

# WORMS AT WORK: LONG-RUN IMPACTS OF A CHILD HEALTH INVESTMENT\*

SARAH BAIRD  
JOAN HAMORY HICKS  
MICHAEL KREMER  
EDWARD MIGUEL

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 non-agricultural self-employment, are more likely to hold manufacturing jobs, and miss one fewer meal per week. Women who were in treatment schools as girls are approximately one quarter more likely to have attended secondary school,

\*We thank Kevin Audi, Pierre Bachas, Chris Blattman, Seth Blumberg, Hana Brown, Lorenzo Casaburi, Lisa Chen, 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; seminar participants at UC 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, the International Health Economics Association conference; and the editor and four anonymous referees 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. Michael Kremer declares that he works with USAID, which supports deworming, and was formerly a board member of Deworm the World, a 501(c)3 organization. The content is solely the responsibility of the authors and does not necessarily reflect the views of any of our funders.

© The Author(s) 2016. Published by Oxford University Press, on behalf of President and Fellows of Harvard College. All rights reserved. For Permissions, please email: journals.permissions@oup.com

*The Quarterly Journal of Economics* (2016), 1637–1680. doi:10.1093/qje/qjw022.  
Advance Access publication on July 19, 2016.

halving the gender gap. They reallocate time from traditional agriculture into cash crops and nonagricultural self-employment. We estimate a conservative 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. *JEL Codes:* I10, I20, J24, O15.

## I. INTRODUCTION

The question of whether—and how much—child health gains affect adult outcomes is of major research interest across disciplines and is of 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 labor. Pitt, Rosenzweig, and Hassan (2012) (PRH) further note that time allocation will also depend on how the labor market values increased human capital and improved raw labor capacity, and 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 they provide evidence consistent with this model.

We examine the case of intestinal worms, which globally affect approximately 2 billion people according to the World Health Organization (WHO 2014). Worms (helminths) are spread when fecal matter containing eggs from infected individuals is deposited in the local environment. Intense infections lead to lethargy, anemia, and growth stunting (Stephenson et al. 1993; Stoltzfus et al. 1997; Guyatt et al. 2001; Silva et al. 2003) and may weaken the immunological response to other infections (Kjetland et al. 2006; Kirwan et al. 2010). 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 (Larocque et al. 2006; Hotez 2009).

There is ongoing debate about whether it is appropriate to carry out mass deworming treatment programs in endemic regions. Because treatment is safe and cheap but diagnosis is expensive, WHO recommends periodic mass school-based deworming in high-prevalence areas (WHO 1992). Several other groups also highlight deworming as a cost-effective investment (Disease Control Priorities Project 2008; Hall and Horton 2008; Jameel Poverty Action Lab 2012; Givewell 2013). In contrast, a recent Cochrane Review argues that although 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).<sup>1</sup>

Because of its selection criteria focusing on medical-style randomized control trials, the Cochrane Review includes numerous studies subject to well-known methodological limitations (Bundy et al. 2009) and excludes rigorous social science evidence. For instance, the review excludes Bleakley (2007), which estimates the community-wide impact of deworming in the early twentieth-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%.<sup>2</sup>

The present article 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

1. The Cochrane Reviews are systematic reviews of primary research in human health care and health policy. They are influential among health policy makers.

2. A small body of social science research studies the impact of deworming on labor outcomes. In addition 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, whereas 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 later.

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 one to two 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 communities showed cognitive gains in later tests (Ozier 2014).

As discussed in Miguel and Kremer (2014), the original Miguel and Kremer (2004) paper contained several rounding errors and a coding error in the estimation of cross-school externalities. Correcting this coding error indicates that short-run, one-year epidemiological externalities extend out to 3 km or 4 km, rather than 6 km (Aiken et al. 2015; Clemens and Sandefur 2015; Miguel, Kremer, and Hicks 2015). This coding error has been resolved in the current article. Davey et al. (2015) express concerns that there were differences across schools in the number of visits to measure school attendance in Miguel and Kremer (2004). As noted in Hicks, Kremer, and Miguel (2015), there is no statistical evidence for any imbalance in data collection patterns across treatment and control schools, and the Miguel and Kremer (2004) results are robust to weighting each individual equally in the analysis. This issue is not relevant to the current article, which employs a different data set than Miguel and Kremer (2004).

In the current analysis, we examine health, education, and labor market outcomes a decade later, at which point most subjects were young adults 19–26 years of age. We find improvements in self-reported health, but not in height. 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 nonagricultural self-employment, and are

more likely to hold manufacturing jobs with higher wage earnings. Their living standards improve as well, with men 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 approximately one quarter more likely to have passed the secondary school entrance exam and to have attended secondary school. They reallocate time from traditional agriculture to non-agricultural self-employment and are also more likely to grow cash crops. Estimated effects on labor hours and living standards are larger for those who were older than 12 years of age (the median age) at baseline, who are much more likely to be out of school by the follow-up survey.

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 employed in that paper, and we 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.

Last, 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 32%, a high return.

The rest of the article is organized as follows. Section II discusses the Kenyan context, the deworming project, and the data. Section III presents the estimation strategy. Section IV discusses the main results. Section V 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. All supplementary material is in the Online Appendix.

## II. DATA

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

### *II.A. 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 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 15–18 years of age, is a key time of labor market transition. Secondary education in Kenya, like tertiary education in the United States, 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 sex, 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.

### *II.B. The Primary School Deworming Project (PSDP)*

In 1998 the nongovernmental 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 a third of the sample had moderate to 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 subunit (zone), zones were listed alphabetically within each geographic division, and schools were then listed

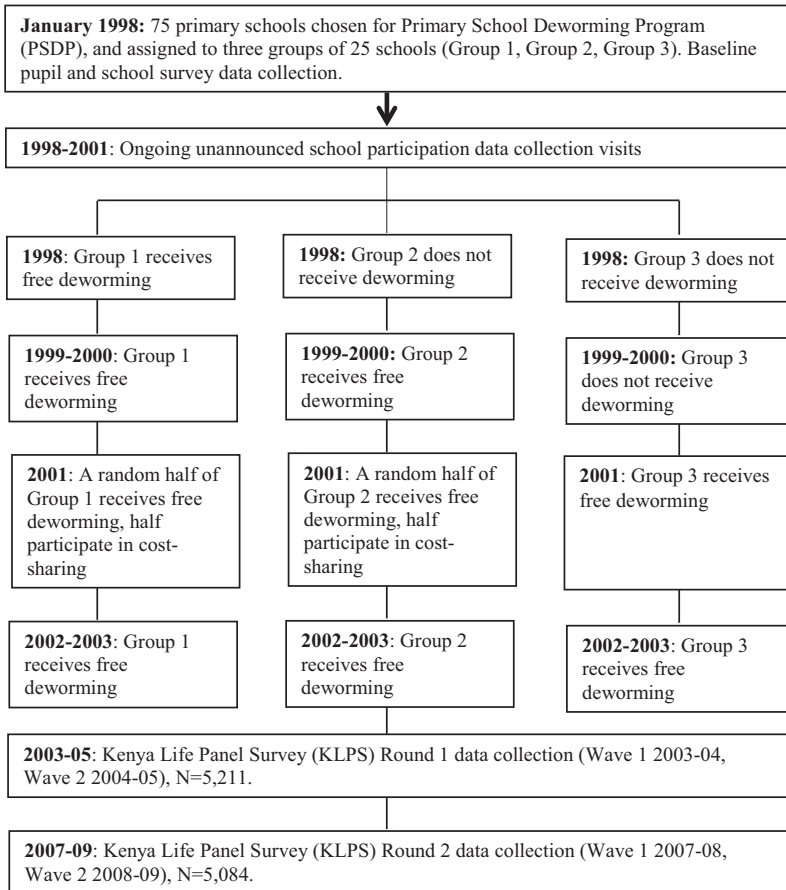


FIGURE I

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

in order of pupil enrollment within each zone, with every third school assigned to a given program group. Figure I presents the project research design and describes the timing of data collection. Online Appendix section A contains a detailed description of the experimental design, provides further information on the sample, and shows that the three groups were well balanced along baseline characteristics (Online 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 (Online 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).

