1999 Kenya Integrated Labour Force Survey, by regressing individual earnings on age, age squared, and indicator variables for female, attained a schooling level of primary/secondary/beyond, and province of residence. Future earnings are also assumed to increase by the average per-capita GDP growth rate in Kenya during the 2001 to 2011 period, namely 1.52% per annum (World Bank Development Indicators). Calculations are available upon request.

Figure 1: Hours worked in self-employment (if working 10 to 80 hours in sector) and earnings, treatment versus control
Panel A: Hours worked in self-employment in last week, males; Panel B: Hours worked in self-employment in last week, females;
Panel C: Log earnings in wage employment in past month, males; Panel D: Log earnings in wage employment in past month, females.



Supplemental Online Appendix for

"Worms at work: Long-run impacts of a child health investment"

Sarah Baird, Joan Hamory Hicks, Michael Kremer and Edward Miguel*

*Corresponding author. E-mail: emiguel@berkeley.edu

November 2015

This PDF file includes:

Supplementary Text
Figures S1 to S4
Tables S1 to S24
Additional references

Contents

A. Study setting and research design

A.1 Selection of Schools for the Primary School Deworming Project (PSDP) Sample

In January 1998, there were a total of 92 primary schools in the study area of Budalangi and Funyula divisions, across eight geographic zones. Seventy-five of these 92 schools were selected to participate in PSDP. The 17 excluded schools include: town schools that were quite different from other local schools in terms of student socioeconomic background; single-sex schools; a few schools located on islands in Lake Victoria (posing severe transportation difficulties); and those few schools that had in the past already received deworming and other health treatments under an earlier small-scale ICS (NGO) program.

In particular, four primary schools in Funyula Town were excluded due to large perceived income differences between their student populations and those in other local schools. Specifically, three schools charged schools fees well in excess of neighboring primary schools, and thus attracted the local “elite”. Another is a private boarding school for girls, charging even higher fees, and was similarly excluded.

Four other primary schools in Budalangi division were excluded from the sample due to geographic isolation, which introduced logistic difficulties and would have complicated deworming treatment and data collection. Three of these schools are located on islands in Lake Victoria, and the fourth is separated from the rest of Budalangi by a marshy area.

Two additional schools were excluded. One served as the pilot school for the PSDP in late 1997, receiving deworming treatment before other local schools, and thus it was excluded from the evaluation. The other was excluded since it was a newly opened school in 1998 with few pupils in the upper standards (grades), and thus was not comparable to the other sample schools.

Seven schools had participated in the ICS Child Sponsorship Program/School Health Program (CSP/SHP). In 1998, it was felt that identification of treatment effects in these schools could be complicated by the past and ongoing activities in those schools, including health treatment (and deworming in particular), and hence they were excluded from the sample. The NGO’s earlier criteria in selecting these particular seven schools (in 1994-1995) is not clear.

The PSDP sample was roughly a quarter of the total population (across all ages) of Budalangi and Funyula divisions, which was 127,231 (1999 Census). The 1998 Kenya Demographic and Health Survey finds that 85% of 8 to 18 year olds in western Kenya were enrolled in school, indicating that our school-based sample is broadly representative of children in the region.

Drugs for STH (albendazole) were offered twice per year and for schistosomiasis (praziquantel) once per year.

A.2 Prospective Experimental Procedure

Miguel and Kremer (2004) contains a partial description of the prospective experimental list randomization procedure, and we expand on it here; further detail on the study design is presented in Miguel et al. (2014). Schools were first stratified by geographical area (division, then zone), and the zones were listed alphabetically (within each division), and then within each zone the schools were listed in increasing order of student enrolment. There are two divisions, Budalangi and Funyula, containing a total of eight zones (Agenga/Nanguba, Bunyala Central, Bunyala North, Bunyala South, Bwiri, Funyula, Namboboto, and Nambuku). Table S1 shows there is no significant difference between average school populations in the treatment and control groups.

While the original plan had been to stratify by participation in other NGO programs, the actual randomization was not carried out this way. Schools participating in the intensive CSP/SHP program were dropped from the sample (as detailed above), while 27 primary schools with less intensive NGO programs were retained in the sample. These 27 schools were receiving assistance in the form of either free classroom textbooks, grants for school committees, or teacher training and bonuses. It is worth emphasizing that the randomized evaluations of these interventions did not find statistically significant average project impacts on a wide range of educational outcomes (Glewwe, Kremer, and Moulin, 2009). The schools that benefited from these previous programs were found in all eight geographic zones; the distribution of the 27 schools across the eight zones is: Agenga/Nanguba (5 schools), Bunyala Central (1), Bunyala North (4), Bunyala South (2), Bwiri (4), Funyula (5), Namboboto (1), Nambuku (5). The results in the current paper are similar when including controls for participation in these other NGO programs (results not shown).

The schools were "stacked" as follows. Schools were divided by geographic division, then zone (alphabetically), and then listed according to school enrolment (as of February 1997, for grades 3 through 8) in ascending order. If there were, say, four schools in a zone, they would be listed according to school enrolment in ascending order, then they would be assigned consecutively to Group 1; Group 2; Group 3; Group 1. Then moving onto the next zone, the first school in that stratum would be assigned to Group 2, the next school to Group 3, and so on. Thus the group assignment "starting value" within each stratum was largely arbitrary, except for the alphabetically first zone (in the alphabetically first division), which assigned the school with the smallest enrolment in the zone to Group 1. Finally, there were three primary schools excluded from the original stacking of 72 schools that were added back into the sample for the original randomization, to bring the sample up to 75. These schools were originally excluded for similar reasons as listed above – e.g., two are rather geographically isolated, and the third is a relatively high quality school located near Funyula Town. However, in the interests of boosting sample size, these three schools were included in the list randomization alphabetically as the "bottom" three schools in the list.

Deaton (2010) raises concerns about the list randomization approach, in the case where the first school listed in the first randomization "triplet" is different than other schools (in our case, the concern would be that it has lower than average school enrolment); the same concerns would apply to several other well-known recent field experiments in development economics, most notably Chattopadhyay and Duflo (2004). However, this is not a major threat to our empirical approach. Following Bruhn and McKenzie (2009) we include all variables used in the randomization procedure (such as baseline school enrolment) as explanatory variables in our regressions, thus controlling for any direct effect of school size, and partially controlling for unmeasured characteristics correlated with school size. Coefficient estimates on the deworming treatment indicator are largely unchanged whether or not these additional explanatory variables are included, suggesting that any bias is likely to be small. The difference in average school enrollment between the treatment and control groups is small and not statistically significant (Table S2). Moreover, even if the first school in the first randomization triplet were an outlier along some unobserved dimension (which seems unlikely), given our sample size of 75 schools and 25 randomization triplets, and the fact that school size is not systematically related to treatment group assignment for the other 24 randomization triplets (as discussed above), approximately 96% of any hypothesized bias would be eliminated. Taken together, the prospective experimental design in the current paper is likely to yield reliable causal inference. Figure S1 further summarizes the research design.

Miguel and Kremer (2004) present evidence of balance across treatment groups along a fuller set of baseline covariates for the treatment and control groups (in their Table I), reproduced here as Table S1. The same balance on predetermined characteristics is also evident among the subsample of respondents no longer enrolled in school and among those currently working for wages (see Tables S4 and S14), two subsamples that feature in some of the analysis of this paper.

While it is not necessary to utilize baseline data in a randomized experiment, since treatment versus control differences yield unbiased effect estimates even when relying solely on follow-up data, some readers may be concerned that baseline data on school participation was lacking based on incorrect claims (e.g., Taylor-Robinson *et al.* 2012, 2015). Miguel and Kremer (2004) show three pieces of baseline data demonstrating balance on educational variables: (i) baseline data from school registers, which show nearly identical measured attendance across the three treatment groups; (ii) baseline data for Group 2 versus Group 3 from the unannounced school attendance checks during 1998 (when both groups were "control"), showing no statistically significant differences in school participation (and in fact, school participation was, if anything, slightly lower in Group 2 schools in that year, making the large positive school participation difference between Group 2 and Group 3 in 1999 even more noteworthy); and (iii) baseline balance along a wide range of other educational, health and socioeconomic measures (as reproduced in Table S1 below).

B. Econometric estimation of externalities

B.1 Estimating Treatment Effects in the Presence of Externalities

Define the complete vector of lagged deworming treatment saturation levels at all distances and in all time periods (excluding $P_{j,0,\delta}$, which we have already accounted for) as P_{j,t^*}, $P_{j,t^*} \equiv \{P_{j,0,2\delta}, P_{j,0,3\delta}, \dots, P_{j,0,t^*\delta}, P_{j,1,\delta}, P_{j,1,2\delta}, \dots, P_{j,1,t^*\delta}, \dots, P_{j,t^*,\delta}, P_{j,t^*,2\delta}, \dots, P_{j,t^*,t^*\delta}\}$, where the subscript $n\delta$ denotes deworming treatment saturation at distances between $(n-1)\delta$ and $n\delta$ from school j.

We can generalize our empirical quantities of interest taking into account these additional externality effects. Consider the impact of a program in which the share of nearby population receiving deworming in the original period is $P_{j,0,\delta} = p_0$ and the vector of additional externality exposure is $P_{j,t^*} = \underline{p}_1$:

$$\pi_{t^*}\left(p_0, \underline{p}_1\right) \equiv E\left[Y_{ijt^*} \mid T_{j,0} = 1, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right] - E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = 0, P_{j,t^*} = 0\right] \quad \text{(eqn. B1)}$$

(where $P_{j,t^*} = 0$ indicates that all elements of the vector are equal to zero). As above, define the expected outcome in untreated communities surrounded only by other untreated communities (i.e., "pure control" communities uncontaminated by exposure to treatment schools) as $y_{0,t^*} \equiv E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = 0, P_{j,t^*} = 0\right]$. The generalized difference in expected outcomes between treated and untreated communities at given local treatment saturation exposure is:

$$\lambda_{1t^*}\left(p_0, \underline{p}_1\right) \equiv E\left[Y_{ijt^*} \mid T_{j,0} = 1, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right] - E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right] \quad \text{(eqn. B2)}$$

The difference in average outcomes between untreated communities at initial treatment saturation ($P_{j,0,\delta} = p_0$) versus those only benefiting from the additional externalities is:

$$\lambda_{2t^*}\left(p_0, \underline{p}_1\right) \equiv E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right] - E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = 0, P_{j,t^*} = \underline{p}_1\right] \quad \text{(eqn. B3)}$$

The new term to consider is the difference between those communities only benefiting from the additional externalities versus the pure control communities:

$$\lambda_{3t^*}\left(p_0, \underline{p}_1\right) \equiv E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = 0, P_{j,t^*} = \underline{p}_1\right] - y_{0,t^*} \quad \text{(eqn. B4)}$$

The sum of these three effects is $\pi_{t^*}\left(p_0, \underline{p}_1\right) \equiv \lambda_{1t^*}\left(p_0, \underline{p}_1\right) + \lambda_{2t^*}\left(p_0, \underline{p}_1\right) + \lambda_{3t^*}\left(p_0, \underline{p}_1\right)$.

Closely following the proof to proposition 1 in the text, Assumption 1 implies that the new externality term $\lambda_{3t^*}\left(p_0, \underline{p}_1\right)$ is non-negative, and thus that once again an analysis that does not account for cross-community spillover effects and focuses on $\lambda_{1t^*}\left(p_0, \underline{p}_1\right)$ yields a lower bound on both quantities of empirical interest, $\pi_{t^*}\left(1, \underline{p}_1\right)$ and $\pi_{t^*}\left(p_0, \underline{p}_1\right)$.

Proposition B1 (Bounding the treatment effect): Suppose for all $\left(p_0, \underline{p}_1\right)$, $E\left[Y_{ijt^*} \mid T_{j,0} = 1, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right] \geq E\left[Y_{ijt^*} \mid T_{j,0} = 0, P_{j,0,\delta} = p_0, P_{j,t^*} = \underline{p}_1\right]$, then $\pi_{t^*}\left(1, \underline{p}_1\right) \geq \pi_{t^*}\left(p_0, \underline{p}_1\right) \geq \lambda_{1t^*}\left(p_0, \underline{p}_1\right)$ for all $\left(p_0, \underline{p}_1\right)$.

Proof: The proof that $\pi_{t^*}\left(1,\underline{p_1}\right) \geq \pi_{t^*}\left(p_0,\underline{p_1}\right)$ follows directly from the proof to Proposition 1. We next show that $\pi_{t^*}\left(p_0,\underline{p_1}\right) \geq \lambda_{1t^*}\left(p_0,\underline{p_1}\right)$. It is sufficient to show that both $\lambda_{2t^*}\left(p_0,\underline{p_1}\right)$ and $\lambda_{3t^*}\left(p_0,\underline{p_1}\right)$ are non-negative. The proof that $\lambda_{2t^*}\left(p_0,\underline{p_1}\right) \geq 0$ follows directly from the proof to Proposition 1. For the sign of $\lambda_{3t^*}\left(p_0,\underline{p_1}\right)$, consider the vector of saturation exposure $\underline{p_1}$ where $p_{\tilde{t},\tilde{\delta}} \geq 0$ for each element of the vector. The monotonicity assumption (Assumption 1) implies that $\lambda_{3t^*}\left(p_0,\underline{p_1}\right) = E\left[Y_{ijt^*} | T_{j,0} = 0, P_{j,0,\delta} = 0, P_{j,t^*} = \underline{p_1}\right] - y_{0,t^*} \geq 0$. The result follows. □

B.2. Understanding externalities and treatment interactions across multiple outcomes

We also estimated the interaction between the treatment indicator and local treatment saturation. The sign of this interaction is theoretically ambiguous. While there are more infections to eliminate in more highly infected areas and this would naturally lead to larger impacts in such areas, areas with higher prevalence will also typically have conditions more conducive to transmission of the disease (i.e., soil moisture). Thus re-infection is likely to occur more rapidly in these areas, dampening treatment impacts relative to areas where it takes longer for re-infection to occur. Empirically, we typically do not find significant interaction effects, even when jointly testing for significance across multiple outcomes (see discussion below). Nor do we generally find statistically significant effects on non-linear terms in local treatment saturation, leading us to focus on a linear functional form of $\lambda_2(p) = p\lambda_2^*$ (where λ_2^* is a constant) for the externality effect.

