Worms at Work: Long-run Impacts of Child Health Gains*

Sarah Baird
George Washington University

Joan Hamory Hicks
University of California, Berkeley CEGA

Michael Kremer
Harvard University and NBER

Edward Miguel
University of California, Berkeley and NBER

First version: October 2010
This version: August 2012

Abstract: We use experimental variation in a Kenyan deworming program to calibrate the Grossman (1972) model, in which health investments increase future endowments of healthy time, and estimate the labor market and fiscal impacts of such investments. Ten years after the start of the program, the treatment group has better self-reported health, consume more meals, spend more time in entrepreneurship, and are more likely to grow cash crops. Kenyan women who participated in the program as girls have fewer miscarriages and reallocate labor time from agriculture to entrepreneurship. Men who participated as boys work 3.4 more hours each week, and are more likely to hold manufacturing jobs with higher wage earnings. The deworming program generates positive externalities from reduced disease transmission. A calibration suggests that fully subsidizing deworming costs less than the additional net present value of government revenue it generates, creating an "expenditure Laffer effect" in which government subsidies for health investments allow for reduced tax rates.

* Acknowledgements: Kevin Audi, Pierre Bachas, Chris Blattman, Hana Brown, Lorenzo Casaburi, Lisa Chen, Garret Christensen, Lauren Falcao, Francois Gerard, Eva Arceo Gomez, Jonas Hjort, Gerald Ipapa, Maryam Janani, Anne Karing, Andrew Fischer Lees, Jamie McCasland, Owen Ozier, Kristianna Post, Adina Rom, Martin Rotemberg, Changcheng Song, Sebastian Stumpner, Paula Vinchery, Michael Walker, Paul Wang, and Ethan Yeh provided excellent research assistance on the KLPS project. We thank Michael Anderson, Jere Behrman, David Card, Alain de Janvry, Erica Field, Fred Finan, Michael Greenstone, Jim Heckman, Adriana Lleras-Muney, Steve Luby, Isaac Mbiti, Mark Rosenzweig, T. Paul Schultz, Jim Smith, John Strauss, Glen Weyl, Alix Zwane and seminar participants at U.C. Berkeley, USC, Harvard, the J-PAL Africa Conference, the Pacific Conference on Development Economics, UCSF, the Gates Foundation WASH Convening in Berkeley, Yale, the 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, and Makerere University for helpful suggestions. We gratefully acknowledge our NGO collaborators (International Child Support and Innovations for Poverty Action Kenya), 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. All errors remain our own.

1. Introduction

We use experimental variation in a Kenyan deworming program to calibrate the Grossman (1972) model, in which health investments increase future endowments of healthy time, and estimate the labor market and public finance impacts of such investments. Miguel and Kremer (2004) found that children who were dewormed are healthier and spend substantially more time in school. We follow participants a decade later, when most were 19 to 26 years old, and find that the treatment group has better self-reported health, spends more time in entrepreneurship, and is more likely to grow cash crops. Living standards improve as well, with treatment respondents missing one fewer meal per week. Kenyan women who participated in the program as girls have fewer miscarriages and reallocate labor time from agriculture to entrepreneurship. Men who participated as boys work 3.4 more hours each week as young adults (on a base of 20.3 hours), and are more likely to hold manufacturing jobs with higher wage earnings. The deworming program also generated positive epidemiological externalities from reduced disease transmission. The analysis is based on a new longitudinal dataset with an effective survey tracking rate of 84% over roughly ten years.

One implication of combining the Grossman (1972) model with a simple public finance model is that, in the presence of pre-existing distortionary taxes, subsidizing investments in health could generate a positive *fiscal externality*. The extra endowment of healthy time generated by these subsidies can increase work hours and thus tax revenue, which benefits other citizens; the argument follows Kaplow (2008, 2009).[1] Thus the social cost of financing one dollar's worth of public health expenditure may be less than a dollar, rather than more than a dollar, which would be the case under the standard deadweight loss of taxation logic. In some cases, the social cost of financing a dollar of investment may be negative: it may raise social

[1] This connection was brought to our attention by Glen Weyl.

welfare to subsidize health investments which no individual would choose to make in the absence of subsidies, and this could hold even in the absence of epidemiological externalities (like those that are relevant for infectious diseases). In such cases, instituting a health treatment subsidy can be a Pareto improvement, benefiting not only those residents who take it up but also those who do not, who benefit from lower tax rates.