### *II.C. 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 82.5%, and 83.9% among those still alive (see Online 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 (Orr et al. 2003; Kling, Liebman, and Katz 2007).

These are high tracking rates for any age group over a decade, 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 (Online Appendix Table S2).

## III. ESTIMATION STRATEGY

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

### *III.A. 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. Therefore we 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  $\delta$  km in a single year, for instance, due to the natural movement of and interaction among the local population. Miguel and Kremer (2004, 2014) and Hicks, Kremer, and Miguel (2015) estimate substantial and significant short-run (after one year) cross-school externalities on worm infections within 3 km of treatment schools.

Consider an outcome  $Y_{ijt}$  for individual  $i$  in school  $j$  at time  $t$ , for example, 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  $\delta$  km 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”,  $R_{j,t-1,\delta}$ —that is, 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.<sup>3</sup> For now we focus on saturation, which is the epidemiologically relevant quantity, and we return to the distinction between saturation and coverage in the empirical implementation later.

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:

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

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, that is,  $\pi_t(1)/(\text{Cost of } P_{j,t-1,\delta}=1)$  and  $\pi_t(p)/(\text{Cost of } P_{j,t-1,\delta}=p)$ .

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  $y_{0,t} \equiv E[Y_{ijt}|T_{j,t-1}=0, P_{j,t-1,\delta}=0]$  and define the difference in expected outcomes between treated and untreated communities at a given local treatment saturation proportion  $p$  as:

$$(2) \quad \lambda_{1t}(p) \equiv E[Y_{ijt}|T_{j,t-1}=1, P_{j,t-1,\delta}=p] - E[Y_{ijt}|T_{j,t-1}=0, P_{j,t-1,\delta}=p].$$

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

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

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 nondecreasing in worm load in that location and neighboring areas within distance  $\tilde{\delta}$  at lagged time  $t - \tilde{t}$ . Both own and

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

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  $p$ ,

$$\begin{aligned} E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-\tilde{t},\delta} = p] &\geq E[Y_{ijt}|T_{j,t-1} = 0, P_{j,t-\tilde{t},\delta} = p], \\ \text{then for any two levels of local treatment saturation } p'' > p', \\ E[Y_{ijt}|T_{j,t-1} = \mu, P_{j,t-\tilde{t},\delta} = p''] &\geq E[Y_{ijt}|T_{j,t-1} = \mu, P_{j,t-\tilde{t},\delta} = p'] \\ \text{for all } \mu \in \{0, 1\}. \end{aligned}$$

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  $p$ ,  $E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-1,\delta} = p] \geq E[Y_{ijt}|T_{j,t-1} = 0, P_{j,t-1,\delta} = p]$ , 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  $\pi_t(p'') - \pi_t(p') = (E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-1,\delta} = p''] - y_{0,t}) - (E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-1,\delta} = p'] - y_{0,t}) = E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-1,\delta} = p''] - E[Y_{ijt}|T_{j,t-1} = 1, P_{j,t-1,\delta} = p']$ . This is greater than or equal to 0 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  $\lambda_{2t}(p) \equiv E[Y_{ijt}|T_{j,t-1} = 0, P_{j,t-1,\delta} = p] - E[Y_{ijt}|T_{j,t-1} = 0, P_{j,t-1,\delta} = 0] \geq 0$ . 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 expression, which is effectively a weighted average across different saturation proportions  $p$  in this set.

The foregoing discussion abstracts away from other covariates. As we discuss later, their inclusion in a regression analysis is important given the nature of the experimental design and stratified sampling and 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 Online Appendix and in Miguel and Kremer (2004) that the local numbers of all primary school pupils and 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 any potential bias in the coefficient estimate on the treatment school indicator 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 reinfection with worms, and so on.

Denote worm prevalence at location  $j$  at time  $t$  by  $\omega_{jt}$ . Given the geographic spread of worm infections by  $\delta$  km per year,  $\omega_{jt}$  will be a nondecreasing 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 ( $= 10 \text{ years} \times 3 \text{ km a year}$ ) and perhaps beyond. Of course, these effects may fade over time, but no school in our

study area of roughly  $15 \text{ km} \times 40 \text{ km}$  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 article, as the spatial and temporal variation in our data do not allow us to precisely estimate the wide range of potentially relevant parameters, but in Online 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, because there were never more than 12 schools within 4 km, but does lead to less precisely estimated overall effects between 3 and 6 km from a school; Miguel and Kremer (2014) and Ahuja et al. (2015) contain a complete discussion of the updated empirical results. We consider cross-school externalities up to 6 km in the analysis in this article for two reasons. First, spillover effects are likely to diffuse spatially over time, as discussed already. 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 0 in the 3–6 km range for the outcomes we consider ( $p$ -value  $< .001$ ), indicating that their inclusion is appropriate (see Online 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 later.

### III.B. Estimation

The econometric approach relies on the PSDP's prospective experimental design, namely, that the program exogenously provided individuals in treatment (Groups 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 men and women 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).<sup>4</sup>

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

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

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 ( $P_j$ ), a vector  $X_{ij,0}$  of baseline individual and school controls, and a disturbance term  $\varepsilon_{ij}$ , which is clustered at the school level. 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 preprogram 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.

4. This study is registered on the American Economic Association RCT registry (#AEARCTR-0001191). We did not register a preanalysis plan, as they were uncommon in economics when data collection for this study was completed in 2009.

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 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 main coefficient of interest is  $\lambda_1$ , which captures gains accruing to individuals in treatment schools relative to the control; because 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 because it is a lower bound on the overall effect of deworming (Proposition 1). Another coefficient of some interest is  $\lambda_2$ , which captures the spillover effects for nearby schools, following the approach in Miguel and Kremer (2004), in which cross-school externalities are estimated by taking advantage of variation in the local density of treatment schools induced by the randomization. As explained further in that paper, since reinfection rates are 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 nonlinearities in saturation (Online Appendix B), but we 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 although we do not find evidence of interaction effects or nonlinear externalities, it remains possible that such effects would emerge at treatment levels outside the support of values we observe. One case of potential interest is one where 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, that is, 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 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  $\frac{1}{R}$  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  $\frac{1}{(2/3)} = 1.5$ , whereas the total benefits are assumed to remain unchanged. We present bounds using this approach in Section V.<sup>5</sup>

#### IV. 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.

##### IV.A. Long-Run Health Impacts

Although treatment dramatically reduced moderate-heavy infections in the short run (Table I), adult helminth life spans 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 likely instead be due to effects on other diseases through an immunological channel or to the effects of changes in schooling or labor outcomes.