Given the range of outcomes we explore in Tables 1 to 4 – 28 in total (not including the 2001 health result in Table 1) – it is useful to carry out a summary test to assess the existence of deworming treatment externalities across schools. The simplest such test is to assess whether the externality effect has the same "sign" as the direct deworming treatment estimate across all 28 outcomes. This test effectively tests the null hypothesis that the externality effect is symmetric with a mean of zero, in which case the estimated effects should be evenly distributed on both sides of zero. Examining the 28 outcomes, we immediately see that the externality estimates disproportionately have the same sign as the direct deworming effect (i.e., the coefficient estimate on the treatment school indicator). Specifically, the two signs are the same in 23 out of 28 outcomes in the full sample (examining males and females together). This pattern is extremely unlikely to occur by chance. In the case where the externality effect was pure "noise", the likelihood of a sign "match" between the two terms would be distributed as a binomial distribution with p=0.5. In that case, 23 of 28 pairs of estimates would have the same sign roughly six times in 10,000 cases. This pattern provides empirical support for the monotonicity assumption in section 3.1.

This "sign test" has limitations, as it ignores information on the magnitude of the estimated effects, and does not take into account that some of the outcomes are correlated with others (i.e., total earnings are correlated with total hours worked). An alternative test that accounts for the first of these concerns estimates the correlation between the t-statistics for the direct effect and the externality effect (across all outcomes). We obtain a correlation between the pairs of t-statistics of 0.655 (P-value < 0.002). The results are very similar when considering the correlation between the coefficient estimates across these outcomes instead of the t-statistic (not shown), but the t-statistic approach provides a useful normalization. These results confirm the finding from the simple sign test discussed above, and in both cases we reject the hypothesis of no externality effect at high levels of confidence.

We carried out a related analysis in order to assess whether there is robust evidence of interactions between treatment assignment and cross-school externality effects across all outcomes. Specifically, we examined the correlation between the t-statistics for the coefficient estimates on the direct effect and the interaction (Treatment x Externality) effects across all outcomes. Note that we use the zonal-level baseline infection rate, rather than individual-level data (which was not collected at baseline for the control group for ethical reasons); using zonal averages is likely to introduce some measurement error and attenuation bias, and thus these interaction effect estimates may understate the true extent of differential impacts in high worm infection areas. We obtain a relatively weak and not statistically significant correlation of 0.306 (P-value=0.189) for the full sample. Thus we cannot reject the null hypothesis of no relationship between the treatment effect and the interaction effect.

Finally, using a seemingly unrelated regression (SUR) estimation across 24 of the 28 regressions (ignoring the results that were run on different samples, including the 2001 health result and the miscarriage result in Table 1, the out-of-school subsample results in Tables 3 and 4, and the trimmed profits result in Table 4), we reject the hypothesis that the cross-school externality effect from 0-6 km is zero (P<0.001).

As we note in the main text, the original study found externality impacts out to 3 km upon correction of a coding error (detailed in Miguel and Kremer, 2014 and Hicks et al., 2015). However, spillover effects are likely to diffuse spatially over time. We perform this same SUR estimation including separate terms for 0-3 km and 3-6 km spillovers, and conditional on 0-3 km we also reject the hypothesis that the effects from 3-6 km are zero. For this reason, we include externality impacts from 0-6 km in our primary analysis.

C Discussion of additional empirical results

C.1. Sample tracking and attrition

As time progressed and the pace of locating respondents slowed, a representative (random) subsample containing approximately one quarter of still-unfound respondents was drawn. Those sampled were tracked "intensively" (in terms of enumerator time and travel expenses) for the remaining months, while those not sampled were no longer actively tracked. We re-weight those chosen for the "intensive" sample by their added importance to maintain the representativeness of the sample. As a result, all figures reported here are "effective" tracking rates (ETR), calculated as a fraction of those found, or not found but searched for during intensive tracking, with weights adjusted appropriately. This is analogous to the approach in the U.S. Moving To Opportunity study (Kling, Liebman and Katz, 2007; Orr *et al.,* 2003). The effective tracking rate (ETR) is a function of the regular phase tracking rate (RTR) and intensive phase tracking rate (ITR) as follows:

$$\mathrm{ETR} = \mathrm{RTR} + (1 - \mathrm{RTR}) * \mathrm{ITR}\,. \tag{S1}$$

The RTR in KLPS-2 is 65.0% and the ITR is 62.1%, which implies that the ETR = 86.7% when including all those surveyed, plus those who refused or were found but were unable to be surveyed, and the deceased.

A midterm round (KLPS-1) was collected in 2003-05. We focus on the KLPS-2 since it was collected at a more relevant time point to assess adult life outcomes: most respondents are adults by 2007-09 (median age 22 years vs. 18 in KLPS-1), the vast majority have completed school, many have married, and a growing share are employed.

Table S4 shows that, other than the treatment saturation proportion, there are no significant differences across the treatment groups in the out-of-school subsample along observable dimensions, and Table S5 similarly shows that there is minimal selection into the out-of-school subsample along observable characteristics across the treatment and control groups. There are no significant deworming impacts on migration out of the study district or to urban areas (not shown).

C.2. Additional results related to Table 1

Worms' average lifespan in the human body is only one to three years (Anderson and May, 1991; Bundy and Cooper, 1989). So deworming in a school 10 years ago would affect current worm load insofar as it had a persistent epidemiological effect on worm load seven years later, an effect that is likely to be extremely small given the high reinfection rates in our data. Thus any health impacts are likely to work through other channels, as discussed in the main text. Note that we see no evidence that students in treatment schools are more likely to purchase deworming medicine as adults (not shown).

Figure S2 visually presents the difference in moderate-to-heavy worm infection rates among the three program group (Group 1, Group 2 and Group 3) during 1998-2002. The impact of treatment on infection rates in early 2001 is presented in Table 1. It is apparent that there are high levels of moderate-to-heavy worm infections in the study area, and that mass deworming leads to sharp reductions in worm infection rates.

Point estimates suggest women in the treatment group have had somewhat fewer pregnancies and are less likely to be married or the parent of a child (not shown), although effects are not significant.

C.3. Additional results related to Table 2

Given that KLPS-2 school enrollment data misses out on attendance impacts, which are sizeable, a plausible lower bound on the total increase in time spent in school induced by the deworming intervention is the 0.137 gain in school participation from 1998-2001 plus the school enrollment gains from 2002-2007 (multiplied by average attendance conditional on enrollment), which works out to nearly 0.3 additional years of schooling (not shown).

There is little evidence of differential selection along observables into the out-of-school sample between the treatment and control groups (Tables S4-S5). Note that the out-of-school variable cannot necessarily be taken as an indicator of more (or less) schooling, since it could indicate either rapid on-time completion of secondary school, or more post-secondary school education.

One potential concern with longitudinal data collected over such a long time span is attrition. Reassuringly, we cannot reject that treatment effect estimates are equal in the regular tracking and the intensive tracking subsamples for the outcomes in Table 2 (results not shown).

C.4. Additional results related to Table 3

There is no significant change in the proportion in the treatment group working at all (greater than zero hours in the past week), which is roughly 68% overall (Table S15, Panel A) and 73% for the out-of-school subsample (not shown). There is thus a considerable degree of "non-activity" for a young adult population. In the full sample, females are somewhat more likely to be classified as non-active, which may be related to the fact that most out-of-school females have had at least one pregnancy. However, note that some females are engaged in home production or child-rearing activities that were not collected in detail in the KLPS-2 survey and thus not classified as "work hours" here. This possible under-reporting of total work hours, which is likely to be particularly important for females, provides another reason to conduct the analysis separately for males and females.

77% of self-employed women are in retail. While 44% of the male self-employed also work in retail, others work in occupations such as commercial fishing (21%), small manufacturing (12%) and passenger transport (9%), several of which require substantial physical strength and regularly take the respondents farther afield, and thus may be more difficult to combine with child care.

In further evidence consistent with the PRH model, we find suggestive evidence of a more general shift into high work hour occupations for males, but not for females. To explore this, we assigned each individual in the sample to a broad occupation group (or set of groups, if the individual worked in more than one occupation at the time of the follow-up survey), and created a measure of average work hours (among control individuals) within each occupation group. These included farming, six different self-employment occupations, and 13 different wage employment categories. We then regressed this measure of average work hours on the treatment measure, an interaction between treatment and gender, and our standard regression controls. Among males, we find suggestive evidence of a shift into occupations characterized by higher average work hours (coefficient estimate 1.71, SE 1.01, $P < 0.10$). Further detail on patterns of employment among wage earners is provided in Table S16.

We find no evidence of differential effects by age cohort (Tables S6-S9, column 7). We examine impacts in geographic zones within the sample with different levels of baseline worm infection rates, but do not find significantly different treatment effects for hours worked or meals eaten, nor when we separately examine geohelminths and schistosomiasis (where the latter is proxied by proximity to Lake Victoria, where schistosomiasis is concentrated, Tables S10-S11). Deworming treatment effects typically remain positive, similar in magnitude and significant among schools located more than 5 km from Lake Victoria, in areas where schistosomiasis is rare, suggesting that the results are not driven by schistosomiasis alone (Tables S10-S13). Brooker et al. (2000) argues that distance to Lake Victoria is a good proxy for schistosomiasis infection in this region of Kenya.

ICS, the NGO which undertook the PSDP program, typically required cost-sharing, and in 2001, a randomly chosen half of the Group 1 and 2 schools took part in a program in which parents had to pay a small positive price (US$0.27 on average) to purchase the drugs, while the other half of Group 1 and 2 schools received free treatment (as did all Group 3 schools); the randomization was carried out with a computer random number generator. In 2002 and 2003, all schools received free treatment. Cost-sharing reduced deworming take up from approximately 75% to 18% (Kremer and Miguel 2007).

We estimate negative coefficients on the 2001 cost-sharing indicator variable when the dependent variable is hours worked, meals eaten, passing the primary school leaving exam, or log wage earnings (Tables S6-S9). The fact that coefficient estimates on the cost-sharing indicator are of opposite sign compared to the direct treatment effect, and typically smaller in absolute value, is reassuring.

C.5. Additional results related to Table 4

The wage earnings result (in Table 4, Panel B) is robust to several alternative specifications. It changes little in response to trimming the top 1% of earners, so the result is not driven by outliers, and to including a full set of gender-age fixed effects (results not shown).

A decomposition along the lines of Oaxaca (1973) – which uses mean earnings by occupation in the control group as a reference point – indicates that among wage earners, 75% of the higher earnings in the treatment group can be accounted for by occupational shifts (not shown), for instance, the shift into manufacturing and out of casual labor.

Trimming the top 5% of self-reported profits results in a similarly sized treatment effect of 341 shillings ($P < 0.10$). We obtain similar results on firm profits (Table 4, Panel C) using both the inverse sine hyperbolic transformation and log profits (not shown). There are large, positive but not statistically significant impacts on a monthly profit measure based directly on revenues and expenses reported in the survey (not shown). We focus here on self-reported profits in the last month, which appear to be less noisy. De Mel, McKenzie and Woodruff (2009) argue in favor of focusing on self-reported profits rather than computed profits in their work on small firms in low-income countries.

There appears to be a shift in the distribution of log self-employed profits for men, but it is less prominent for women (Figures S3-S4).

A consumption expenditure module was collected as a small pilot for 255 respondents. The estimated effect on total per capita consumption is near zero and not statistically significant but the confidence interval is large and includes both substantial gains and losses (not shown).

There is no evidence that the quality of labor in agriculture in the treatment group was lower than in the control group based on observable characteristics. For instance, average education levels among those working in agriculture are, if anything, slightly higher in the treatment group, and the difference is statistically significant at the 10% level among males. The share of male labor in agriculture is also, if anything, slightly higher in the treatment group. Both of these patterns are likely to be associated with somewhat higher agricultural productivity in the treatment, since male agricultural labor is higher paid in the area (in our data) and education tends to be associated with higher labor productivity in agriculture.

The multiple testing False Discovery Rate (FDR) adjustments presented in Tables S18-S21 and discussed in the text are carried out separately for the outcome variables within each table. Other recent economics research adopts a similar approach to multiple testing adjustments within domains of related hypotheses, including Casey et al (2012) and Finkelstein et al (2012). Regarding the multiple testing adjustment by gender subsamples, for women, q-value < 0.10 for improvements in self-reported health, reduction in miscarriage, increases in secondary school enrollment, passing the secondary school entrance exam and years enrolled, and for men, q-value < 0.10 for increases in total hours worked (Tables S18-S21). Anderson (2008) presents multiple instances in which the FDR q-value is smaller than the per-comparison (naïve) P-value, and this occurs for several outcomes in our data. This pattern may occur in both FDR and Family-wise Error Rate (FWER) adjustments because both methods enforce the original monotonic ordering of unadjusted p-values; see Anderson (2008) for details.