Analysis of the welfare impacts of subsidies for health investments requires combining information on the price of responsiveness of demand with information on the labor market and fiscal consequences of these investments. Because we have an unusual dataset containing information on both deworming take-up (from Kremer and Miguel 2007) and long-run impacts on work hours and other labor market outcomes (in this paper), we can directly calibrate the model.[2] We find that full subsidies for deworming generate greater social welfare than either zero or partial subsidies in this setting, and the point estimates indicate that fully subsidizing deworming costs less than the additional net present value of government revenue it generates, creating an "expenditure Laffer effect" in which government subsidies for health human capital allow for reduced tax rates.

The increased work hours for men that we document helps shed light on the determinants of labor supply, an understudied issue in development economics. While there is considerable discussion about how work hours in wealthy countries differ with tax rates and labor market institutions (Prescott 2004, Costa 2000), differences in labor hours associated with economic development across space and time have been less studied, despite the fact that they are often larger than differences across wealthy countries. Some historians see a move to a work life governed by long, regular hours and factory discipline as an important aspect of the industrial

---

[2] For further discussion of the price-responsiveness of health treatment take-up, refer to Kremer and Holla (2009), Cohen and Dupas (2010), Ashraf, Berry and Shapiro (2010), and Dupas (2011).

revolution (Clark 1994). Work hours are low in some rural low-income contexts. For example, in Sahelian Burkina Faso, Fafchamps (1993) finds that farmers only work an average of two to three hours per day, arguing that this is due to low marginal productivity of labor in rain-fed agriculture during much of the year (Lewis 1954). Colonial observers advanced racial or ethnic theories of Africans' "laziness", love of leisure and lack of ambition (see Abudu 1986). A growing body of work in labor economics emphasizes cultural factors as key drivers of labor supply decisions (Fernandez and Fogli 2009; Alesina, Giuliano and Nunn 2012), while others explore behavioral economics factors (Clark 1994; Kaur, Kremer, and Mullainathan 2011). Our work suggests that child health can affect long-run labor supply.[3]

Beyond finding impacts on total work hours, we see a particular increase in entrepreneurship (non-agricultural self-employment). One interpretation is that the marginal product of labor in agriculture is relatively low, and hence people seek new opportunities to use their time more productively. There seems to be a movement toward "higher intensity" work, with a tripling of manufacturing employment (albeit on a relatively low base) and less casual labor. Earnings increase among wage workers by more than 20%. These patterns are consistent with the hypotheses that the ability to do regular, full-time work allows people to get better jobs, as manufacturing jobs are among the most demanding in our dataset, with long average work weeks. In an Oaxaca-style decomposition, these shifts in employment occupation account for

[3] This study contributes to the literature on long-run impacts of child health gains. The seminal INCAP experiment in Guatemala (Hodinott *et al.* 2008, Maluccio *et al.* 2009, Behrman *et al.* 2009) provided nutritional supplementation to children in two villages while two others served as a control, and finds gains in male wages of one third, improved cognitive skills among both men and women, and positive intergenerational effects on the nutrition of beneficiaries' children. Beyond the small sample size, a limitation of these studies is their 40% attrition rate over the 35 years of follow-up. Other studies have studied long-run economic impacts of child health, including effects of war-induced famine in Zimbabwe (Alderman *et al.*, 2006a) and economic shocks driven by rainfall variation in Indonesia (Maccini and Yang, 2009). Other work that finds that moderate increases in morbidity affect labor supply (Ichino and Moretti 2009, Hanna and Oliva 2011). Schultz and Tansel (1997) and Habyarimana, Mbakile and Pop-Eleches (2010) find links between disease and absence from work. Other noteworthy micro-empirical contributions on nutrition, health and productivity include Glewwe *et al.* (2001), Schultz (2005), Jukes et al. (2006), Alderman (2007), Thomas *et al.* (2008), and Pitt, Rosenzweig and Hassan (2011). Related U.S. work includes Currie, Garces and Thomas (2002), Currie (2009), Smith (2009), and Case and Paxson (2010).

most of the earnings gains in the treatment group. Point estimates also suggest higher profits for small non-agricultural entrepreneurs (although these latter estimates are imprecise).