5. Of course, if  $\pi(p) = \pi(p')$  and  $\lambda_2(p) = \lambda_2(p')$  for all  $p' > p$ , policy makers 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.

TABLE I  
DEWORMING IMPACTS ON HEALTH

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female	(4) All	(5) All	(6) Male	(7) Female
Moderate-heavy worm infections in 2001	−0.166*** (0.026)	−0.191*** (0.028)	−0.144*** (0.032)	−0.074 (0.223)	0.327 (0.469) 2,297	0.319 (0.466) 1,216	0.337 (0.473) 1,081
Self-reported health “very good” indicator at KLPS-2	0.040** (0.018)	0.023 (0.025)	0.051** (0.025)	0.128 (0.115)	0.673 (0.469) 5,070	0.713 (0.452) 2,585	0.629 (0.483) 2,485
Height at KLPS-2	−0.152 (0.272)	0.041 (0.376)	−0.367 (0.396)	−2.136 (1.632)	167.3 (7.9) 5,057	171.7 (6.4) 2,579	162.4 (6.4) 2,478
Body mass index at KLPS-2	0.121 (0.104)	−0.131 (0.112)	0.358** (0.167)	0.138 (0.539)	21.50 (2.36) 5,048	21.31 (2.10) 2,576	21.71 (2.62) 2,472
Miscarriage indicator (obs. at pregnancy level) at KLPS-2 (for females—their selves; for males—their partners)	−0.015* (0.008)	0.000 (0.004)	−0.028** (0.013)	−0.078** (0.037)	0.030 (0.171) 5,022	0.015 (0.123) 1,622	0.039 (0.194) 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 in columns (1)–(3) is from a separate OLS regression, except the miscarriage outcomes, 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  $\lambda_1$  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  $\lambda_2$  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 1 if the answer to the question “Would you describe your general health as somewhat good, very good, or not good?” is “very good,” and 0 otherwise. Sixteen unreasonably low values for height are dropped, but the height results are substantively the same if these are not dropped. Sixteen unreasonably high values (at 98 and above) for BMI are dropped (these are the same 16 observations that were dropped for height); note that the female BMI effect estimate is not statistically significant at traditional confidence levels if these observations are retained (not shown).

Although we find no long-term effects on height or body mass index in the full sample, 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 (std. err. 1.8,  $p < .05$ ), 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. We detect gains in body mass index among treated women ( $p < .05$ ). Furthermore, deworming reduced miscarriage rates among treatment group women by 2.8 percentage points (std. err. 1.3,  $p < .05$ ) on a base of 3.9% 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.

#### IV.B. *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 average academic performance did not improve, with higher enrollment translating into higher rates of grade repetition but no increase in educational attainment and no significant differences between the treatment and control groups in rates of passing the secondary school exam or enrolling in secondary school (Table II). We do not have data on whether increased primary school enrollment improved noncognitive skills, a possible channel for later 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 15–18) 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

TABLE II  
DEWORMING IMPACTS ON EDUCATION

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female		(5) All	(6) Male	(7) Female
Total years enrolled in school, 1998–2007	0.294** (0.145)	0.150 (0.166)	0.354* (0.179)	1.015 (0.839)	6.69 (2.97) <i>5,037</i>	7.05 (2.93) <i>2,567</i>	6.29 (2.96) <i>2,470</i>
Total years enrolled in primary school, 1998–2007	0.155** (0.075)	0.238** (0.102)	0.026 (0.098)	0.784 (0.485)	4.38 (2.48) <i>5,038</i>	4.43 (2.42) <i>2,568</i>	4.32 (2.55) <i>2,470</i>
Repetition of at least one grade (1998– 2007) indicator	0.063*** (0.018)	0.072*** (0.025)	0.053* (0.030)	0.099 (0.123)	0.672 (0.470) <i>5,084</i>	0.669 (0.471) <i>2,595</i>	0.676 (0.468) <i>2,489</i>
Grades of schooling attained by 2007	0.150 (0.143)	−0.030 (0.148)	0.261 (0.171)	0.323 (0.842)	8.72 (2.21) <i>5,084</i>	9.06 (2.28) <i>2,595</i>	8.34 (2.07) <i>2,489</i>
Attended secondary school indicator	0.030 (0.035)	−0.035 (0.038)	0.090** (0.038)	−0.032 (0.217)	0.421 (0.494) <i>5,084</i>	0.504 (0.500) <i>2,595</i>	0.329 (0.470) <i>2,489</i>
Passed secondary school entrance exam during 1998–2007 indicator	0.050 (0.031)	0.004 (0.030)	0.096** (0.040)	0.220 (0.161)	0.505 (0.500) <i>4,974</i>	0.590 (0.492) <i>2,541</i>	0.409 (0.492) <i>2,433</i>
Out-of-school (at 2007–2009 survey) indicator	−0.006 (0.022)	0.022 (0.030)	−0.029 (0.026)	0.185 (0.142)	0.75 (0.43) <i>5,058</i>	0.70 (0.46) <i>2,582</i>	0.80 (0.40) <i>2,476</i>

WORMS AT WORK  
1655

*Notes.* For details on the regressions, see the notes for Table I. Each entry in columns (1)–(3) is from a separate OLS regression. Significant at 90% (\*), 95% (\*\*), 99% (\*\*\*) confidence.

percentage points ( $p < .05$ ) on a base of 41%. This increase of roughly one quarter 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 a third (Online Appendix Table S3), and increasing overall years of school enrollment for women by 0.354 years (std. err. 0.179,  $p < .10$ ) (Table II). The estimated increase in girls' educational attainment is 0.261 years (std. err. 0.171,  $p = .13$ ), as some of the increased enrollment translated into increase grade repetition, as was the case for males.

#### IV.C. *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 a week (std. err. 1.42,  $p < .05$ , Table III, Panel A). In contrast, estimated effects on nonhousehold work hours among women are small. It is worth noting that one quarter of both the treatment and control groups were still in school by the time of the survey (Table II), and labor market outcomes are less meaningful for this group. We next focus on a subpopulation that is largely older than school age, which we operationalize as those who were older than 12 years (the median age) at baseline, and thus at least 22 or 23 years of age at follow-up: only 5% of control individuals in this age group were still enrolled in any school at follow-up, compared with 39% among younger control individuals. In this older subpopulation, average hours worked per week in the control group is somewhat higher: 28.2 hours for men and 21.7 hours for women. For this subgroup among both sexes, deworming increased time spent working by 13.0%, or 3.29 hours a week (std. err. 1.80,  $p < .10$ ), and treated men worked 3.74 more hours a week ( $p < .10$ ). Treated women worked 2.01 more hours a 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.