C.6. Additional results related to Figure 1

The distribution of work hours in the last week in our data is similar to several other recent labor surveys in Africa (in the International Income Distribution Database, I2D2, World Bank: Development Research Group, Poverty and Inequality Unit), including for average work hours, suggesting that our sample of workers is not unusual. For instance, a sizeable 25% of KLPS-2 males who work positive hours for wages worked more than 60 hours in the last week, while in the 2007 Tanzania survey in I2D2, 23% of rural male wage earners worked more than 60 hours.

C.7. Additional results related to Table 5 and the internal rate of return calculation
For consumption, the marginal tax rate in our sample of mostly rural residents may be lower than the national average since some consumption comes in the form of food produced on their own farms. Home produced food constitutes roughly 10% of total household consumption (in data from a pilot consumption survey); note that it is likely that an even smaller share of the marginal income gains experienced by deworming beneficiaries are in the form of home produced food, given the documented shift towards cash crops and entrepreneurship. Even if home produced food is entirely untaxed, the tax revenue generated still outweighs the deworming subsidy by a ratio of roughly 10 to 1 if remaining elements of consumption are taxed at the national average.

Wanjala (2007) discusses VAT collection in Kenya since the 1990s. The VAT burden has risen over time due to the progressive formalization of many sectors of the Kenyan economy (increasing payments) and a recent tax reform that expanded the number of good subject to VAT. We find that future revenue exceeds subsidy costs in the no health spillover case (health spillover case) as long as taxes capture 8% (3%) of the additional lifetime earnings of $142.43 ($766.81).

We can alternatively focus solely on income gains among the subsample of wage earners and the self-employed (outside of agriculture), as a larger share of their earnings are likely to be subject to tax than subsistence farmers. This exercise implicitly sets earnings gains among subsistence farmers to zero. An analysis that focuses solely on earnings and profits yields similarly large increases in government revenue, at US$17 dollars in net revenue for each dollar of subsidies (Appendix Table S22).

The main result also holds if earnings gains are assumed to remain constant over time, ruling out further gains over the life cycle due to work experience and ruling out further economic growth, yielding $1.96 in net revenue for each dollar of deworming subsidies (Appendix Table S23). We also estimate the impact of partial deworming subsidies assuming that their labor market impacts are proportional to the number of people dewormed. As noted above, theory provides little guidance on the shape of the function linking deworming treatment intensity to outcomes, but we cannot reject this hypothesis of linearity. Point estimates suggest increasing treatment rates from 19% to 75% less than proportionally increases the benefits of treatment, but estimated effects are imprecisely estimated.

The result that the NPV of revenue exceeds subsidy costs also holds if we consider both the opportunity cost of additional time spent in secondary school by adolescents as well as the benefits in terms of future wage growth, yielding $12 in future revenue for each $1 in subsidies (Appendix Table S24).

Note that Kremer and Miguel (2007) find no evidence that people with serious worm infections are more likely to pay for deworming treatment during the cost-sharing phase of the project in 2001.

We use the estimated $\lambda_{1,\gamma}$ and $\lambda_{2,\gamma}$ values from year 10 post-treatment onwards, and then use the pattern of lifecycle earnings reported in the most recent publicly available data, the 1998/1999 Kenya Integrated Labour Force Survey, to scale these effect sizes over time. This assumes that earnings effects will stay the same in percentage terms over the life cycle. To the extent that experience and education are complementary (as argued, for example, by Heckman among others); that investments in establishing new businesses may yield longer-run payoffs, and that differences in earnings and meals may not yet have appeared among those still in school, it seems more likely that effects would grow over time than that they would shrink over time. The coefficient estimates on years of work experience and years squared in the 1998/1999 Kenyan labor data are 0.102 and -0.001, respectively. Future earnings are also assumed to increase by average annual per-capita GDP growth rate in Kenya during 2001 to 2011, namely 1.52% (World Bank Development Indicators). This is a conservative assumption, given that annual Kenyan per-capita GDP growth has been faster than 1.52% since 2011.

Note that being free of worms, having a lower miscarriage rate and better self-reported health, and having more education could also be considered additional benefits of deworming, but we do not include these in the IRR calculation, making it a conservative calculation.

Departing from the implicit assumption of a small open economy with capital supplies from abroad would indeed allow for GE effects that we could not measure. Allowing for these effects could potentially either increase or decrease the estimated rate of return to deworming, and its fiscal impact, but we think it is likely to increase these effects and unlikely to reverse them.

Consider first the purely pecuniary effects of increased labor supply by those who are dewormed. The increase in labor supply will cause a decrease in wages for workers and an increase in returns for owners of capital. To the extent that owners of capital pay greater marginal tax rates and consume fewer government services, the resulting redistributional effect will increase total net tax payments.

The analysis becomes more complicated if the supply of labor and capital is endogenous. The increase in returns to capital would spur more capital accumulation under most models. The decline in wages could lead to either an increase or decrease in labor supply depending on the balance of income and substitution effects, but this would less than fully offset the direct effect of increased labor supply due to better health. The rate of return and fiscal effects could thus be either larger or smaller than under the small open economy assumptions.

It is worth noting that because the experimental assignment is at the school level (rather than the individual level), we already capture that portion of the GE effects that occur within the local school catchment area and in the young adult population we survey. Since a considerable fraction of our population produce locally-consumed goods and services (e.g., retail trade), and we see no change in migration rates with the program, a considerable fraction of GE spillovers may take place within the school catchment area. Finally, note that epidemiological externalities outside the village and cohort are likely to have positive impacts on earnings and net tax payments. It thus strikes us as unlikely that effects on those outside our sample would change the main conclusion on the desirability of deworming subsidies.

Figure S1: Project Timeline of the
Primary School Deworming Program (PSDP) and the Kenya Life Panel Survey (KLPS)

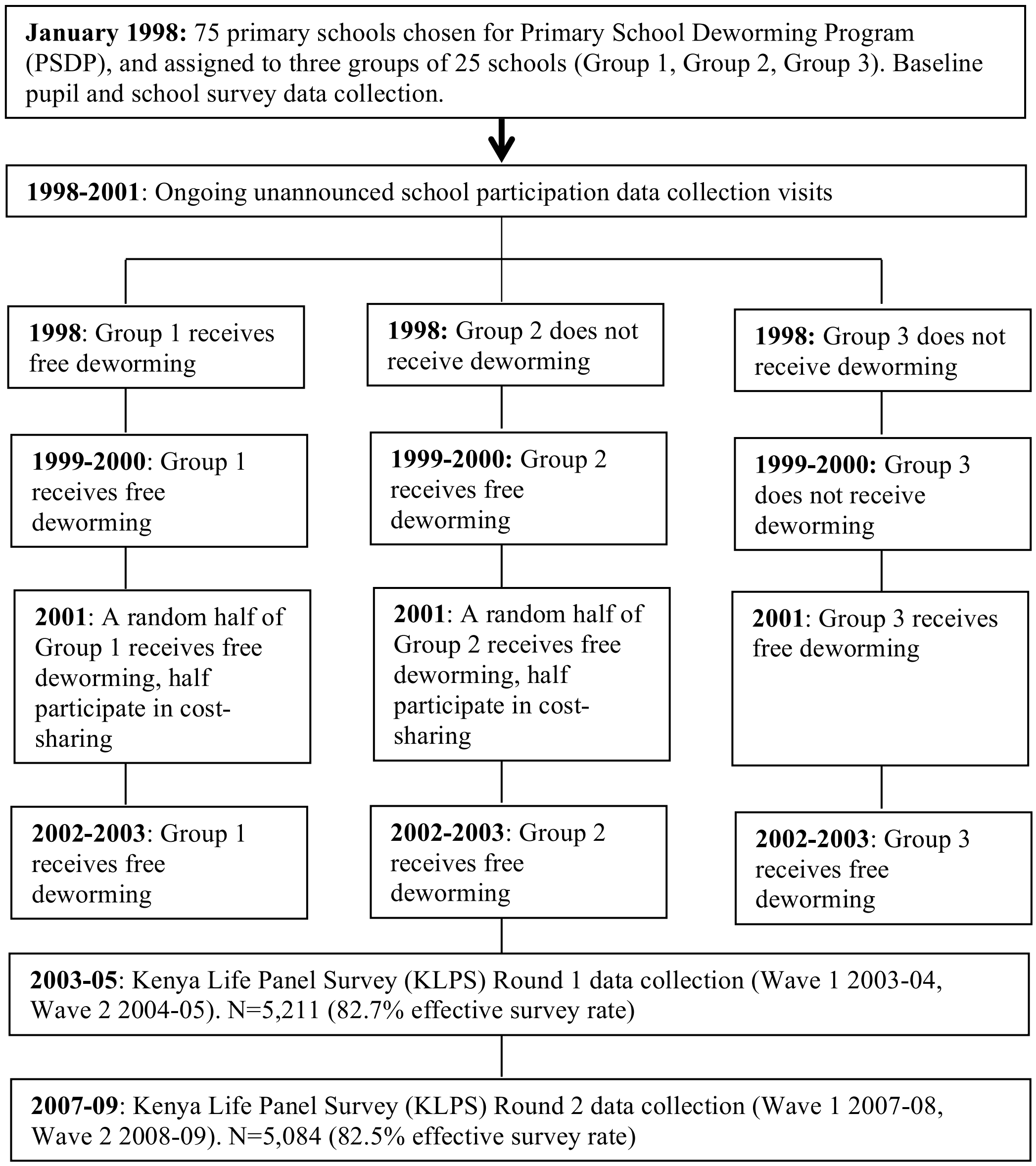
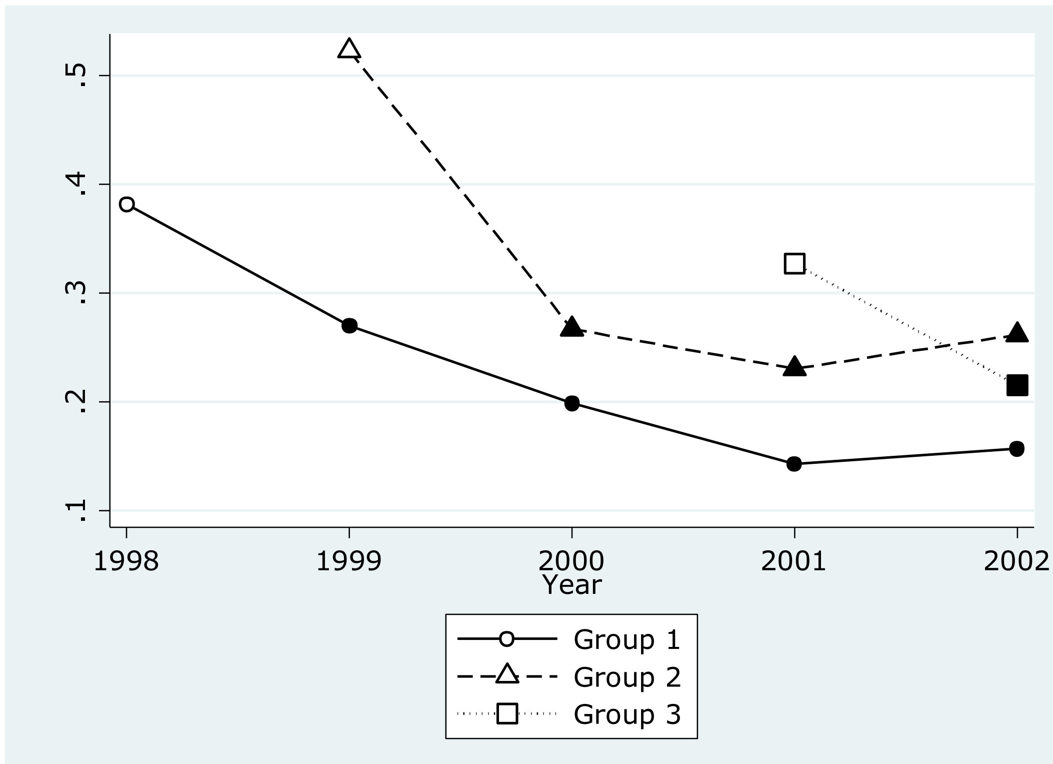
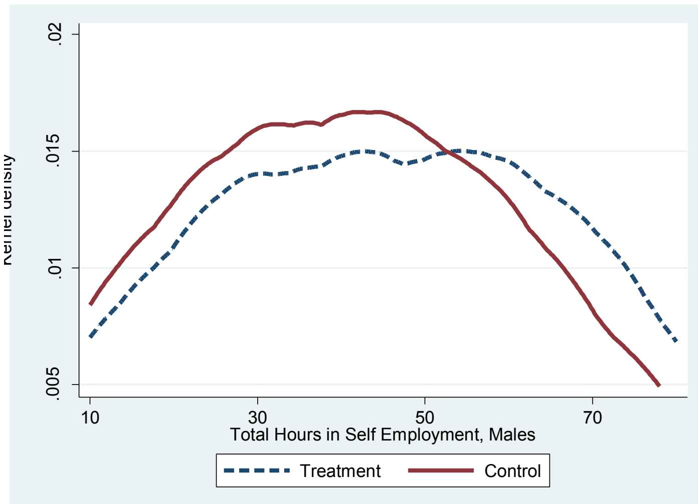


Figure S2: Worm infection rates over time, by treatment group

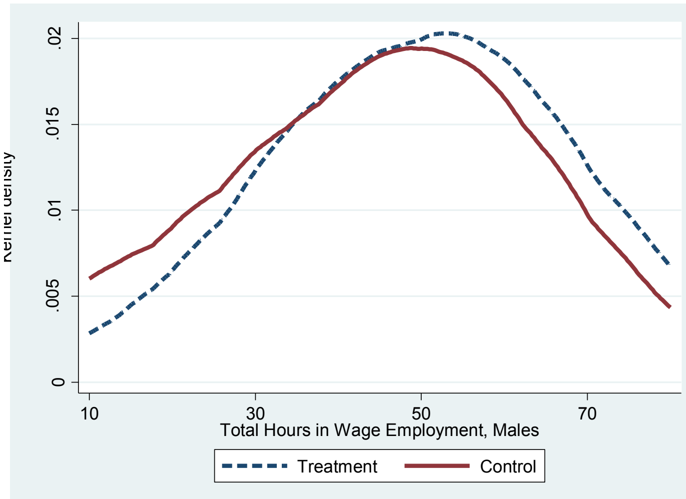


Notes: This figure is produced using the data from individuals who received parasitological testing between 1998 and 2002. The hollow symbols denote control (pre-treatment) group-year observations (i.e., Group 1 in early 1998, Group 2 in early 1999, and Group 3 in early 2001), and the filled symbols denote treatment observations (Group 1 in 1999-2002, Group 2 in 2000-2002, and Group 3 in 2002).