The somewhat different pattern of results by gender is noteworthy. Pitt, Rosenzweig and Hassan (2011) argue that health investments are likely to have a larger impact on males than females in physical "brawn" based economies, like rural Kenya, and this point may help explain why some labor market impacts among our respondents are larger for males, especially for work hours. Yet women who received deworming do show improvements in self-reported health and have a lower miscarriage rate, an indication of improved overall health status. Women's family circumstances may also partially explain the pattern. Women are roughly twice as likely as male respondents to be married at the follow-up, much more likely to have young children, and work significantly fewer hours outside the home. While there is no significant impact on women's total work hours, there are once again notable increases in more productive economic activities: there is a significant shift out of agricultural work and into non-agricultural entrepreneurship, and among those who remain in agriculture, a large increase in cash crops cultivation.

The findings have several implications for our understanding of links between health and productivity. First, many existing studies of the impact of health on productivity track production within a firm, for example, examining the impact of contemporaneous health on plantation workers' productivity (e.g., Fox et al. 2004). Our evidence suggests this approach misses important gains, in particular on how health investments may lead to shifts across employment occupations, sectors, and activities. Second, while many studies argue that early childhood health gains *in utero* or before age three have the largest impacts (World Bank 2006, Hodinott *et al.* 2008, Almond and Currie 2010), our findings show that even health investments made in primary school-age children can have important long-run effects. Finally, we cannot decisively distinguish the extent to which we are observing the direct impact of health as opposed to the indirect effect of health through endogenous changes in education or other behaviors and attitudes. There is some evidence that deworming increased years enrolled in school and, among those currently out-of-school, improved exam performance.

The rest of the paper is organized as follows. Section 2 presents a simple model of the costs and benefits of health subsidies in a health capital framework related to Grossman (1972), and works out their fiscal implications. Section 3 discusses the Kenyan context, the deworming project, and the survey. Section 4 lays out the estimation strategy and describes the impacts of deworming on health, education, work hours, and meal consumption. Section 5 combines the data on price responsiveness of take-up and long-run deworming impacts to calibrate the model, and finds that full subsidies for deworming yield greater welfare than partial or no subsidies, and "pay for themselves", in that the net present value of additional taxes generated by deworming exceeds its cost. Section 6 breaks out the data by economic sector, arguing that work hours increase most outside traditional agriculture, and presents evidence for productivity gains, especially among wage earners. The final section concludes, discussing external validity and implications for research and policy.

2. A Model of Health Subsidies and their Labor and Fiscal Impacts

In the classic Grossman (1972) model, better health status increases "the total amount of time [one] can spend producing money earnings and commodities" (p. 224). We consider a variant of this model in which health investments may lead to increased endowments of healthy time not just for the individual but also for neighbors through an epidemiological externality. Furthermore, we examine implications for optimal deworming subsidies and their fiscal impacts.

The way changes in an individual's endowment of healthy time affect time spent in work and leisure depends on the form of the utility function. We demonstrate that under certain fairly general conditions, it is sufficient to examine observable changes in labor supply and health treatment take-up rates to determine if subsidies for health investments are Pareto-improving. In particular, such subsidies are Pareto-improving if they generate an increase in the net present value (NPV) of taxes exceeding the additional cost of the subsidies. Diamond and Mirrlees (1971) and Atkinson and Stern (1974) both note that optimal subsidization depends on whether the subsidized good is complementary to taxed goods, which in turn decreases the relative cost of the subsidies. It is even possible that the increase in taxable consumption generated by the subsidy policy could outweigh the cost of the subsidy itself, although Diamond and Mirrlees (1971) do not focus on this case. In section 5, we use the empirical estimates of long-run labor supply impacts from this paper, combined with earlier results on the responsiveness of deworming take-up to price (Kremer and Miguel 2007) to argue that these conditions are likely to be satisfied in the case of deworming that we study.

Consider a government that can set a linear income or consumption tax, but does not have access to more complicated non-linear taxes. The government can borrow or save at rate r.[4] The government faces an exogenous expenditure requirement equivalent to a net present value of G per person over each citizen's lifetime (e.g., for national defense or schooling). It must choose whether to subsidize an additional health investment for its citizens.