TABLE III  
DEWORMING IMPACTS ON LABOR HOURS AND OCCUPATIONAL CHOICE

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female	(4) All	(5) All	(6) Male	(7) Female
Panel A: Hours worked							
Hours worked in all sectors in last week, full sample	1.58 (1.04)	3.49** (1.42)	0.32 (1.36)	10.20 (7.80)	18.4 (23.1) <i>5,084</i>	20.3 (24.6) <i>2,595</i>	16.3 (21.1) <i>2,489</i>
Hours worked in all sectors in last week, older than school age subsample (older than 12 years of age at baseline)	3.29* (1.80)	3.74* (2.21)	2.01 (2.45)	18.0 (11.8)	25.4 (26.1) <i>2,235</i>	28.2 (27.2) <i>1,201</i>	21.7 (24.1) <i>1,034</i>
Panel B: Sectoral time allocation (full sample)							
Hours worked in nonagricultural self-employment in last week	1.51*** (0.55)	1.35* (0.73)	1.86** (0.81)	6.00* (3.23)	3.3 (12.8) <i>5,084</i>	3.8 (13.7) <i>2,595</i>	2.7 (11.7) <i>2,489</i>
Hours worked in agriculture in last week	−0.07 (0.42)	1.03* (0.55)	−1.27** (0.56)	−0.55 (3.41)	8.3 (11.4) <i>5,084</i>	7.8 (11.6) <i>2,595</i>	8.8 (11.2) <i>2,489</i>
Hours worked in wage earning in last week	0.14 (0.84)	1.11 (1.32)	−0.27 (1.08)	4.74 (5.07)	6.9 (18.5) <i>5,084</i>	8.8 (20.0) <i>2,595</i>	4.8 (16.5) <i>2,489</i>

WORMS AT WORK

1657

TABLE III  
(CONTINUED)

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of</i> <i>observations</i>		
	(1) All	(2) Male	(3) Female	(4) All	(5) All	(6) Male	(7) Female
Panel C: Occupational choice (full sample)							
Manufacturing job indicator	0.0110*** (0.0040)	0.0192** (0.0077)	0.0050 (0.0035)	0.0531** (0.0250)	0.0049 (0.0698) <i>5,084</i>	0.0068 (0.0824) <i>2,595</i>	0.0027 (0.0522) <i>2,489</i>
Construction/casual labor job indicator	−0.0053** (0.0026)	−0.0031 (0.0030)	−0.0073 (0.0045)	−0.0196 (0.0154)	0.0048 (0.0691) <i>5,084</i>	0.0040 (0.0628) <i>2,595</i>	0.0057 (0.0756) <i>2,489</i>
Domestic service job indicator	−0.0050 (0.0061)	0.0016 (0.0038)	−0.0134 (0.0129)	−0.0097 (0.0322)	0.0192 (0.1372) <i>5,084</i>	0.0067 (0.0813) <i>2,595</i>	0.0331 (0.1791) <i>2,489</i>
Grows cash crop indicator	0.0136** (0.0060)	0.0068 (0.0071)	0.0207** (0.0094)	0.0111 (0.0260)	0.0073 (0.0850) <i>5,068</i>	0.0080 (0.0890) <i>2,588</i>	0.0065 (0.0803) <i>2,480</i>

*Notes.* For details on the regressions, see the notes for Table I. Each entry in columns (1)–(3) is from a separate OLS regression. “Older than school age” denotes those older than 12 years of age (the median age) at baseline in 1998. Agricultural work in Panel B includes both farming and pastoral activities. Significant at 90% (\*), 95% (\*\*), 99% (\*\*\*) confidence.



Deworming changes how work hours are allocated across sectors and occupations, with important distinctions by gender (Table III, Panel B). Considering the genders together, hours in nonagricultural self-employment increase by 45% ( $p < .01$ ), and results are shown by gender in Figure II (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, and we cannot reject the hypothesis of equal percentage increases across sectors (Table III, Panel B). In contrast, women increase time in nonagricultural self-employment by 1.86 hours (std. err. 0.81,  $p < .05$ ) on a base of 2.7 hours, nearly 70%, and reduce hours worked in agriculture by 1.27 hours (std. err. 0.56,  $p < .05$ ). This shift from agricultural work into nonagricultural self-employment could potentially be interpreted as consistent with PRH, although the evidence is not dispositive. Seventy-seven percent 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 working in agriculture and those in nonagricultural self-employment.

Deworming treatment also leads to shifts in occupational choice (Table III, Panel C). Treatment respondents are three times more likely to work in manufacturing (coefficient 0.0110,  $p < .05$ ) from a low base of 0.005. On the flip side, casual labor—which typically does not require regular work hours—falls significantly ( $p < .05$ ). Manufacturing jobs require more hours a week than other occupations: they average 53 hours a week, compared with 42 hours for all wage-earning jobs, 34 hours for self-employment, and 15 hours for agriculture. Workers in manufacturing tend to miss relatively few work days due to poor health, at just 1.1 days in the past month (in the control group), compared with 1.5 days among all wage earners. Manufacturing jobs are highly paid, with average earnings more than double those in casual labor (Table S17). Deworming also leads to an increase in cash crop cultivation for the entire sample (Table III, Panel C), with a gain of 1.36 percentage points ( $p < .05$ ) on a low base of 0.73%.

Estimates of occupational effects by sex are less precise, but there are significant increases in manufacturing among men and in growing cash crops among women. The particularly large effect

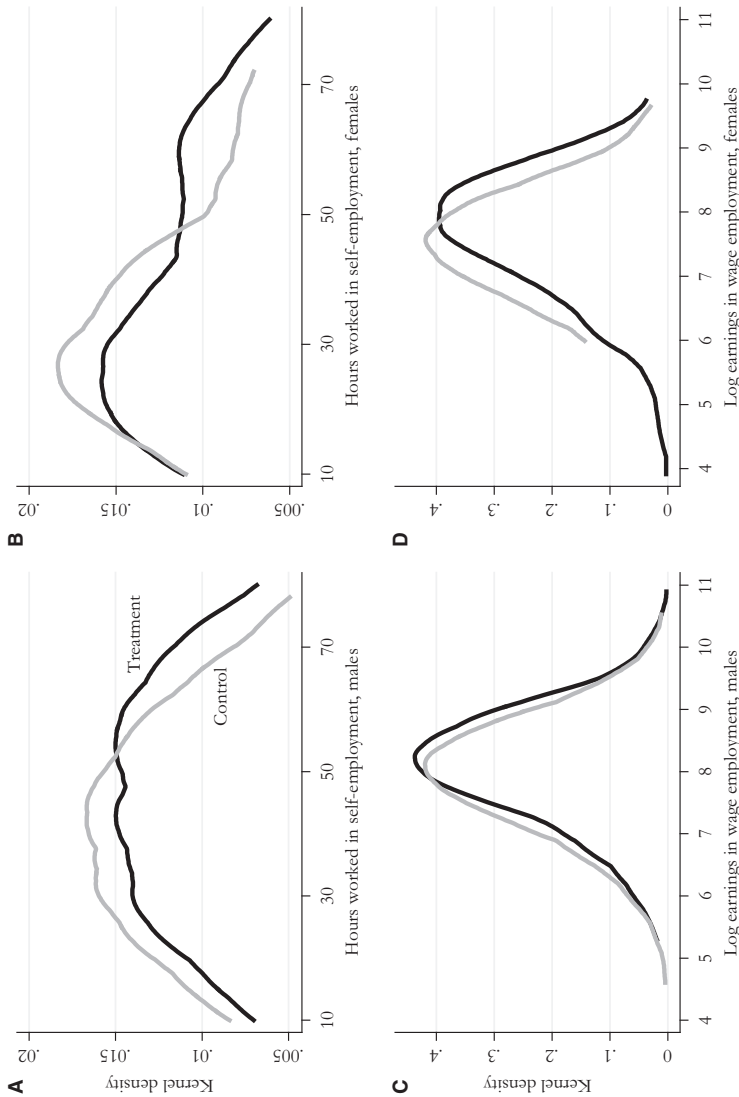


FIGURE II

Kernel Densities of Hours Worked in Self-Employment and Log Earnings in Wage Employment, Treatment versus Control