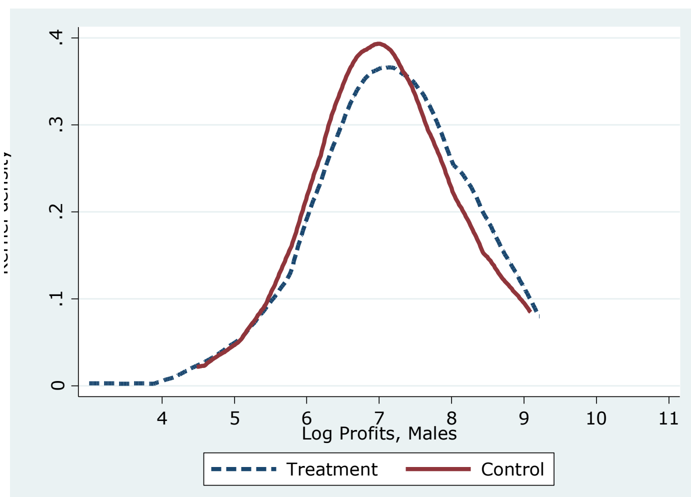
Figure S3: Hours worked (if working 10 to 80 hours in sector) and earnings among males, treatment versus control
Panel A: Hours worked in self-employment in last week; Panel B: Hours worked in wage employment in last week;
Panel C: Log self-employed profits in last month (top 5% trimmed); Panel D: Log earnings in wage employment in past month.



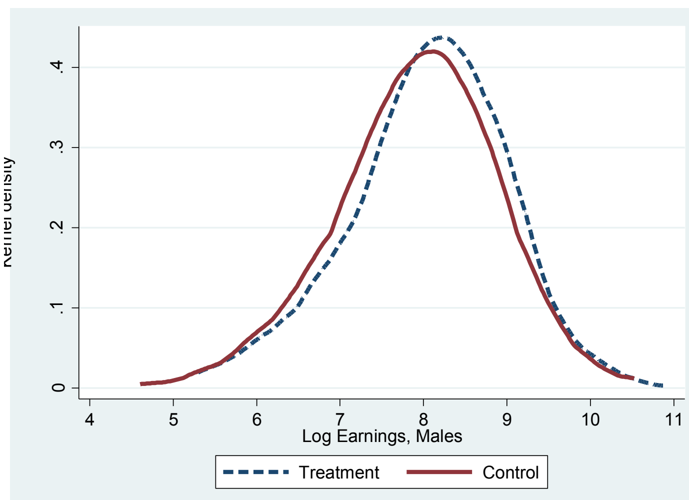
(A)



(B)



(C)



(D)

Figure S4: Hours worked (if working 10 to 80 hours in sector) and earnings among females, treatment versus control
Panel A: Hours worked in self-employment in last week; Panel B: Hours worked in wage employment in last week;
Panel C: Log self-employed profits in last month (top 5% trimmed); Panel D: Log earnings in wage employment in past month.

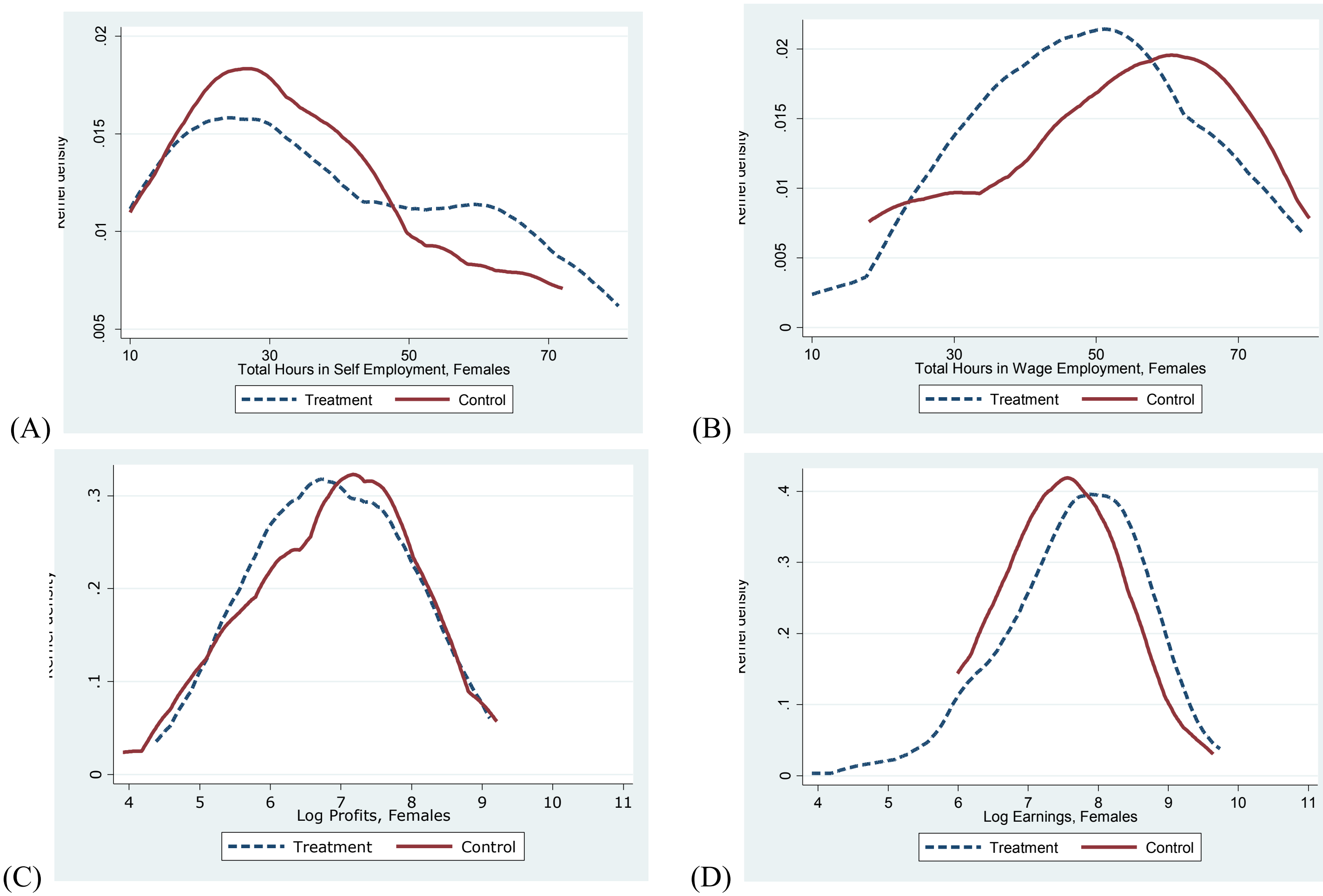
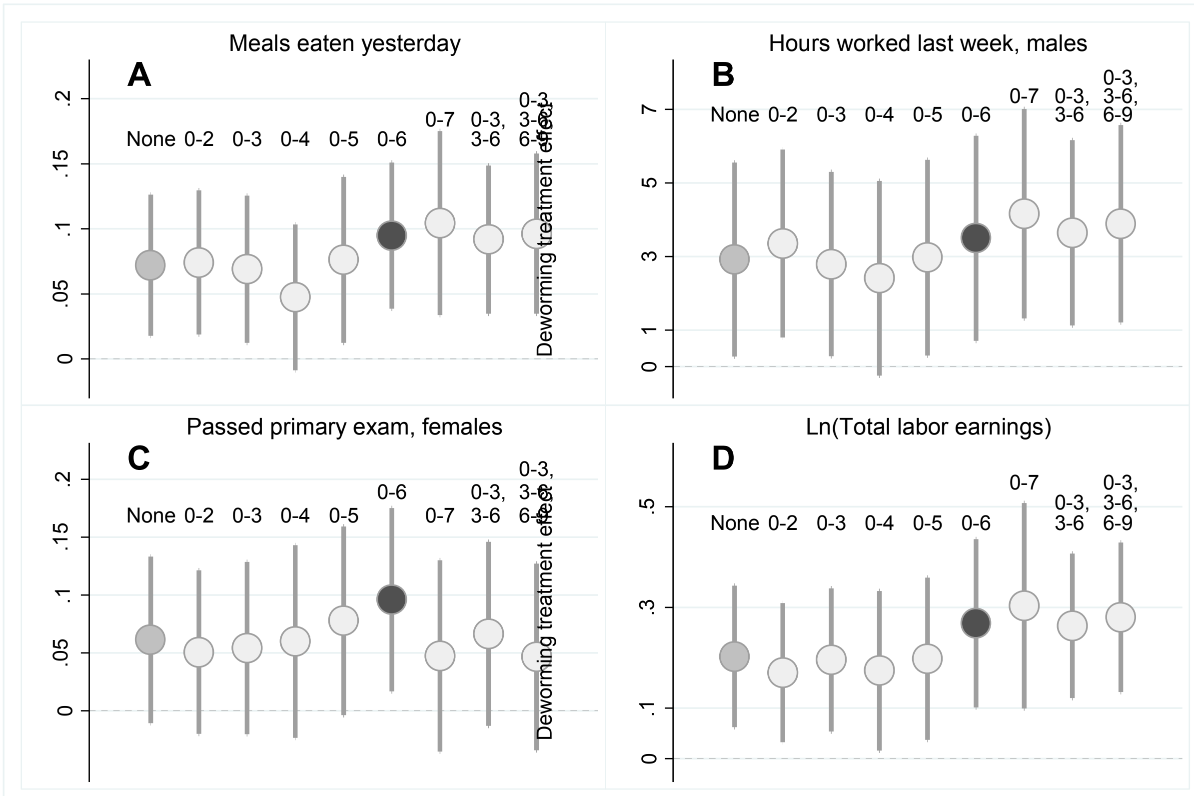


Figure S5: Deworming treatment effect estimates conditional on different specifications of the cross-school externality effect
Panel A: Number of meals eaten, Panel B: Hours worked last 7 days, all sectors (males)
Panel C: Passed primary school leaving exam (females), Panel D: Ln(Total labor earnings), past month



Notes: Each circle denotes an estimate from a separate regression of the outcome on the treatment indicator, the standard set of regression controls (from Tables 1-4), but accounting for cross-school externality effects out to different distances. The dark grey estimate does not contain cross-school externality controls, the black estimate is the main specification shown in Tables 1-4 of the paper, and the light grey estimates account for cross-school externality effects at alternative distances (in km) as denoted in the figure. The horizontal lines denote the 95% confidence interval. The p-values for the point estimates presented in panel A range from 0.001 to 0.098; in panel B from 0.004 to 0.075; in panel C from 0.018 to 0.261; and in panel D from 0.000 to 0.031.

Table S1: 1998 Average pupil and school characteristics, pre-treatment[†]

	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[‡]	-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[††]	430.4	433.2	344.5	85.9	88.7
				(116.2)	(116.2)
Group 1 pupils within 3-6 km	1157.6	1043.0	1297.3	-139.7	-254.4
				(199.3)	(199.3)
Total primary school pupils within 3 km	1272.7	1369.1	1151.9	120.8	217.2
				(208.1)	(208.1)
Total primary school pupils within 3-6 km	3431.3	3259.8	3502.1	-70.8	-242.3
				(366.0)	(366.0)

[†]This table replicates Table 1 from Miguel and Kremer (2004), using the final data and coding as presented in Miguel and Kremer (2014). School averages weighted by pupil population. Standard errors in parentheses. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence. Data from the 1998 ICS Pupil Namelist, 1998 Pupil Questionnaire and 1998 School Questionnaire. [‡]1996 District exam scores have been normalized to be in units of individual level standard deviations, and so are comparable in units to the 1998 and 1999 ICS test scores (under the assumption that the decomposition of test score variance within and between schools was the same in 1996, 1998, and 1999).