Let x denote the change in work hours for those who deworm, and x_{EXT} denote any spillover effect of deworming on others' work hours through epidemiological externalities. While we focus on this "reduced form" increase in work hours below, it is worthwhile to briefly

[4] To the extent that governments are credit-constrained, there could be potentially severe consequences from these constraints on tax instruments.

consider the micro-foundations of such an increase. Suppose individuals have utility over consumption (c_i), leisure (l_i), and non-pecuniary benefits from health (H_i), $U_i(c_i, l_i, H_i)$. An individual i makes a choice over health investments, which determines both their health and total time endowment (Z_i), which they divide into leisure and labor (L_i), consuming labor earnings.

In the case of Cobb-Douglas utility with separable health benefits, $U_i(c_i, l_i, H_i) = c_i^\alpha (Z_i - L_i)^{1-\alpha} + f(H_i)$. When health investments are costless, it is straightforward to show that individuals will choose to work a fraction α of their time endowment, allowing the econometrician to identify their total endowment of healthy time as $\frac{L_i}{\alpha}$. With this utility function and hourly wage rate w, an exogenous increase in healthy time by $\tilde{Z}$ hours increases utility by the same amount as an exogenous cash transfer of $\tilde{Z}w$. The econometrician can thus further determine the utility gain (in money-metric terms) generated by deworming using both the observable wage rate and the change in work hours: $\frac{x}{\alpha}w$.

There would likely be heterogeneous responses to health changes across individuals and demographic groups in the population (e.g., by gender), where those whose utility functions put more "weight" on consumption relative to leisure (higher α) in turn have a higher elasticity of work hours with respect to the total endowment of healthy time. To illustrate, if women place more weight on "leisure" (non-work) time than men on average, perhaps because they prefer to spend more time at home with their children, then their α would be lower and thus their total work hours would respond less to health investments than men. Similarly, women could place less value on consumption if there are fewer socially acceptable consumption options for women in rural Kenya than for men (for instance, it is socially acceptable for married men but not women to drink in bars with friends), in which case once again α could be lower and their work hours would be less responsive to health gains. These are meant to be illustrative possibilities, but regardless of each individual's exact utility function, the only assumption that we rely on in the welfare analysis framework below is that dewormed individuals who choose to work more hours have (weakly) greater utility than those who did not receive deworming.

Abstracting away from the exact utility function discussed above, suppose there is a mass 1 of individuals, each of whom, in the absence of the health investment, chooses to spend E hours working. Deworming medicine is competitively provided at price p, and hourly wages are w. Define $\theta(s)$ as the proportion of the population that takes deworming medicine given subsidy level s, where individuals choose to deworm ($D_i = 1$) if doing so weakly increases their lifetime utility. Thus at subsidy level s each individual i chooses to work $E + xD_i + x_{EXT}\theta(s)$ hours, where we assume (as discussed below in the estimation section) that the externality impacts increase linearly in the proportion of other citizens who take deworming.

A linear income tax $\tau(s)$ that equates the NPV's of expenditure and revenue must satisfy:

$$G + s\theta(s) = \frac{1}{1-r} w[E + \theta(s)(x + x_{EXT})]\tau(s).$$

While we impose that r equals the interest rate for notational and analytical convenience, the qualitative results that follow hold for any coefficient that maps future tax revenue to current value. As a result, for a given subsidy s we can define a unique tax rate satisfying:

$$\tau(s) = \frac{G + s\theta(s)}{w[E + \theta(s)(x + x_{EXT})]/(1-r)}.$$

Proposition 1: **Given two subsidy levels $s_1 > s_2$, if the net present value of extra revenue associated with subsidy s_1 (relative to s_2) under tax rate $\tau(s_2)$ is greater than the extra expenditures associated with s_1, i.e., if**

$$s_1\theta(s_1) - s_2\theta(s_2) < \frac{1}{1-r} w\big[\big(\theta(s_1) - \theta(s_2)\big)(x + x_{EXT})\big]\tau(s_2), \qquad \textbf{(eqn. 1)}$$

then the higher subsidy rate represents a Pareto-improvement over the lower one.

Proof:

$\tau(s_2)$ solves

$$G + s_2\theta(s_2) = \frac{1}{1-r} w\big[E + \big(\theta(s_2)\big)(x + x_{EXT})\big]\tau(s_2),$$

so

$$G + s_2\theta(s_2) + s_1\theta(s_1) - s_2\theta(s_2) < \frac{1}{1-r} w\big[E + \big(\theta(s_2)\big)(x + x_{EXT})\big]\tau(s_2) + \frac{1}{1-r} w\big[\big(\theta(s_1) - \theta(s_2)\big)(x + x_{EXT})\big]\tau(s_2),$$

which implies that

$$G + s_1\theta(s_1) < \frac{1}{1-r} w\big[E + \big(\theta(s_1)\big)(x + x_{EXT})\big]\tau(s_2),$$

so $\tau(s_1) < \tau(s_2)$.