Kernel density in the treatment group is shown in black, and in the control group shown in gray. Panel A displays hours worked in self-employment in the last week (among those working 10–80 hours in the sector) for males, and Panel B displays the same for females. Panel C displays log earnings in wage employment in the past month (among those with positive earnings) for males, and Panel D displays the same for females.

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 Online Appendix C).

The increase in secondary education, nonagricultural self-employment, 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.

#### *IV.D. 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 we do have data on the number of meals consumed. Treatment respondents eat 0.095 more meals per day (std. err. 0.029,  $p < .01$ , Table IV, Panel A). The increase in meals eaten is larger for men, at 0.125 meals/day ( $p < .01$ ) than for women (0.051 meals), implying that treatment males miss just under one fewer meal each week than control males. Treatment effects are particularly large for the older than school age subsample (across both genders), at 0.119 more meals per day ( $p < .01$ ).

Total earnings are the sum of earnings in wage labor, in nonagricultural self-employment, and in agriculture, each weighted by the proportions working in each sector. We begin by considering total nonagricultural earnings (the sum of wage labor earnings and nonagricultural self-employment profits), which are likely to be more accurately captured than agricultural production in this setting. Those with no nonagricultural earnings are included in the analysis (with zero earnings). In the full sample, treatment respondents' total nonagricultural earnings are 15.0% higher (112 shillings, std. err. 96, Table IV, Panel A), although the effect is not statistically significant. In the older than school age subsample, the effect is considerably larger at 22.6% (278 shillings, std. err. 167,  $p = .101$ ).

We next consider each source of income separately. In principle, the proportions working in different sectors could differ by treatment group, but note that there are no significant differences by treatment status (Online Appendix Table S5, odd

TABLE IV  
DEWORMING IMPACTS ON LIVING STANDARDS AND LABOR EARNINGS

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female	(4) All	(5) All	(6) Male	(7) Female
Panel A: Consumption and nonagricultural earnings							
Number of meals eaten yesterday, full sample	0.095*** (0.029)	0.125*** (0.041)	0.051 (0.043)	0.415*** (0.124)	2.16 (0.64) <i>5,083</i>	2.10 (0.65) <i>2,595</i>	2.23 (0.62) <i>2,488</i>
Number of meals eaten yesterday, older than school age subsample (older than 12 years of age at baseline)	0.119*** (0.042)	0.147*** (0.051)	0.070 (0.063)	0.406* (0.236)	2.11 (0.66) <i>2,234</i>	2.04 (0.67) <i>1,201</i>	2.20 (0.63) <i>1,033</i>
Total nonagricultural earnings (wage earnings plus self-employed profits), past month, full sample	112 (96)	139 (171)	98 (68)	226 (694)	749 (2,132) <i>5,084</i>	1,115 (2,703) <i>2,595</i>	340 (1,075) <i>2,489</i>
Total nonagricultural earnings (wage earnings plus self-employed profits), past month, older than school age subsample (older than 12 years of age at baseline)	278 (167)	312 (265)	188 (139)	1,152 (971)	1,231 (2,440) <i>2,235</i>	1,774 (2,903) <i>1,201</i>	527 (1,375) <i>1,034</i>

TABLE IV							
(CONTINUED)							
	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female		(5) All	(6) Male	(7) Female
Panel B: Wage earnings (among wage earners)							
Ln(Total labor earnings), past month	0.269*** (0.085)	0.244** (0.109)	0.165 (0.175)	1.141 (0.869)	7.79 (0.88) <i>710</i>	7.92 (0.87) <i>542</i>	7.46 (0.81) <i>168</i>
Ln(Wage = Total labor earnings / hours), past month, if $\geq 10$ hours per week of work	0.197* (0.102)	0.181 (0.128)	0.225 (0.194)	0.378 (0.898)	2.68 (0.91) <i>601</i>	2.88 (0.89) <i>448</i>	2.21 (0.81) <i>153</i>
Ln(Total labor earnings), most recent month worked since 2007	0.225*** (0.070)	0.221** (0.097)	0.178* (0.104)	0.941 (0.597)	7.83 (0.91) <i>1,175</i>	7.97 (0.89) <i>819</i>	7.54 (0.89) <i>356</i>

TABLE IV  
(CONTINUED)

	Coefficient estimate (std. err.) on deworming treatment indicator			Coeff. est. (std. err.) externality term	Control group mean (std. dev.); <i>number of observations</i>		
	(1) All	(2) Male	(3) Female	(4) All	(5) All	(6) Male	(7) Female
Panel C: Nonagricultural self-employment outcomes (among nonagricultural self-employed)							
Total self-employed profits (self-reported) past month	384 (308)	111 (465)	250 (265)	−77 (1,646)	1,766 (2,619)	2,135 (3,235)	1,265 (1,261)
Total self-employed profits past month, top 5% trimmed	341* (177)	259 (309)	80 (219)	440 (1,256)	585 (1,151)	313 (1,056)	272 (1,261)
Total employees hired (excluding self)	0.416 (0.361)	0.245 (0.403)	0.603 (1.275)	−0.886 (2.547)	0.188 (0.624)	0.253 (0.614)	0.097 (0.630)
					633	343	290

*Notes.* For details on the regressions, see the notes for Table I. Each entry in columns (1)–(3) is from a separate OLS regression, except for “total employees hired” in Panel C, which uses a negative binomial regression. “Older than school age” denotes those older than 12 years of age (the median age) at baseline in 1998. 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 total nonagricultural earnings measure in Panel A includes those with zero reported earnings and profits. The wage, earnings, and profits results in Panels B and C are among those who reported wage employment or nonagricultural self-employment, respectively. When computing wages, we exclude those with fewer than 10 hours a week to address division bias from noise in estimation of number of hours worked. “Total employees hired” is among those who are self-employed. Significant at 90% (\*), 95% (\*\*), 99% (\*\*\*) confidence.