[††] This includes girls less than 13 years old, and all boys (those eligible for deworming in treatment schools).

Table S2: Baseline (1998) summary statistics and PSDP randomization checks, and KLPS (2007-09) survey attrition patterns

	Treatment – Control (s.e.)			Control group mean (s.d.)		
Panel A: Baseline summary statistics	All	Male	Female	All	Male	Female
Age (1998)	-0.04	-0.16	0.08	12.0	12.2	11.7
	(0.11)	(0.17)	(0.12)	(2.6)	(2.7)	(2.5)
Grade (1998)	-0.03	-0.07	0.02	4.25	4.26	4.24
	(0.05)	(0.07)	(0.08)	(1.66)	(1.67)	(1.65)
Female indicator	-0.004			0.473		
	(0.019)					
School average test score (1996)	-0.013	-0.038	0.014	0.038	0.042	0.032
	(0.109)	(0.108)	(0.114)	(0.406)	(0.404)	(0.408)
Primary school located in Budalangi division indicator	-0.017	-0.030	-0.002	0.381	0.387	0.374
	(0.137)	(0.141)	(0.136)			
Population of primary school	58	49	68	436	445	426
	(54)	(51)	(57)	(146)	(145)	(146)
Total primary school students within 6 km	-34	1	-74	4,732	4,717	4,749
	(389)	(399)	(386)	(1,555)	(1,553)	(1,558)
Primary school students within 6 km in treatment schools (Group 1,2)	-296	-290	-302	3,381	3,375	3,388
	(260)	(271)	(255)	(1,022)	(1,022)	(1,024)
Saturation (P_j): Proportion of treated students within 6 km	-0.046***	-0.049***	-0.042**	0.542	0.543	0.541
	(0.017)	(0.018)	(0.017)	(0.059)	(0.059)	(0.060)
Years of assigned deworming treatment, 1998-2003	2.41***	2.45***	2.37***	1.68	1.68	1.67
	(0.08)	(0.10)	(0.09)	(1.23)	(1.24)	(1.23)
Panel B: Sample attrition, KLPS (2007-09)						
Found indicator[a]	-0.007	0.007	-0.021	0.867	0.878	0.854
	(0.017)	(0.022)	(0.025)			
Surveyed indicator	-0.003	0.016	-0.023	0.827	0.834	0.820
	(0.018)	(0.023)	(0.025)			
Not surveyed, deceased indicator	0.004	0.003	0.006	0.014	0.016	0.012
	(0.004)	(0.005)	(0.005)			

Notes: Panel A data is from the PSDP, and includes those surveyed in KLPS2. N=5,084 observations, with 2,595 males and 2,489 females (except for age, where N=5,072). Years of assigned deworming treatment is calculated using the treatment group of the respondent's school and grade. Respondents who "age out" of primary school are no longer considered assigned to treatment. School average test scores are from the 1996 Busia mock exam, and are converted to normalized individual s.d. units. Panel B includes all individuals surveyed, refused participation, deceased, found but unable to survey, and not found but sought in intensive tracking, for 5,569 respondents (3,686 treatment and 1,883 control; 2,827 males and 2,742 females). Observations are weighted to maintain initial population proportions. The "Treatment – Control" differences are derived from a linear regression on a constant and the treatment indicator. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence. [a] "Found" includes pupils surveyed, refused, deceased, and found but unable to survey.

Table S3: Deworming impacts on school enrollment, by year and gender

	1999	2000	2001	2002	2003	2004	2005	2006	2007	Total
Panel A: Primary School										
Overall	0.027**	0.033**	0.024*	0.025	0.020	0.004	0.004	0.013	0.005	0.155**
	(0.011)	(0.013)	(0.015)	(0.018)	(0.017)	(0.017)	(0.020)	(0.016)	(0.010)	(0.075)
Male	0.024	0.032**	0.022	0.028	0.045**	0.023	0.032	0.020	0.011	0.238**
	(0.015)	(0.016)	(0.017)	(0.021)	(0.019)	(0.022)	(0.025)	(0.021)	(0.017)	(0.102)
Female	0.029**	0.024	0.023	0.018	-0.006	-0.012	-0.029	-0.010	-0.011	0.026
	(0.014)	(0.018)	(0.026)	(0.024)	(0.024)	(0.024)	(0.023)	(0.019)	(0.014)	(0.098)
Panel B: Secondary School										
Overall	-0.002	0.007	0.026**	0.024**	0.027	0.023	0.028	0.006	0.008	0.149
	(0.002)	(0.007)	(0.011)	(0.011)	(0.019)	(0.025)	(0.029)	(0.029)	(0.025)	(0.130)
Male	-0.000	0.003	0.023*	-0.000	-0.018	-0.019	-0.029	-0.023	-0.013	-0.077
	(0.001)	(0.009)	(0.014)	(0.016)	(0.025)	(0.030)	(0.033)	(0.033)	(0.029)	(0.147)
Female	-0.004	0.010	0.028*	0.050***	0.067***	0.056**	0.068**	0.029	0.020	0.325***
	(0.005)	(0.010)	(0.015)	(0.012)	(0.018)	(0.024)	(0.029)	(0.030)	(0.029)	(0.124)
Panel C: Primary and Secondary School										
Overall	0.024**	0.040**	0.049**	0.049**	0.046**	0.027	0.036	0.017	0.008	0.313**
	(0.011)	(0.016)	(0.020)	(0.021)	(0.022)	(0.025)	(0.027)	(0.027)	(0.025)	(0.133)
Male	0.024	0.035**	0.042*	0.027	0.027	0.004	0.007	-0.007	-0.005	0.203
	(0.015)	(0.017)	(0.022)	(0.025)	(0.026)	(0.025)	(0.031)	(0.031)	(0.032)	(0.153)
Female	0.023*	0.034	0.052*	0.068**	0.059*	0.043	0.044	0.023	0.003	0.336*
	(0.014)	(0.021)	(0.028)	(0.028)	(0.032)	(0.035)	(0.032)	(0.030)	(0.029)	(0.179)

Notes: Each entry is from a separate OLS regression. For details on the regressions, see the "Notes" for Table 1. The analysis in this table uses KLPS-2 school enrollment data, which misses out on any additional school attendance impacts; in Miguel and Kremer (2004), pupil enrollment and attendance information were combined in the school participation measure. The seemingly paradoxical negative point estimates on female primary school enrollment (during 2003-2007) are likely driven by a combination of lower primary school repetition rates and higher rates of advancement to secondary school (shown in Panel B). Note that nearly the entire gain in female school enrollment during the period is driven by higher secondary school enrollment, while schooling gains are concentrated in primary school among males.

Table S4: Baseline (1998) summary statistics across treatment groups, for "out-of-school" subsample

	Treatment – Control (s.e.)			Control group mean (s.d.)		
	All	Male	Female	All	Male	Female
Age (1998)	-0.11	-0.26	0.02	12.7	13.2	12.3
	(0.12)	(0.21)	(0.12)	(2.4)	(2.5)	(2.2)
Grade (1998)	-0.07	-0.12	-0.03	4.61	4.69	4.54
	(0.06)	(0.08)	(0.10)	(1.59)	(1.58)	(1.60)
Female	-0.012			0.508		
	(0.022)					
School average test score (1996)	-0.011	-0.025	0.003	0.020	0.018	0.023
	(0.105)	(0.104)	(0.109)	(0.400)	(0.397)	(0.404)
Primary school located in Budalangi division	-0.033	-0.045	-0.021	0.408	0.423	0.394
	(0.139)	(0.145)	(0.137)			
Population of primary school	65	56	73	433	443	423
	(55)	(53)	(58)	(148)	(148)	(146)
Total primary school students within 6 km	-7	-19	5	4,667	4,685	4,650
	(400)	(415)	(400)	(1,571)	(1,563)	(1,579)
Primary school students within 6 km who are treatment (Group 1,2) pupils	-264	-276	-253	3,335	3,346	3,324
	(271)	(286)	(267)	(1,046)	(1,043)	(1,049)
Saturation (P_j)	-0.043**	-0.043**	-0.043**	0.541	0.541	0.542
	(0.017)	(0.017)	(0.017)	(0.059)	(0.057)	(0.060)
Years of assigned deworming treatment, 1998-2003	2.42***	2.44***	2.40***	1.42	1.39	1.45
	(0.09)	(0.11)	(0.10)	(1.21)	(1.21)	(1.21)

Notes: Data is from the PSDP, and includes individuals surveyed in KLPS2 who were not enrolled in school at the time of survey. N=3,873 observations, with 1,869 males and 2,004 females (except for age, where N=3,866). Years of assigned deworming treatment is calculated using the treatment group of the respondent's school and their grade, but is not adjusted for the treatment ineligibility of females over age 13 or assignment to cost-sharing in 2001; respondents who "age out" of primary school are no longer considered assigned to treatment. School average test scores are from the 1996 Busia District mock exam, and are converted to normalized individual standard deviation units. Observations are weighted to maintain initial population proportions. The "Treatment – Control" differences are derived from a linear regression on a constant and the treatment indicator. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence.

Table S5: Selection into the out-of-school subsample and into employment types

	Out-of-school at survey		In Wage Employment		In Self-Employment		In Agriculture	
	(1)	(2)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	-0.008	-0.011	-0.015	0.031	0.014	-0.132***	-0.011	-0.003
	(0.022)	(0.078)	(0.018)	(0.056)	(0.012)	(0.041)	(0.026)	(0.082)
Female	0.089***	0.108***	-0.151***	-0.130***	-0.004	-0.025**	0.043**	0.059*
	(0.014)	(0.023)	(0.013)	(0.021)	(0.011)	(0.011)	(0.017)	(0.035)
Grade	0.087***	0.099***	0.036***	0.040***	0.025***	0.023***	-0.023***	-0.022**
	(0.005)	(0.008)	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)	(0.011)
School average test score (1996)	-0.089***	-0.079***	-0.022	-0.085	-0.007	0.016	-0.060*	0.016
	(0.032)	(0.025)	(0.020)	(0.061)	(0.013)	(0.016)	(0.032)	(0.044)
Population of primary school	-0.004	-0.191**	-0.015	-0.047	-0.025	-0.297***	0.032	-0.035
	(0.040)	(0.084)	(0.033)	(0.099)	(0.029)	(0.054)	(0.046)	(0.085)
Primary school located in Budalangi division	-0.004	0.005	0.030	-0.007	0.031	0.066*	-0.047	0.073
	(0.034)	(0.029)	(0.029)	(0.060)	(0.023)	(0.034)	(0.045)	(0.045)
Saturation (P_j), demeaned	0.157	0.038	-0.032	0.034	0.025	0.317***	0.154	0.548**
	(0.152)	(0.240)	(0.105)	(0.290)	(0.058)	(0.094)	(0.150)	(0.236)
Total primary school students within 6 km, demeaned	-0.023***	-0.025***	0.004	0.009	-0.004	0.006	0.010	0.010
	(0.007)	(0.009)	(0.005)	(0.010)	(0.004)	(0.004)	(0.009)	(0.007)
Female * Treatment		-0.028		-0.033		0.030		-0.019
		(0.029)		(0.026)		(0.019)		(0.039)
Grade * Treatment		-0.018*		-0.005		0.002		-0.001
		(0.010)		(0.007)		(0.006)		(0.013)
School average test score * Treatment		-0.018		0.072		-0.032		-0.098*
		(0.044)		(0.064)		(0.020)		(0.051)
Population of primary school * Treatment		0.222**		0.026		0.341***		0.102
		(0.097)		(0.103)		(0.057)		(0.099)
Budalangi division * Treatment		0.014		-0.022		-0.076***		-0.141**
		(0.059)		(0.055)		(0.026)		(0.059)
Saturation (P_j) * Treatment		0.177		-0.091		-0.313***		-0.460
		(0.290)		(0.311)		(0.115)		(0.302)
Total primary school students within 6 km * Treatment		0.003		-0.005		-0.015**		-0.001
		(0.014)		(0.012)		(0.006)		(0.014)
R^2	0.137	0.141	0.074	0.081	0.025	0.032	0.025	0.033
Observations	5,058	5,058	5,081	5,081	5,083	5,083	5,043	5,043
Mean in the control group	0.748	0.748	0.166	0.166	0.100	0.100	0.555	0.555

Notes: For details on the regressions, see the "Notes" for Table 1. The outcomes are indicator variables, and the employment variables take on a value of one if the respondent worked positive hours in the activity. F-tests of the joint significance of the treatment indicator and all treatment interaction terms give p-values of 0.142 for out-of-school, <0.001 for in agriculture, 0.068 for in wage employment, and <0.001 for in self-employment.