The population can be split into three subgroups defined by conditions under which they choose to take deworming, and each subgroup weakly gains from the higher subsidy:

1. "Always-Takers," who deworm at either subsidy level. When the subsidy increases they are strictly better off, regardless of the tax change, since they are receiving a net transfer from the government in the form of the higher subsidy.
2. "Never-Takers," who deworm at neither subsidy level. As a result, they are only affected by the tax changes. If the tax rate falls, they are better off because their consumption choice set has expanded.
3. "Compliers," who deworm only with the higher subsidy level. By revealed preference, ignoring the tax change, they are no worse off deworming with a higher subsidy than they were by not deworming with a lower subsidy. As a result, they are strictly better off if the tax rate falls.

□

While the higher subsidy level s_1 is Pareto-improving relative to s_2 under the conditions outlined above, it may also be preferred by a social planner under much weaker conditions, including possibly the case of no epidemiological externalities, i.e., $x_{EXT} = 0$.

The unusual dataset we use in this paper, combined with Kenyan public finance data, provide estimates for each of the parameters in equation 1, allowing us to assess the welfare impacts of different subsidy policies. The take-up results in Kremer and Miguel (2007) provide estimates of $\theta(s)$ at multiple subsidy levels (including the no subsidy case). The long-run follow-up data provides estimates of both direct (x) and externality (x_{EXT}) deworming impacts on hours worked, as well as wages (w). Government of Kenya data contains information on tax rates (τ) and interest rates (r).

We next show that the effect of higher health investment subsidies on the tax rate is magnified at higher levels of existing government expenditures.

Proposition 2: **When Equation 1 holds, given two subsidy levels ($\boldsymbol{s_1 > s_2}$) and two levels of government expenditure ($\boldsymbol{G_a > G_b}$), then the resulting decrease in tax rates (at the higher subsidy level) is larger when the level of other government expenditures is higher, i.e.,**

$$\boldsymbol{\tau(s_2|G_a) - \tau(s_1|G_a) > \tau(s_2|G_b) - \tau(s_1|G_b).}$$

Proof:

Using the definition of the tax rate and twice differentiating yields:

$$\frac{\partial^2 \tau(s)}{\partial s \partial G} = -\frac{\theta'(s)(x + x_{EXT})}{[E + \theta(s)(x + x_{EXT})]^2 * \frac{w}{(1-r)}} < 0.$$

The change in take-up with respect to the subsidy ($\theta'(s)$) is positive, as are all other terms, implying that the decrease in tax rates is larger given higher outside government expenditures, *G*. □

Note that in the extreme case where $\theta(s_1) = \theta(s_2)$ – in other words, take-up of the health treatment is not responsive to the subsidy – a higher subsidy does not increase the endowment of total healthy time, and therefore does not increase production in the economy or tax revenue. As a result, the lower subsidy is associated with a lower tax rate, and, to the extent that taxes are distortionary, associated with greater social welfare. Furthermore, if all individuals are homogeneous, and all choose to deworm at the exact same subsidy level $s^* > 0$, and there are no health externalities ($x_{EXT} = 0$), then that subsidy level will generate more revenue than no subsidy if $s^* < \frac{1}{1-r} wx\tau(0)$. This is a special case of equation 1.

3. Study Background

This section describes the context, the deworming program, and the follow-up survey, including our respondent tracking approach, and sample summary statistics.

3.1 The context

We study the impact of a school-based deworming program in Busia district, a densely-settled farming region of rural western Kenya adjacent to Lake Victoria that is somewhat poorer than the Kenyan average. Survey respondents originally attended rural schools, and at the time of the follow-up data collection were young adults, mainly in their early twenties.

Agriculture in Busia is rain-fed with two cropping seasons per year, and there are few draft animals. Unlike other parts of Kenya, where many farmers have turned to growing vegetables for local markets or flowers, coffee or tea for international markets, there is little intensification of agricultural production, with only 1% of respondents (in the control group) growing cash crops, as discussed below. The Lewis (1954) model assumption that young adults working in traditional family agriculture receive a share of output rather than their marginal product is plausible in this context. Markets for agricultural land and labor exist in this area but are relatively thin, as evidenced by our sample, in which the majority of respondents engage in agricultural activity but less than 1% do paid agricultural work. Young adults have the option of staying on their parents' farms or leaving home to seek paid work, to start businesses, or, if female, to marry (Government of Kenya 1986). Sons typically inherit land from their parents, with many receiving *inter vivos* land transfers.