TABLE V  
FISCAL IMPACTS OF DEWORMING SUBSIDIES

	No subsidy	Partial subsidy	Full subsidy	Notes
Panel A: Calibration parameters				
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 $\times$ take-up rate
Mean per person increase in work hours/week: $\lambda_1$	0.00	0.44	1.75	Men: increase of 3.49 hours/week; women: no change (Table III). Partial subsidy multiplied by $\frac{Q(S)}{Q(\text{full})}$
Mean increase in work hours/week from externality: $p\lambda_2$	0.00	1.76	5.21	10.20 (Table III) $\times$ Coverage of treatment school students within 6 km ( $R$ , 68.1%) $\times$ [ $Q(S)$ for full subsidy, $\frac{Q(S)}{Q(\text{full})}$ for partial subsidy]
Mean increase in schooling costs	0.00	2.71	10.71	NPV of (additional secondary schooling costs per pupil-year (\$116.85) $\times$ direct increase in secondary schooling). Partial subsidy multi- plied by $\frac{Q(S)}{Q(\text{full})}$ .
Mean increase in schooling costs from externality	0.00	3.40	13.42	NPV of (additional secondary schooling costs per pupil-year (\$116.85) $\times$ externality in- crease in secondary schooling). Partial sub- sidy multiplied by $\frac{Q(S)}{Q(\text{full})}$ .



TABLE V  
(CONTINUED)

	No subsidy	Partial subsidy	Full subsidy	Notes
Panel B: No health spillovers				
Annual increase in per person earnings	\$0.00	\$3.91	\$15.44	$\lambda_1 \times \text{starting wage} \times 52$
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 $\times$ 16.575% tax rate – Direct schooling costs
Panel C: With health spillovers				
Annual increase in per person earnings	\$0.00	\$26.77	\$83.11	$(\lambda_1 + (\frac{p}{h}) \lambda_2) \times \text{starting wage} \times 52$
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	—	\$34.83	\$102.97	NPV earnings $\times$ 16.575% 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 \$0.17, a weighted average of wages by sector with weights corresponding to control group mean hours per sector (Table IV). We use Suri’s (2011) mean wage of \$0.16 as the agricultural wage, and the control group mean of \$0.23 (Table IV, Panel A) for those working for wages. Self-employed wages are calculated by dividing control group monthly profits (Table IV, Panel B) by 4.5 times the hours worked a 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 2001–2011 period, namely, 1.52% per annum (World Bank Development Indicators).

columns). Although weighted earnings by sector can always be summed to generate total earnings, the treatment versus control differences within particular sectors presented above reflect a combination of treatment and selection effects. Treatment and control individuals work as wage laborers at similar rates and have similar selection patterns along observable dimensions (Tables S5, S14–S15), but there are significantly different patterns of selection into wage employment and nonagricultural self-employment by treatment status (Table S5). This suggests that selection concerns are potentially important and that it may not be appropriate to interpret the differences between treatment and control individuals within employment sectors as causal impacts. Recall that the consumption and total nonagricultural earnings results above (in Table IV, Panel A) are based on the full sample and the issue of sorting across employment sectors does not apply.

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 II, Panel C) and women (Panel D) in the treatment group relative to control. Log earnings (Table IV, Panel B) are 26.9 log points (std. err. 8.5,  $p < .01$ ) 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 pay better. Log wages computed as earnings per hour worked (among those who work at least 10 hours a week) are 19.7 log points (std. err. 10.2,  $p < .10$ ) 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 ( $p < .01$ ) during the most recent earnings period.

The data on self-employment profits are likely measured with somewhat more noise. Monthly profits are 22% larger in the treatment group, but the difference is not significant (Table IV, 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% ( $p < .10$ ).

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 hours worked in agriculture did not change. Most important, 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 (std. err. 0.033) in this group (Online Appendix C).

#### *IV.E. Heterogeneous Treatment Effects and Alternative Specifications*

Although 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 Online Appendix Section C.4 and Tables S6–S13).

Estimated deworming impacts are largely robust to whether we account for the cross-school spillovers at all and to accounting for cross-school externalities at different distances (Online Appendix Tables S6–S9, column (5)). Online Appendix Figure S4 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 women,  $p$ -values range from .02 to .26). The externality results are similar if we focus on the number of local pupils, rather than the proportion, in treatment schools (Online Appendix Tables S6–S9, column (2)).

#### *IV.F. 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.

Online 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, 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 nonagricultural self-employment and manufacturing employment is significant at the 5% level, and the reduction in casual labor jobs, the increase in cash crops, and trimmed self-employed profits 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 (Online Appendix Section C.5).

#### *IV.G. 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 26 of the 30 outcomes presented in Tables I–IV (excluding the first outcome in Table I, 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 0 ( $p < .001$ ); see Online Appendix Section B for further details.

#### *IV.H. 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 I–IV are significant and large in magnitude, including for miscarriage, manufacturing employment, and meals eaten ( $p < .05$ ). Under the null hypothesis of no epidemiological externalities, there should be no correlation with the direct treatment effect. In 26 of the 30 post-2001 specifications in Tables I–IV, 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 0.750 between the  $t$ -statistics for the direct effect and the externality effect across outcomes ( $p < .001$ ); and using SUR, we reject the hypothesis that the 0–6 km cross-school externality effects are 0 ( $p < .001$ ); see Online Appendix B. The existence of cross-school externalities provides additional evidence on the robustness of the deworming impacts and reassurance that estimated effects are not simply due to some form of reporting bias in the treatment schools.

## V. 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 financial 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 somewhat speculative by necessity. 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: (i) estimated differences in earnings among wage workers are larger than differences in hours (Table IV, Panel B); (ii) 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, particularly if education and experience are complements (Card 1999); (iii) there is increased nonagricultural self-employment, particularly among women, and it seems plausible that some of this consists of investments that could pay off in increased earnings later; and (iv) estimated effects on hours worked and nonagricultural earnings are larger among those who are older and more likely to be out of school.

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

$$(5) \quad S_2Q(S_2) - S_1Q(S_1) < \sum_{\gamma} N_{\gamma} \left[ \tau(S_1) \sum_{t=0}^{t=50} \left( \frac{1}{1+r} \right)^t w_t (\lambda_{1,\gamma} + \frac{p\lambda_{2,\gamma}}{R}) - K \sum_{t=0}^{t=50} \left( \frac{1}{1+r} \right)^t \Delta \bar{E}_{\gamma t}(S_1, S_2) \right].$$

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 nondecreasing 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 V, 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  $S_2$ .  $N_\gamma$  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  $\gamma$ ;  $\lambda_{2,\gamma}$  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  $\gamma$ , 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  $\gamma$  when the deworming subsidy increases from  $S_1$  to  $S_2$ . To compute the right-hand side of equation (5), we use a combination of estimates from this article and other Kenyan data. The hours worked estimates (Table III) 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  $p\lambda_{2,\gamma}$ .<sup>6</sup> Because 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

6. Results are similar when externalities are disaggregated by gender (not shown).

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 ( $\frac{1}{R}$ ) since the control group also benefits from externalities.

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%.<sup>7</sup> We assume that the sample population begins working 10 years after they first began receiving deworming and retires after 40 years of work.<sup>8</sup> From year 10 post-treatment onward, we combine estimated  $\lambda_{1,\gamma}$  and  $\lambda_{2,\gamma}$  values from Tables III–IV with the pattern of life cycle 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 \$0.17 an 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 IV).<sup>9</sup> Kenyan taxes (mainly on consumption) absorb roughly 16.6% of GDP, so we set the tax rate under no subsidy to 16.6%.<sup>10</sup>

7. See <http://www.centralbank.go.ke/securities/bonds/manualresults.aspx> and World Bank Development Indicators. This is a conservative assumption because other potential funders of deworming subsidies (e.g., international organizations, private donors) are likely to face lower interest rates; use of a lower interest rate greatly increases the returns to deworming in the calculations described later.