Table S6: Heterogeneous deworming impacts, full sample

	Hours worked last 7 days, all sectors						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Deworming Treatment indicator	1.58	1.53	1.78	1.58	0.977	3.47**	1.49
	(1.04)	(1.03)	(1.12)	(1.04)	(1.073)	(1.48)	(1.17)
Female	-6.63***	-6.63***	-6.62***	-6.63***	-6.61***	-3.96**	-6.61***
	(0.94)	(0.94)	(0.94)	(0.94)	(0.94)	(1.71)	(0.93)
Female * Treatment						-3.94**	
						(2.00)	
Grades 5-7 in 1998							7.67***
							(1.62)
Grades 5-7 * Treatment							0.113
							(1.978)
Saturation (P_j), demeaned	10.20		18.43	10.32		10.32	11.40
	(7.80)		(12.87)	(7.78)		(7.75)	(7.81)
Saturation (P_j), demeaned * Treatment			-10.12				
			(12.20)				
Saturation (P_j), demeaned and squared				6.67			
				(42.26)			
Deworming treatment school students within 6 km (in '000s), demeaned		1.71					
		(1.43)					
Total primary school students within 6 km (in '000s), demeaned	0.194	-0.989	0.218	0.206		0.189	0.147
	(0.364)	(1.124)	(0.359)	(0.369)		(0.359)	(0.363)
Cost-sharing school (2001) indicator	-1.60*	-1.49*	-1.54*	-1.60*	-1.37	-1.64**	-1.64*
	(0.84)	(0.85)	(0.84)	(0.84)	(0.85)	(0.83)	(0.85)
R^2	0.061	0.061	0.061	0.061	0.060	0.062	0.055
Observations	5,084	5,084	5,084	5,084	5084	5,084	5,084
Mean (s.d.) in control group	18.4	18.4	18.4	18.4	18.4	18.4	18.4
	(23.1)	(23.1)	(23.1)	(23.1)	(23.1)	(23.1)	(23.1)

Notes: For details on the regressions, see the "Notes" for Table 1. Column (1) replicates the main results from Table 2, while columns (2) through (7) add or remove terms from this base specification. Column (2) uses the base specification but replaces saturation with the total number of deworming treatment pupils within 6 km. Column (3) includes interactions of the deworming treatment indicator with demeaned saturation, and column (4) instead includes demeaned saturation squared. Column (5) drops the externality term. Column (6) includes an interaction between the female indicator and treatment, and Column (7) includes an interaction between baseline grade level and treatment.

Table S7: Heterogeneous deworming impacts, full sample

	Number of meals eaten						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Deworming Treatment indicator	0.095***	0.096***	0.085***	0.095***	0.072**	0.125***	0.068*
	(0.029)	(0.028)	(0.028)	(0.028)	(0.028)	(0.041)	(0.036)
Female	0.079***	0.078***	0.078***	0.078***	0.079***	0.122**	0.079***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.052)	(0.026)
Female * Treatment						-0.064	
						(0.059)	
Grades 5-7 in 1998							-0.006
							(0.030)
Grades 5-7 * Treatment							0.059
							(0.041)
Saturation (P_j), demeaned	0.415***		0.037	0.378***		0.417***	0.426***
	(0.124)		(0.246)	(0.101)		(0.125)	(0.126)
Saturation (P_j), demeaned * Treatment			0.465				
			(0.287)				
Saturation (P_j), demeaned and squared				-2.091**			
				(0.826)			
Deworming treatment school students within 6 km (in '000s), demeaned		0.080***					
		(0.023)					
Total primary school students within 6 km (in '000s), demeaned	-0.014	-0.070***	-0.015*	-0.018*		-0.014	-0.015
	(0.010)	(0.018)	(0.009)	(0.010)		(0.009)	(0.009)
Cost-sharing school (2001) indicator	-0.073**	-0.069**	-0.075**	-0.070**	-0.062*	-0.074**	-0.073**
	(0.032)	(0.031)	(0.031)	(0.031)	(0.032)	(0.031)	(0.032)
R^2	0.035	0.035	0.035	0.035	0.033	0.035	0.032
Observations	5,083	5,083	5,083	5,083	5,083	5,083	5,083
Mean (s.d.) in control group	2.16	2.16	2.16	2.16	2.16	2.16	2.16
	(0.64)	(0.64)	(0.64)	(0.64)	(0.64)	(0.64)	(0.64)

Notes: For details, see "Notes" on Table S6.

Table S8: Heterogeneous deworming impacts, full sample

	Passed primary school leaving exam						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Deworming Treatment indicator	0.050	0.048	0.056*	0.050*	0.037	0.050	0.042
	(0.031)	(0.031)	(0.031)	(0.030)	(0.028)	(0.033)	(0.037)
Female	-0.183***	-0.183***	-0.183***	-0.183***	-0.183***	-0.185***	-0.183***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.031)	(0.019)
Female * Treatment						0.002	
						(0.040)	
Grades 5-7 in 1998							0.033
							(0.040)
Grades 5-7 * Treatment							0.019
							(0.044)
Saturation (P_j), demeaned	0.220		0.437*	0.240		0.220	0.235
	(0.161)		(0.252)	(0.151)		(0.161)	(0.160)
Saturation (P_j), demeaned * Treatment			-0.267				
			(0.273)				
Saturation (P_j), demeaned and squared				1.148			
				(1.082)			
Deworming treatment school students within 6 km (in '000s), demeaned		0.032 (0.029)					
Total primary school students within 6 km (in '000s), demeaned	0.007 (0.010)	-0.015 (0.021)	0.007 (0.009)	0.009 (0.010)		0.007 (0.010)	0.006 (0.010)
Cost-sharing school (2001) indicator	-0.039	-0.037	-0.038	-0.041	-0.035	-0.039	-0.041
	(0.027)	(0.026)	(0.027)	(0.027)	(0.026)	(0.027)	(0.027)
R^2	0.071	0.070	0.071	0.071	0.070	0.071	0.067
Observations	4,974	4,974	4,974	4,974	4,974	4,974	4,974
Mean (s.d.) in control group	0.505	0.505	0.505	0.505	0.505	0.505	0.505
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)

Notes: For details, see "Notes" on Table S6.

Table S9: Heterogeneous deworming impacts, full sample

	Ln(Total labor earnings), past month						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Deworming Treatment indicator	0.269***	0.267***	0.206**	0.284***	0.202***	0.233**	0.303**
	(0.085)	(0.089)	(0.081)	(0.085)	(0.072)	(0.097)	(0.144)
Female	-0.440***	-0.439***	-0.436***	-0.442***	-0.428***	-0.515***	-0.451***
	(0.090)	(0.090)	(0.090)	(0.090)	(0.090)	(0.142)	(0.093)
Female * Treatment						0.124	
						(0.191)	
Grades 5-7 in 1998							0.502***
							(0.155)
Grades 5-7 * Treatment							-0.058
							(0.180)
Saturation (P_j), demeaned	1.141		-1.620**	1.103		1.128	1.193
	(0.869)		(0.765)	(0.711)		(0.864)	(0.902)
Saturation (P_j), demeaned * Treatment			3.567***				
			(1.042)				
Saturation (P_j), demeaned and squared				-9.404**			
				(4.241)			
Deworming treatment school students within 6 km (in '000s), demeaned		0.196					
		(0.163)					
Total primary school students within 6 km (in '000s), demeaned	0.036	-0.100	0.027	0.017		0.034	0.036
	(0.026)	(0.124)	(0.026)	(0.026)		(0.026)	(0.026)
Cost-sharing school (2001) indicator	-0.153*	-0.142	-0.173**	-0.147*	-0.123	-0.148*	-0.170**
	(0.087)	(0.087)	(0.084)	(0.083)	(0.093)	(0.086)	(0.086)
R^2	0.196	0.196	0.208	0.203	0.190	0.197	0.182
Observations	710	710	710	710	710	710	710
Mean (s.d.) in control group	7.79	7.79	7.79	7.79	7.79	7.79	7.79
	(0.88)	(0.88)	(0.88)	(0.88)	(0.88)	(0.88)	(0.88)

Notes: For details, see "Notes" on Table S6.

Table S10: Heterogeneous deworming impacts, by gender and helminth type

	Hours worked in last week, all sectors								
	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Deworming Treatment indicator	1.695	1.534	2.048	3.664**	3.726**	3.661*	0.169	-0.557	0.165
	(1.176)	(1.461)	(1.484)	(1.532)	(1.853)	(1.873)	(1.559)	(1.859)	(2.595)
Moderate-heavy worm infection rate at the zonal level (1998), demeaned	4.24			2.45			4.98		
	(6.74)			(10.54)			(9.98)		
Moderate-heavy infection rate * Treatment	2.68			9.59			-0.46		
	(7.45)			(11.29)			(11.84)		
Moderate-heavy geohelmith rate at the zonal level (1998), demeaned		2.39			-6.30			9.90	
		(6.41)			(11.71)			(9.03)	
Moderate-heavy geohelmith rate * Treatment		11.71			22.60*			2.78	
		(7.73)			(13.66)			(11.29)	
Indicator for school within 5 km of lake		-1.42			-0.80			-2.57	
		(1.54)			(2.34)			(2.12)	
Indicator for within 5 km of lake * Treatment		-1.04			-2.52			1.32	
		(1.77)			(2.57)			(2.52)	
R^2	0.059	0.061	0.068	0.052	0.055	0.050	0.053	0.056	0.082
Observations	5,084	5,084	3,652	2,595	2,595	1,893	2,489	2,489	702
Mean (s.d.) in control group	18.43	18.43	18.76	20.32	20.32	20.62	16.32	16.32	16.61
	(23.09)	(23.09)	(23.44)	(24.55)	(24.55)	(25.01)	(21.15)	(21.15)	(21.31)

Notes: For details on the regressions, see the "Notes" for Table 1. Columns (3), (6) and (9) restrict the sample to respondents who live more than 5 km from the lake, which allows us to examine impacts in areas that are likely only infected by geohelminths.

Table S11: Heterogeneous deworming impacts, by gender and helminth type

	Number of meals eaten yesterday								
	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Deworming Treatment indicator	0.074**	0.066*	0.084*	0.102**	0.071	0.045	0.037	0.062	0.181***
	(0.031)	(0.036)	(0.043)	(0.043)	(0.054)	(0.050)	(0.043)	(0.052)	(0.033)
Moderate-heavy worm infection rate at the zonal level (1998), demeaned	-0.433***			-0.564**			-0.191		
	(0.148)			(0.253)			(0.267)		
Moderate-heavy infection rate * Treatment	0.024			0.300			-0.257		
	(0.212)			(0.331)			(0.330)		
Moderate-heavy geohelmith rate at the zonal level (1998), demeaned		-0.356*			-0.239			-0.429	
		(0.191)			(0.367)			(0.287)	
Moderate-heavy geohelmith rate * Treatment		-0.160			-0.199			-0.076	
		(0.271)			(0.435)			(0.396)	
Indicator for school within 5 km of lake		-0.031			-0.096			0.055	
		(0.038)			(0.074)			(0.049)	
Indicator for within 5 km of lake * Treatment		0.039			0.149*			-0.080	
		(0.046)			(0.081)			(0.063)	
R^2	0.029	0.030	0.034	0.041	0.042	0.057	0.028	0.030	0.069
Observations	5,083	5,083	3,651	2,595	2,595	1,893	2,488	2,488	702
Mean (s.d.) in control group	2.162	2.162	2.189	2.103	2.103	2.152	2.229	2.229	2.232
	(0.637)	(0.637)	(0.608)	(0.649)	(0.649)	(0.619)	(0.618)	(0.618)	(0.593)

Notes: For details on the regressions, see the "Notes" for Table 1. Columns (3), (6) and (9) restrict the sample to respondents who live more than 5 km from the lake, which allows us to examine impacts in areas that are likely only infected by geohelminths.

Table S12: Heterogeneous deworming impacts, by gender and helminth type

	Passed secondary school entrance exam during 1998-2007 indicator								
	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Deworming Treatment indicator	0.041	0.029	0.019	-0.009	-0.017	-0.044	0.091**	0.082**	0.101**
	(0.027)	(0.030)	(0.031)	(0.027)	(0.032)	(0.034)	(0.038)	(0.041)	(0.050)
Moderate-heavy worm infection rate at the zonal level (1998), demeaned	-0.311*			-0.690***			0.109		
	(0.171)			(0.204)			(0.198)		
Moderate-heavy infection rate * Treatment	0.368*			0.557**			0.024		
	(0.204)			(0.235)			(0.261)		
Moderate-heavy geohelmith rate at the zonal level (1998), demeaned		-0.139			-0.554**			0.325	
		(0.198)			(0.235)			(0.223)	
Moderate-heavy geohelmith rate * Treatment		-0.076			0.238			-0.474	
		(0.258)			(0.318)			(0.299)	
Indicator for school within 5 km of lake		-0.010			-0.013			0.003	
		(0.044)			(0.050)			(0.048)	
Indicator for within 5 km of lake * Treatment		0.096*			0.099			0.067	
		(0.055)			(0.063)			(0.063)	
R^2	0.069	0.071	0.068	0.036	0.037	0.054	0.069	0.072	0.055
Observations	4,974	4,974	3,568	2,541	2,541	1,852	2,433	2,433	689
Mean (s.d.) in control group	0.505	0.505	0.528	0.590	0.590	0.610	0.409	0.409	0.431
	(0.500)	(0.500)	(0.499)	(0.492)	(0.492)	(0.488)	(0.492)	(0.492)	(0.496)

Notes: For details on the regressions, see the "Notes" for Table 1. Columns (3), (6) and (9) restrict the sample to respondents who live more than 5 km from the lake, which allows us to examine impacts in areas that are likely only infected by geohelminths.