Unemployment and under-employment are considered major social problems in Kenya. At the time of survey, just over half of the sample (56% of the control group) work on family farms, primarily for subsistence. Among those who work in traditional agriculture, work hours are low at approximately 15 per week, but work hours are much higher among those in wage employment or self-employment (at approximately 40 hours per week). Average work hours in the last week is under 20 hours for the control group. Roughly one quarter of the sample is still in school. Fifteen percent of study participants are employed in wage labor and 10% are in non-agricultural self-employment. These proportions do not add to 100% since many respondents who work in agriculture also engage in other activities.[5] There are large differences in family and labor market patterns by gender, with nearly half of female respondents married, compared to only a quarter of males.

The health problem we examine, intestinal worm infections, is extremely widespread, with roughly one in four people globally infected with hookworm, whipworm, roundworm, or schistosomiasis (Bundy 1994, de Silva *et al*. 2003). Although light worm infections are often asymptomatic, more intense infections can lead to lethargy, anemia and growth stunting. Schistosomiasis can also have more severe consequences including enlargement of the liver and spleen. Disease burden estimates suggest that schistosomiasis alone accounts for up to 70 million disability-adjusted life years lost per year with thousands of deaths annually in Africa (Hotez and Fenwick 2009).

Treating worm infections (once to twice per year) can improve child appetite, growth and physical fitness (Stephenson *et al.* 1993), and reduce anemia (Guyatt *et al.* 2001, Stoltzfus *et al.* 1997). It also can strengthen children's immunological response to other infections, potentially producing broader health benefits, such as reduced infection prevalence with *Plasmodium*, the malaria parasite (Kirwan *et al.* 2010). Chronic parasitic infections in childhood generate inflammatory (immune defense) responses and elevated cortisol levels that lead substantial energy to be diverted from growth, and there is mounting evidence that this can produce adverse health consequences throughout the life course, including organ damage, atherosclerosis, impaired intestinal transport of nutrients, and cardiovascular disease (Crimmins and Finch 2005).

Geohelminth eggs are spread when children defecate in the "bush" surrounding their home or school, while the schistosomiasis parasite is spread through contact with infected fresh water. Treatment externalities for schistosomiasis are likely to take place across larger areas than

[5] We elaborate on the sample summary statistics below.

geohelminth externalities due to the differing modes of disease transmission, since the water-borne schistosome may be carried considerable distances by stream and lake currents, and the snails that serve as its intermediate hosts are themselves mobile.

Previous work in the study sample shows that deworming treatment led to large medium-run gains in school attendance and health. Due to worms' infectious nature, sizeable externality benefits also accrued to the untreated within treatment communities and to those living near treatment schools (Miguel and Kremer 2004), as well as to younger children in the treatment communities. Ozier (2010) shows that children who were 0 to 2 years old when the deworming program was launched and lived in the catchment area of a treatment school experienced large cognitive gains ten years later, with average test score gains of 0.3 standard deviation units, equivalent to over half a year of school learning in his sample. Bleakley (2007, 2010), examines the impact of a large-scale deworming campaign in the U.S. South during the early 20th century, by comparing heavily infected versus lightly infected regions over time in a difference-in-difference design. He finds that deworming raised adult income by roughly 17%, and, extrapolating these findings to the higher worm infection rates found in tropical Africa, estimates that deworming in Africa could lead to income gains of 24% (similar to our estimated earnings gains for wage earners discussed below).[6]

3.2 The Primary School Deworming Program (PSDP)

[6] There has been a debate in public health and nutrition about the cost-effectiveness of deworming (see Taylor-Robinson, Jones and Garner 2007). Early work by Schapiro (1919) using a first-difference research design found wage gains of 15-27% on Costa Rican plantations after deworming. Weisbrod et al (1973) document small correlations between worm infections and labor productivity and test scores in St. Lucia. Bundy *et al*. (2009) argue that many studies understate deworming's benefits since they fail to consider externalities by using designs that randomize within schools; focus almost exclusively on biomedical criteria and ignore cognitive, education and income gains; and do not address sample attrition. The current paper attempts to address these three concerns. Beyond Miguel and Kremer (2004) and the current paper, Alderman *et al.* (2006b) and Alderman (2007) also use a cluster randomized controlled design and find large positive child weight gains from deworming in Uganda.