8. This 10-year gap roughly corresponds to the time elapsed from the start of PSDP until the KLPS-2 survey (2007–2009). By ignoring the time before KLPS-2 data were 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 forgone earnings are likely to be small (Udry 1996).

9. In Suri (2011), the mean agricultural wage is \$0.16, and the control group mean is \$0.23 (Table IV, Panel B) for those working for wages. Self-employed wages are calculated by dividing control group monthly profits (Table IV, Panel C) by 4.5 times the hours worked per week among those working in self-employment, for a wage of \$0.14.

10. 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  $0.195 * 0.85 = 0.166$ .



We estimated deworming impacts on school enrollment by gender and year (Online Appendix Table S3) and 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 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 the government would need to incur to maintain the secondary school pupil-teacher ratio using our estimated per student secondary school teacher cost of \$116.85 per year (Table V, Panel A).

Assuming no externality gains,  $\sum_{\gamma} N_{\gamma} \sum_{t=0}^{t=50} \left( \frac{1}{(1+r)} \right)^t w_t \lambda_{1,\gamma} = \$142.43$ , implying that individuals gain an average of \$119 in take-home pay and the net present value (NPV) of government revenue increases by \$23 per person (Table V, 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, that is, 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 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 Online 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, taxed or untaxed: in the above notation,  $MQ(S) = \sum_{\gamma} N_{\gamma} \left[ \sum_{t=0}^{t=50} \left( \frac{1}{1+r} \right)^t w_t \left( \lambda_{1,\gamma} + \frac{p\lambda_{2,\gamma}}{R} \right) - K \sum_{t=0}^{t=50} \left( \frac{1}{1+r} \right)^t \Delta \bar{E}_{\gamma t}(0, S) \right]$ . The annualized social IRR with no health spillovers ( $\lambda_{2,\gamma} = 0$ ) is very high at 31.8%, and with health spillovers is a massive 51.0%.

These fiscal and IRR calculations are speculative for several reasons, including the projection of future earnings, as noted. 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 (Online Appendix C contains a discussion). They are also relatively imprecisely estimated: we bootstrapped standard errors (with 1,000 runs), and find that net revenue gains are less than zero 29% of the time for the case of no health spillovers. So although estimates indicate that the expected net revenue effects of deworming are large, there remains considerable uncertainty around these estimates.

Yet these calculations are conservative in several dimensions. For one thing, 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 IRR exercises 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; the same could be said for any inherent welfare 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 (birth to age 2) leads to improved cognitive and academic performance 10 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.

## VI. CONCLUSION

Previous work (Miguel and Kremer 2004) found that a primary-school deworming program increased school participation. This article shows that some education and labor market outcomes improve a decade after deworming. These gains could have substantial positive welfare impacts for households living near subsistence, like many in our Kenyan sample. A conservative estimate of the annualized financial IRR to deworming is high at 31.8%. Our best estimate is that deworming subsidies will generate more in future government revenue than they cost in up-front expenditures.<sup>11</sup>

The high rate of return to deworming in our Kenyan context is consistent with the finding of sizable deworming impacts on education and incomes in the twentieth-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).

The results also have implications for several related literatures. Many studies argue that early childhood health gains *in*

11. 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 (Ditsuwana et al. 2013).

*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.

Although 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 although 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) 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 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 may be a more appropriate initial response.

GEORGE WASHINGTON UNIVERSITY

UNIVERSITY OF CALIFORNIA, BERKELEY

HARVARD UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

UNIVERSITY OF CALIFORNIA, BERKELEY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

## SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at QJE online ([qje.oxfordjournals.org](http://qje.oxfordjournals.org)).

## REFERENCES

- Ahuja, Amrita, Sarah Baird, Joan Hamory Hicks, Michael Kremer, Edward Miguel, and Shawn Powers, "When Should Governments Subsidize Health? The Case of Mass Deworming," *World Bank Economic Review*, 29 (supplement) (2015), S9–S24.
- Aiken, Alexander, M., Davey Calum, James R. Hargreaves, and Richard J. Hayes, "Re-Analysis of Health and Educational Impacts of a School-Based Deworming Programme in Western Kenya: A Pure Replication," *International Journal of Epidemiology*, 44 (2015), 1572–1580.
- Albonico, Marco, Dirk Engels, and Lorenzo Savioli, "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 (2004), 1205–1210.
- Almond, Douglas, and Janet Currie, "Human Capital Development before Age Five," NBER Working Paper 15827, 2010.
- Altonji, Joseph G., and Rebecca M. Blank, "Race and Gender in the Labor Market," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds. (Amsterdam: Elsevier, 1999).
- Anderson, Michael L., "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 (2008), 1481–1495.
- Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli, "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate," *Biometrika*, 93 (2006), 491–507.
- Bertrand, Marianne, "New Perspectives on Gender," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds. (Amsterdam: Elsevier, 2011).
- Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, 122 (2007), 73–117.
- , "Health, Human Capital, and Development," *Annual Review of Economics*, 2 (2010), 283–310.
- Brooker, Simon, Megan Rowlands, Laurence Haller, Lorenzo Savioli, and Donald A. P. Bundy, "Towards an Atlas of Helminth Infection in Sub-Saharan Africa: The Use of Geographical Information Systems (GIS)," *Parasitology Today*, 16 (2000), 303–307.
- Bruhn, Miriam, and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1 (2009), 200–232.
- Bundy, Donald A. P., Michael Kremer, Hoyt Bleakley, Matthew C. H. Jukes, and Edward Miguel, "Deworming and Development: Asking the Right Questions, Asking the Questions Right," *PLoS Neglected Tropical Diseases*, 3 (2009), e362. doi:10.1371/journal.pntd.0000362.
- Card, David, "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics* Volume 3A, Orley Ashenfelter and David Card, eds. (Amsterdam: Elsevier, 1999).
- Clemens, Michael, and Justin Sandefur, "Mapping the Worm Wars: What the Public Should Take Away from the Scientific Debate about Mass Deworming," *Center for Global Development Blog*, 2015.
- Crimmins, Eileen M., and Caleb E. Finch, "Infection, Inflammation, Height, and Longevity," *Proceedings of the National Academy of Sciences*, 103 (2005), 498–503.