Table S13: Heterogeneous deworming impacts, by gender and helminth type, wage earner subsample

	Ln(Total labor earnings), past month								
	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Deworming Treatment indicator	0.312***	0.276***	0.272**	0.288***	0.255**	0.256*	0.228	0.119	0.448*
	(0.085)	(0.094)	(0.114)	(0.109)	(0.123)	(0.140)	(0.175)	(0.178)	(0.234)
Moderate-heavy worm infection rate at the zonal level (1998), demeaned	0.342			-0.551			2.535***		
	(0.378)			(0.483)			(0.934)		
Moderate-heavy infection rate * Treatment	-0.896			0.066			-3.167**		
	(0.573)			(0.654)			(1.469)		
Moderate-heavy geohelmith rate at the zonal level (1998), demeaned		1.224***			0.339			3.048***	
		(0.433)			(0.601)			(0.940)	
Moderate-heavy geohelmith rate * Treatment		-1.445**			-0.074			-5.518***	
		(0.700)			(0.902)			(1.578)	
Indicator for school within 5 km of lake		-0.166			-0.167			-0.103	
		(0.108)			(0.132)			(0.229)	
Indicator for within 5 km of lake * Treatment		0.015			-0.107			0.749***	
		(0.152)			(0.174)			(0.276)	
R^2	0.188	0.194	0.244	0.166	0.175	0.218	0.256	0.301	0.261
Observations	710	710	508	542	542	383	168	168	159
Mean (s.d.) in control group	7.794	7.794	7.829	7.923	7.923	7.983	7.459	7.459	7.481
	(0.878)	(0.878)	(0.918)	(0.873)	(0.873)	(0.911)	(0.806)	(0.806)	(0.842)

Notes: For details on the regressions, see the "Notes" for Table 1. Columns (3), (6) and (9) restrict the sample to respondents who live more than 5 km from the lake, which allows us to examine impacts in areas that are likely only infected by geohelminths.

Table S14: Baseline (1998) summary statistics across treatment groups, for wage-earner subsample

	Treatment – Control (s.e.)			Control group mean (s.d.)		
	All	Male	Female	All	Male	Female
Age (1998)	-0.28	-0.13	-1.09***	13.4	13.6	12.9
	(0.27)	(0.32)	(0.42)	(2.5)	(2.7)	(1.9)
Grade (1998)	-0.05	-0.03	-0.16	4.91	4.93	4.85
	(0.14)	(0.17)	(0.31)	(1.57)	(1.59)	(1.52)
Female	-0.071			0.280		
	(0.045)					
School average test score (1996)	-0.050	-0.020	-0.122	0.024	-0.010	0.111
	(0.106)	(0.103)	(0.138)	(0.391)	(0.357)	(0.460)
Primary school located in Budalangi division	0.052	0.026	0.115	0.378	0.405	0.310
	(0.144)	(0.149)	(0.156)			
Population of primary school	78	72	94	425	432	407
	(56)	(54)	(69)	(136)	(141)	(120)
Total primary school students within 6 km	0	-75	250	4,730	4,759	4,655
	(420)	(382)	(633)	(1,598)	(1,495)	(1,846)
Primary school students within 6 km who are treatment (Group 1,2) pupils	-268	-324	-63	3,382	3,390	3,363
	(282)	(254)	(420)	(1,064)	(987)	(1,250)
Saturation (P_j)	-0.044***	-0.044***	-0.041**	0.542	0.539	0.548
	(0.015)	(0.016)	(0.019)	(0.055)	(0.052)	(0.062)
Years of assigned deworming treatment, 1998-2003	2.32***	2.28***	2.46***	1.23	1.23	1.24
	(0.14)	(0.17)	(0.24)	(1.23)	(1.25)	(1.16)

Notes: Data is from the PSDP, and includes individuals surveyed in KLPS2 who were working for wages at the time of survey. N=718 observations, with 549 males and 169 females (except for age, where N=717). Years of assigned deworming treatment is calculated using the treatment group of the respondent's school and their grade, but is not adjusted for the treatment ineligibility of females over age 13 or assignment to cost-sharing in 2001. Respondents who "age out" of primary school are no longer considered assigned to treatment. School average test scores are from the 1996 Busia District mock exam, and are converted to normalized individual standard deviation units. Observations are weighted to maintain initial population proportions. The "Treatment – Control" differences are derived from a linear regression on a constant and the treatment indicator. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence.

Table S15: Hours Worked Decomposition – Extensive versus Intensive Margins

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Total Hours in All Sectors	All	Male	Female	All	All	Male	Female
Hours Worked	1.578	3.494**	0.319	10.197	18.4	20.3	16.3
	(1.040)	(1.424)	(1.358)	(7.803)	(23.1)	(24.6)	(21.1)
Indicator for hours > 0	0.002	0.047*	-0.035	-0.016	0.68	0.68	0.68
	(0.022)	(0.027)	(0.031)	(0.115)	(0.47)	(0.47)	(0.47)
Hours worked, among those with hours > 0	2.282*	3.683**	1.534	16.622*	27.0	29.8	24.0
	(1.222)	(1.588)	(1.554)	(9.414)	(23.4)	(24.5)	(21.8)
					3,579	*1,898*	*1,681*
Panel B: Agriculture							
Hours Worked	-0.066	1.026*	-1.274**	-0.549	8.3	7.8	8.8
	(0.415)	(0.551)	(0.555)	(3.412)	(11.4)	(11.6)	(11.2)
Indicator for hours > 0	-0.011	0.018	-0.039	0.123	0.55	0.53	0.58
	(0.025)	(0.028)	(0.032)	(0.142)	(0.50)	(0.50)	(0.49)
Hours worked, among those with hours > 0	-0.015	1.337	-1.333	-4.961	14.9	14.8	15.1
	(0.608)	(0.859)	(0.830)	(5.334)	(11.7)	(12.3)	(11.1)
					2,916	*1,454*	*1,462*
Panel C: Wage Employment							
Hours Worked	0.138	1.114	-0.269	4.745	6.9	8.8	4.8
	(0.839)	(1.320)	(1.076)	(5.065)	(18.5)	(20.0)	(16.5)
Indicator for hours > 0	-0.013	-0.006	-0.009	0.022	0.15	0.20	0.09
	(0.017)	(0.027)	(0.018)	(0.087)	(0.36)	(0.40)	(0.29)
Hours worked, among those with hours > 0	4.943*	6.149*	-1.242	33.351*	46.5	43.7	53.6
	(2.773)	(3.387)	(3.667)	(17.896)	(21.7)	(21.7)	(20.2)
					625	*470*	*155*
Panel D: Self-Employment (non-agricultural)							
Hours Worked	1.506***	1.355*	1.862**	6.001*	3.3	3.8	2.7
	(0.548)	(0.725)	(0.811)	(3.231)	(12.8)	(13.7)	(11.7)
Indicator for hours > 0	0.022*	0.020	0.028	0.022	0.09	0.09	0.08
	(0.011)	(0.015)	(0.017)	(0.076)	(0.28)	(0.29)	(0.27)
Hours worked, among those with hours > 0	5.742**	5.864	5.617*	36.505**	38.1	40.2	35.3
	(2.903)	(4.452)	(2.926)	(14.765)	(24.0)	(23.1)	(25.1)
					542	*288*	*254*

Notes: Each entry is from a separate OLS regression. For details on the regressions, see the "Notes" for Table 1. For the "Hours Worked" and "Indicator for hours > 0" rows, the sample sizes are 5,084 for "All", 2,595 for "Males", and 2,489 for "Females". Hours worked are in the last week.

Table S16: Deworming impacts on occupation, within the wage earner subsample

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean; *Number of Observations*		
	All	Male	Female	All	All	Male	Female
Agriculture	-0.011	-0.008	0.002	-0.031	0.021	0.008	0.052
	(0.012)	(0.014)	(0.031)	(0.143)	*706*	*540*	*166*
Casual/Construction laborer	-0.040**	-0.019	-0.080	-0.148	0.029	0.018	0.059
	(0.019)	(0.014)	(0.052)	(0.112)	*706*	*540*	*166*
Fishing	-0.014	0.029	-0.052*	-0.619	0.192	0.242	0.064
	(0.059)	(0.065)	(0.031)	(0.479)	*706*	*540*	*166*
Manufacturing	0.076***	0.095***	0.059	0.306**	0.030	0.031	0.028
	(0.024)	(0.034)	(0.047)	(0.150)	*706*	*540*	*166*
Retail and wholesale trade	0.002	-0.030	0.080	0.217	0.182	0.190	0.160
	(0.045)	(0.049)	(0.080)	(0.232)	*706*	*540*	*166*
Services (all)	0.030	0.006	-0.015	0.174	0.423	0.341	0.633
	(0.055)	(0.055)	(0.094)	(0.413)	*706*	*540*	*166*
Domestic	-0.012	0.019	-0.160	-0.166	0.117	0.030	0.340
	(0.032)	(0.019)	(0.109)	(0.202)	*706*	*540*	*166*
Restaurants, cafes, etc.	-0.034	-0.020	-0.054	0.057	0.061	0.042	0.110
	(0.025)	(0.027)	(0.043)	(0.197)	*706*	*540*	*166*
Trade contractors	-0.015	-0.030	-0.002	0.168	0.093	0.128	0.004
	(0.028)	(0.040)	(0.010)	(0.249)	*706*	*540*	*166*

Notes: The sample includes all individuals surveyed in the KLPS2 who report working for pay (with earnings greater than zero) at the time of the survey. Each entry is from a separate OLS regression. For details on the regressions, see the "Notes" for Table 1.

Table S17: Average characteristics of occupations within wage employment

	Mean (s.d.) in Control Group		
	Hours per week worked in sector	Days of work lost to poor health*	Earnings in sector, past month (KSh)
Agriculture	13	2.1	618
	(12)	(1.9)	(258)
Casual/Construction laborer	51	0.4	2,246
	(31)	(1.0)	(1,576)
Fishing	37	2.1	3,017
	(25)	(4.2)	(1,704)
Manufacturing	53	1.1	4,916
	(24)	(1.8)	(2,401)
Retail and wholesale trade	40	0.9	2,126
	(27)	(2.0)	(1,757)
Services (all)	49	1.3	4,398
	(22)	(2.6)	(4,905)
Domestic	61	1.5	2,538
	(17)	(2.5)	(1,558)
Restaurants, cafes, etc.	53	1.2	3,694
	(21)	(2.5)	(3,037)
Trade contractors	27	0.8	3,059
	(22)	(2.5)	(1,980)

Notes: The sample includes all individuals surveyed in the KLPS2 who report working for pay (with earnings greater than zero) at the time of the survey. All observations are weighted to maintain initial population proportions. [a] Note that we only have days of work missed in total, not separated by sector, so among those who work in multiple sectors, there is some overlap.

Table S18: Deworming impacts on health (with multiple testing adjustments)

	Coefficient estimate (s.e.) on deworming treatment indicator		
	All	Male	Female
Self-reported health "very good" indicator at KLPS-2	0.040**	0.023	0.051**
	(0.018)	(0.025)	(0.025)
	[0.122]	[1.000]	[0.075]
Height at KLPS-2	-0.109	0.072	-0.301
	(0.271)	(0.382)	(0.387)
	[0.524]	[1.000]	[0.282]
Body mass index (BMI) at KLPS-2	0.022	-0.012	0.058
	(0.045)	(0.060)	(0.066)
	[0.524]	[1.000]	[0.282]
Miscarriage indicator (obs. at pregnancy level) at KLPS-2	-0.015*	0.000	-0.028**
(for females – themselves; for males – their partners)	(0.008)	(0.004)	(0.013)
	[0.122]	[1.000]	[0.073]

Notes: The analysis is the same as Table 1, with the addition of the multiple testing adjusted FDR q-values in square brackets, as described in the text. The terms in regular parentheses are standard errors. The "Moderate-heavy worm infections in 2001" outcome is not included in the multiple testing adjustment since it is not part of KLPS-2 data collection.

Table S19: Deworming impacts on education (with multiple testing adjustments)

	Coefficient estimate (s.e.) on deworming treatment indicator		
	All	Male	Female
Total years enrolled in school, 1998-2007	0.294**	0.150	0.354**
	(0.145)	(0.166)	(0.179)
	[0.102]	[0.852]	[0.095]
Total years enrolled in primary school, 1998-2007	0.155**	0.238**	0.026
	(0.075)	(0.102)	(0.098)
	[0.102]	[0.071]	[0.294]
Repetition of at least one grade (1998-2007) indicator	0.063***	0.072***	0.053*
	(0.018)	(0.025)	(0.030)
	[0.008]	[0.037]	[0.121]
Grades of schooling attained by 2007	0.150	-0.030	0.261
	(0.143)	(0.148)	(0.171)
	[0.226]	[1.000]	[0.139]
Attended secondary school by 2007 indicator	0.030	-0.035	0.090**
	(0.035)	(0.038)	(0.038)
	[0.251]	[0.852]	[0.084]
Passed secondary school entrance exam during 1998-2007 indicator	0.050	0.004	0.096**
	(0.031)	(0.030)	(0.040)
	[0.121]	[1.000]	[0.084]
Out-of-school (at 2007-09 survey) indicator	-0.006	0.022	-0.029
	(0.022)	(0.030)	(0.026)
	[0.512]	[0.852]	[0.178]

Notes: The analysis is the same as Table 2, with the addition of the multiple testing adjusted FDR q-values in square brackets, as described in the text. The terms in regular parentheses are standard errors.