- Croke, Kevin T., "The Long Run Effects of Early Childhood Deworming on Literacy and Numeracy: Evidence from Uganda," unpublished manuscript, 2014.
- Davey, Calum, M. Alexander Aiken, Richard J. Hayes, and James R. Hargreaves, "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*, 44 (2015), 1581–1592.
- Disease Control Priorities Project, "Deworming Children Brings Huge Health and Development Gains in Low-Income Countries," 2008, available at <https://docs.google.com/file/d/0BwhPKU71ZwQDRmxjS0RYRkdRdFU/edit>.
- Ditsuwana, Vallop, Jacob Lennert Veerman, Melanie Y. Bertram, and Theo Vos, "Cost-Effectiveness of Interventions for Reducing Road Traffic Injuries Related to Driving under the Influence of Alcohol," *Value in Health*, 16 (2013), 23–30.
- Fafchamps, Marcel, "Sequential Labor Decisions under Uncertainty: An Estimable Household Model of West African Farmers," *Econometrica*, 61 (1993), 1173–1197.
- Geerts, Stanny, and Bruno Gryseels, "Drug Resistance in Human Helminths: Current Situation and Lessons from Livestock," *Clinical Microbiology Reviews*, 13 (2000), 207–222.
- , "Anthelmintic Resistance in Human Helminths: A Review," *Tropical Medicine and International Health*, 6 (2001), 915–921.
- GiveWell, "Combination Deworming (Mass Drug Administration Targeting Both Schistosomiasis and Soil-Transmitted Helminths)," 2013, available at <http://www.givewell.org/international/technical/programs/deworming>.
- Grossman, Michael, "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80 (1972), 223–255.
- Guyatt, Helen L., Simon Brooker, Charles M. Kihamia, Andrew Hall, and Donald A. P. Bundy, "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 (2001), 695–703.
- Hall, Andrew, and Sue Horton, "Best Practice Paper: New Advice from CCo8: Deworming," Copenhagen Consensus Center, 2008.
- Heckman, James J., Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Non-Cognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24 (2006), 411–482.
- Hicks, Joan Hamory, Michael Kremer, and Edward Miguel, "Commentary: Deworming Externalities and Schooling Impacts in Kenya: A Comment on Aiken et al. (2015) and Davey et al. (2015)," *International Journal of Epidemiology*, (2015), doi:10.1093/ije/dyv129.
- Hotez, Peter J., "Empowering Women and Improving Female Reproductive Health through Control of Neglected Tropical Diseases," *PLoS Neglected Tropical Diseases*, 3 (2009), 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, *Helminth Infections: Soil-Transmitted Helminth Infections and Schistosomiasis*, 2nd ed. (Washington, DC: World Bank, 2006).
- Jameel Poverty Action Lab, "Deworming: A Best Buy for Development," *J-PAL Policy Bulletin* (2012).
- Kirwan, Patrick, 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, "Impact of Repeated Four-Monthly Anthelmintic Treatment on Plasmodium Infection in Preschool Children: A Double-Blind Placebo-Controlled Randomized Trial," *BMC Infectious Diseases*, 10 (2010), doi:10.1186/1471-2334-10-277.
- Kjetland, Eyrun F., Patricia D. Ndhlovu, Exenevia Gomo, Takafira Mduluzza, Nicholas Midzi, Lovemore Gwanzura, Peter R. Mason, Leiv Sandvik, Henrik Friis, and Svein Gunnar Gundersen, "Association between Genital



- Schistosomiasis and HIV in Rural Zimbabwean Women," *AIDS*, 20 (2006), 593–600.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75 (2007), 83–119.
- Kremer, Michael, Jean Lee, Jonathan Robinson, and Olga Rostapshova, "Behavioral Biases and Firm Behavior: Evidence from Kenyan Retail Shops," *American Economic Review Paper and Proceedings*, 103 (2013), 362–368.
- Kremer, Michael, and Edward Miguel, "The Illusion of Sustainability," *Quarterly Journal of Economics*, 112 (2007), 1007–1065.
- Larocque, Renée, Martin Casapia, Eduardo Gotuzzo, J. Dick MacLean, Julio C. Soto, Elham Rahme, and Theresa W. Gyorkos, "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 (2006), 1485–1495.
- Lightwood, James, and Stanton A. Glantz, "The Effect of the California Tobacco Control Program on Smoking Prevalence, Cigarette Consumption, and Healthcare Costs: 1989–2008," *PLoS One*, 8 (2013), e47145.
- Miguel, Edward, and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72 (2004), 159–217.
- , "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, 2014.
- Miguel, Edward, Michael Kremer, and Joan Hamory Hicks, "Comment on Macartan Humphreys' and Other Recent Discussions of the Miguel and Kremer (2004) Study," CEGA Working Paper 54, 2015.
- Miguel, Edward, Michael Kremer, Joan Hamory Hicks, and Carolyn Nekesa, "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, 2014.
- Orr, Larry, Judith Feins, Robin Jacob, Erik Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling, "Moving to Opportunity: Interim Impacts Evaluation," U.S. Department of Housing and Urban Development, 2003.
- Otieno, Wycliffe, and Christopher Colclough, "Financing Education in Kenya: Expenditures, Outcomes and the Role of International Aid," RECOUP Working Paper 25, 2009.
- Ozier, Owen, "Exploiting Externalities to Estimate the Long-Term Benefits of Early Childhood Deworming," World Bank Working Paper Series 7052, 2014.
- Pitt, Mark M., Mark R. Rosenzweig, and Mohammad Nazmul Hassan, "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy," *American Economic Review*, 102 (2012), 3531–3560.
- Prescott, Edward, "Why Do Americans Work So Much More than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28 (2004), 2–13.
- Pullan, Rachel L., Peter W. Gething, Jennifer L. Smith, Charles S. Mwandawiro, Hugh J. Sturrock, Caroline W. Gitonga, Simon I. Hay, and Simon Brooker, "Spatial Modelling of Soil-Transmitted Helminth Infections in Kenya: A Disease Control Planning Tool," *PLoS Neglected Tropical Diseases*, 5 (2011), e958.
- Schapiro, Louis, "The Physical and Economic Benefits of Treatment for Hookworm Disease," *Journal of the American Medical Association*, 73 (1919), 1507–1509.
- Silva, Nilanthi R. de, Simon Brooker, Peter J. Hotez, Anotnio Montresor, Dirk Engels, and Lorenzo Savioli, "Soil-Transmitted Helminth Infections: Updating the Global Picture," *Trends in Parasitology*, (2003), 547–551.
- Stephenson, Lani S., Michael C. Latham, Elizabeth J. Adams, Stephen N. Kinoti, and Anne Pertet, "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," *Journal of Nutrition*, 123 (1993), 1036–1046.

- Stoltzfus, Rebecca J., Hababu M. Chwaya, James M. Tielsch, Kerry J. Schulze, Marco Albonico, and Lorenzo Savioli, "Epidemiology of Iron Deficiency Anemia in Zanzibari Schoolchildren: The Importance of Hookworms," *American Journal of Clinical Nutrition*, 65 (1997), 153–159.
- Suri, Tavneet, "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 79 (2011), 159–209.
- Taylor-Robinson, David C., Nicola Maayan, Karla Soares-Weiser, Sarah Donegan, and Paul Garner, "Deworming Drugs for Soil-Transmitted Intestinal Worms in Children: Effects on Nutritional Indicators, Haemoglobin, and School Performance," *Cochrane Database of Systematic Reviews*, 7 (2015), 1–157.
- Weisbrod, Burton A., Ralph L. Andreano, Robert E. Baldwin, and Allen C. Kelley, *Disease and Economic Development: The Impact of Parasitic Diseases in St. Lucia* (Madison: University of Wisconsin Press, 1973).
- World Health Organization, "Model Describing Information. Drugs used in Parasitic Diseases," WHO, 1992.
- , "Soil-Transmitted Helminth Infections," in *WHO Fact Sheet 366* (2014).