Table S20: Deworming impacts on labor hours and occupational choice (with multiple testing adjustments)

	Coefficient estimate (s.e.) on deworming treatment indicator		
Panel A: Hours worked	All	Male	Female
Hours worked in all sectors in last week, full sample	1.58	3.49**	0.32
	(1.04)	(1.42)	(1.36)
	[0.098]	[0.071]	[0.373]
Hours worked in all sectors in last week, out-of-school sample	2.93**	4.55**	2.14
	(1.29)	(1.95)	(1.49)
	[0.065]	[0.071]	[0.198]
Panel B: Sectoral time allocation			
Hours worked in non-agricultural self-employment in last week, full sample	1.51***	1.35*	1.86**
	(0.55)	(0.73)	(0.81)
	[0.038]	[0.088]	[0.127]
Hours worked in agriculture in last week, full sample	-0.07	1.03*	-1.27**
	(0.42)	(0.55)	(0.56)
	[0.412]	[0.088]	[0.127]
Hours worked in wage earning in last week, full sample	0.14	1.11	-0.27
	(0.84)	(1.32)	(1.08)
	[0.412]	[0.299]	[0.373]
Panel C: Occupational choice (full sample)			
Manufacturing job indicator	0.011***	0.019**	0.005
	(0.004)	(0.008)	(0.004)
	[0.038]	[0.071]	[0.198]
Construction/casual labor job indicator	-0.005**	-0.003	-0.007
	(0.003)	(0.003)	(0.004)
	[0.071]	[0.244]	[0.198]
Domestic service job indicator	-0.005	0.002	-0.013
	(0.006)	(0.004)	(0.013)
	[0.250]	[0.431]	[0.329]
Grows cash crop indicator	0.010**	0.003	0.019**
	(0.005)	(0.004)	(0.009)
	[0.071]	[0.303]	[0.127]

Notes: The analysis is the same as Table 3, with the addition of the multiple testing adjusted FDR q-values in square brackets, as described in the text. The terms in regular parentheses are standard errors.

Table S21: Deworming impacts on living standards and labor earnings (with multiple testing adjustments)

	Coefficient estimate (s.e.) on deworming treatment indicator		
Panel A: Consumption	All	Male	Female
Number of meals eaten yesterday, full sample	0.095***	0.125***	0.051
	(0.029)	(0.041)	(0.043)
	[0.005]	[0.011]	[1.000]
Number of meals eaten yesterday, out-of-school sample	0.102***	0.158***	0.037
	(0.029)	(0.046)	(0.044)
	[0.005]	[0.009]	[1.000]
Panel B: Wage earnings (among wage earners)			
Ln(Total labor earnings), past month	0.269***	0.244**	0.165
	(0.085)	(0.109)	(0.175)
	[0.005]	[0.044]	[1.000]
Ln(Wage = Total labor earnings / hours), past month, if ≥10 hours per week of work	0.197*	0.181	0.225
	(0.102)	(0.128)	(0.194)
	[0.041]	[0.150]	[1.000]
Ln(Total labor earnings), most recent month worked since 2007	0.225***	0.221**	0.178*
	(0.070)	(0.097)	(0.104)
	[0.005]	[0.044]	[1.000]
Panel C: Non-agricultural self-employment outcomes (among non-agricultural self-employed)			
Total self-employed profits (self-reported) past month	384	111	250
	(308)	(465)	(265)
	[0.084]	[0.439]	[1.000]
Total self-employed profits past month, top 5% trimmed	341*	259	80
	(177)	(309)	(219)
	[0.041]	[0.353]	[1.000]
Total employees hired (excluding self)	0.416	0.245	0.603
	(0.361)	(0.403)	(1.275)
	[0.084]	[0.353]	[1.000]

Notes: The analysis is the same as Table 4, with the addition of the multiple testing adjusted FDR q-values in square brackets, as described in the text. The terms in regular parentheses are standard errors.

Appendix Table S22: Fiscal Impacts of Deworming Subsidies, Using Wage Earnings and Self-employed Profits

	No Subsidy	Partial Subsidy	Full Subsidy	*Notes*
Panel A: Calibration Parameters				
Mean per person increase in earnings/month: μ	$0.00	$0.39	$1.54	Treatment effect for total labor earnings + self-employed profits, past month (=0 for non-earners) (Table 4).
Panel B: No health spillovers				
Annual increase in per-person earnings	$0.00	$4.68	$18.49	μ x 12
NPV increase in per-person earnings (relative to no subsidy)	-	$43.21	$170.56	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	$4.45	$17.56	NPV earnings x 16.5% tax rate under no subsidy - mean schooling costs

Notes: The construction of the data and the calibration parameters are the same as Table 5 (see note), except we use the mean increase in total labor earnings plus self-employment profits in the last month, which is equal to US$1.54. The NPV of lifetime earnings in the no subsidy and no health spillovers case is $1,120.60. The social pecuniary internal rate of return (annualized) in the case of no epidemiological spillovers is 34.1%, and with epidemiological spillovers is 40.4%. Calculations are available upon request.

Appendix Table S23: Fiscal Impacts of Deworming Subsidies, with No Life Cycle Earnings Adjustment or Productivity Growth

	No Subsidy	Partial Subsidy	Full Subsidy	*Notes*
Panel A: Calibration Parameters (same as Table 2)				
Panel B: No health spillovers				
Annual increase in per-person earnings	\$0.00	\$3.91	\$15.44	λ_1 x starting wage * 52
NPV increase in per-person earnings (relative to no subsidy)	-	\$19.36	\$76.43	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	\$0.50	\$1.96	NPV earnings x 16.5% tax rate under no subsidy - direct schooling costs
Panel C: With health spillovers				
Annual increase in per-person earnings	\$0.00	\$26.77	\$83.11	(λ_1 + (p/R) λ_2) x starting wage x 52
NPV increase in per-person earnings (relative to no subsidy)	-	\$113.17	\$411.48	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	\$12.65	\$44.07	NPV earnings x 16.5% tax rate under no subsidy - (direct + externality costs of schooling)

Notes: The construction of the data and the calibration parameters is the same as Table 5 (see note), except we assume no growth in wages over time (from either life cycle adjustments or productivity growth). The NPV of per-person lifetime earnings in the no subsidy and no health spillovers case is \$810.26. The social pecuniary internal rate of return (annualized) in the case of no epidemiological spillovers is 28.2%, and with epidemiological spillovers is 48.2%. Calculations are available upon request.

Appendix Table S24: Fiscal Impacts of Deworming Subsidies, Including Female Education Gains and Opportunity Costs

	No Subsidy	Partial Subsidy	Full Subsidy	*Notes*
Panel A: Calibration Parameters				
Mean per person increase in earnings due to educational gains	0.00%	0.27%	1.06%	Men: no increase. Females: Assume 6% return to an additional year of education (Duflo 2001) x 0.354 additional years of education (Table 2, Panel B).
Mean per person increase in earnings due to educational gains from externality	0.00%	0.31%	1.22%	Men: no increase. Females: Assume 6% return to an additional year of education (Duflo 2001) x 0.408 additional years of education from externality (Table 2, Panel B).
Opportunity costs of additional schooling for females	0.00	4.01	15.84	Mean starting wage scaled by age x 16.3 hours worked by control group females per week (Table 3) x 37 weeks of school per year
Opportunity costs of additional schooling for females from externality	0.00	16.25	64.14	Mean starting wage scaled by age x 16.3 hours worked by control group females per week (Table 3) x 37 weeks of school per year
Panel B: No health spillovers				
NPV increase in per-person earnings (relative to no subsidy)	-	$34.71	$137.03	9.85% Annual (real) interest rate in Kenya - opportunity costs
NPV increase in per-person government revenue	-	$3.04	$12.00	NPV earnings x 16.5% tax rate under no subsidy - direct schooling costs
Panel C: With health spillovers				
NPV increase in per-person earnings (relative to no subsidy)	-	$236.43	$725.16	9.85% Annual (real) interest rate in Kenya
NPV increase in per-person government revenue	-	$33.08	$96.07	NPV earnings x 16.5% tax rate under no subsidy - (direct + externality costs of schooling)

Notes: The construction of the data and the calibration parameters is the same as Table 5 (see note), except we assume females gain from their additional years of education, earning a return of 6% per additional year (following Duflo 2001). We also include the opportunity cost of the extra time females spend in school, scaled linearly upward by age from 0% at age 8 to 100% of the adult wage at age 18. The NPV of per-person lifetime earnings in the no subsidy and no health spillovers case is $1,509.96. The social pecuniary internal rate of return (annualized) in the case of no health spillovers is 27.2%, and with epidemiological spillovers is 35.3%. Calculations are available upon request.

D. References

Aiken AM, Davey C, Hargreaves JR, Hayes RJ. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication. *Int J Epidemiol*. 2015.

Anderson, M.L. (2008). "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Re-evaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *JASA,* 103(484): 1481-1495.

Anderson, R.M. and May, R.M. (1991). *Infectious diseases of humans: Dynamics and control.* Oxford University Press, London.

Benjamini, Y., A. Krieger, and D. Yekutieli. (2006). "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate", *Biometrika*, 16.

Brooker, S. E. Miguel, S. Moulin, A.I. Luoba, D.A.P. Bundy, and M. Kremer. (2000). "Epidemiology of single and multiple species of helminth infections among school children in Busia District, Kenya", *East African Medical Journal*, 77(3), 157-161.

Bruhn, M., and D. McKenzie. (2009). "In pursuit of balance: Randomization in practice in development field experiments", *American Economic Journal: Applied Economics*, 1(4).

Bundy, D.A.P., and E.S. Cooper. (1989). "Trichuris", in *Tropical and Geographical Medicine*, eds. K.S. Warren and A.A.F. Mahmoud. New York: McGraw-Hill.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. (2012). "Reshaping institutions: Evidence on aid impacts using a pre-analysis plan", *Quarterly Journal of Economics*, 127(4), 1755-1812.

Chattopadhyay, R., and E. Duflo. (2004). "Women as policymakers: Evidence from a randomized policy experiment in India", *Econometrica*, 75(2), 1409-1443.

Davey C, Aiken AM, Hayes RJ, Hargreaves JR. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a statistical replication of a cluster quasi-randomized stepped-wedge trial. *Int J Epidemiol*. 2015.

de Mel, S., D. McKenzie, and C. Woodruff. (2009). "Measuring microenterprise profits: Must we ask how the sausage is made?", *Journal of Development Economics*, 88(1): 19-31.

Deaton, A. (2010). "Instruments, Randomization and Learning about Development", *Journal of Economic Literature*, 48, 424-455.

Eble, Alex, Peter Boone, and Diana Elbourne. (2013). "Risk and evidence of bias in randomized controlled trials in Economics", *CEP Discussion Paper No. 1240.*

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph Newhouse, Heidi Allen, Katherine Baicker, and the Oregon Health Study Group. (2012). "The Oregon Health Insurance Experiment: Evidence from the First Year", *Quarterly Journal of Economics*, 127(3), 1057-1106.

Glewwe, P., M. Kremer, and S. Moulin. (2009). "Many Children Left Behind? Textbooks and Test Scores in Kenya", *American Economic Journal: Applied Economics*, 1(1): 112-135.

Heckman, J., J. Stixrud, & S. Urzua. (2006). "The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*.

Hicks JH, Kremer M, Miguel E. Commentary: Deworming externalities and schooling impacts in Kenya: a comment on Aiken et al. (2015) and Davey et al. (2015). *Int J Epidemiol*. (2015).

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. (2007). “Experimental Analysis of Neighborhood Effects.” *Econometrica*, 75 (1) 83-119.

Kremer, Michael, and Edward Miguel. (2007). “The Illusion of Sustainability”, *Quarterly Journal of Economics*, 112(3), 1007-1065.

Kremer, Michael, Jean Lee, Jonathan Robinson, and Olga Rostapshova (2013). "Behavioral Biases and Firm Behavior: Evidence from Kenyan Retail Shops," *American Economic Review Papers and Proceedings* 103:3 (May 2013): 362-368.

Miguel, Edward, and Michael Kremer (2014). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Guide to Replication of Miguel and Kremer (2004)”, *CEGA Working Paper #39.*

Miguel, Edward, Michael Kremer, Joan Hamory Hicks, and Carolyne Nekesa (2014). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Data User’s Guide”, *CEGA Working Paper #40.*

Oaxaca, Roger. (1973). “Male-Female Wage Differentials in Urban Labor Markets”, *International Economic Review*, 14(3), 693-709.

Orr, Larry, et al. (2003). *Moving to Opportunity: Interim Impacts Evaluation.* Washington D.C.: U.S. Dept. of HUD.

Ozier, Owen. (2014). “Exploiting Externalities to Estimate the Long-term Benefits of Early Childhood Deworming”, *World Bank Working Paper Series 7052*.

Suri, Tavneet. (2011). “Selection and Comparative Advantage in Technology Adoption”, *Econometrica*, 79(1): 159-209.

Taylor-Robinson DC, Maayan, N, Soares-Weiser, K. Garner, P. (2012) “Deworming drugs for treating soil-transmitted intestinal worms in children: effects on nutrition and school performance.” *Cochrane Database of Systematic Reviews*.

Udry, Christopher. (1996). “Gender, Agricultural Productivity, and the Theory of the Household”, *Journal of Political Economy*, 104, 1010-1046.

Wanjala, Bernadette (2007). “Design and Performance of Kenya’s Tax System: An Inequality Perspective,” in *Readings on Inequality in Kenya: Sectoral Dynamics and Perspectives*. Nairobi: Society for International Development (SID).