In 1998, the non-governmental organization (NGO) ICS launched the Primary School Deworming Program (PSDP) to provide deworming medication to children enrolled in 75 primary schools. The schools participating in the program consisted of 75 of the 89 primary schools in Budalangi and Funyula divisions in southern Busia (with 14 town schools, all-girls schools, geographically remote schools, and program pilot schools excluded), containing 32,565 pupils at baseline.

Parasitological surveys conducted by the Ministry of Health indicated that these divisions had high baseline helminth infection rates at over 90%. Using modified WHO infection thresholds (described in Brooker *et al.* 2000a), over one third of children in the sample had moderate to heavy infections with at least one helminth at the time of the baseline survey, a high but not atypical rate in African settings (Brooker *et al.* 2000b, Pullan et al. 2011).[7]

The 75 schools involved in this program were experimentally divided into three groups (Groups 1, 2, and 3) of 25 schools each: the schools were first stratified by administrative sub-unit (zone), listed alphabetically by zone, and were then listed in order of pupil enrollment within each zone, and every third school was assigned to a given program group; supplementary appendix A contains a detailed description of the experimental design. The groups are well-balanced along baseline demographic and educational characteristics (Table 1, Panel A).[8]

Due to the NGO's administrative and financial constraints, the schools were phased into the deworming program over the course of 1998-2001 one group at a time. This prospective and staggered phase-in is central to this paper's econometric identification strategy. Group 1 schools

[7] The 1998 Kenya Demographic and Health Survey indicates finds that 85% of 8 to 18 year olds in western Kenya were enrolled in school, indicating that our school-based sample is broadly representative of children in the region.
[8] The same balance on predetermined characteristics is also evident among the subsample of respondents no longer enrolled in school and among those currently working for wages (see appendix tables A1, A2 and A3), two subsamples that feature in the analysis below. Miguel and Kremer (2004) present a fuller set of baseline covariates for the treatment and control groups. Appendix figure A1 summarizes the research design.

began receiving free deworming treatment in 1998, Group 2 schools in 1999, while Group 3 schools began receiving treatment in 2001. The project design implies that in 1998, Group 1 schools were treatment schools while Group 2 and 3 schools were the control schools, and in 1999 and 2000, Group 1 and 2 schools were treatment and Group 3 was control, and so on.

The NGO typically requires cost sharing, and in 2001, a randomly chosen half of the Group 1 and 2 schools took part in a program in which parents had to pay a small positive price to purchase the drugs, while the other half of Group 1 and 2 schools received free treatment (as did all Group 3 schools). Kremer and Miguel (2007) show that cost-sharing led to a sharp drop in deworming treatment, by nearly 60 percentage points. In 2002 and 2003, all sample schools received free treatment.

Children in Group 1 and 2 schools thus were assigned to receive 2.41 more years of deworming than Group 3 children on average (Table 1, Panel A), and these early beneficiaries are what we call the deworming treatment group below. We focus on a single treatment indicator rather than separating out effects for Group 1 versus Group 2 schools since this simplifies the analysis and because we sometimes lack statistical power to distinguish effects across these groups. The fact that the Group 3 schools eventually did receive deworming treatment will tend to dampen any estimated treatment effects relative to the case where the control group was never phased-in to treatment. In other words, a program that consistently dewormed some children throughout childhood while others never received treatment might have even larger impacts. Note, however, that several cohorts "aged out" of Group 3 primary schools (i.e., graduated or dropped out) before treatment was phased-in, meaning that they received little or no deworming and are closer to a more traditional control group.

Deworming drugs for geohelminths (albendazole) were offered twice per year and for schistosomiasis (praziquantel) once per year in treatment schools. We focus on intention-to-treat (ITT) estimates, as opposed to actual individual deworming treatments, in the analysis below. This is natural as compliance rates are high. To illustrate, 81.2% of grades 2-7 pupils scheduled to receive deworming treatment in 1998 actually received at least some treatment. Absence from school on the day of drug administration was the leading reported cause of non-compliance. The ITT approach is also attractive since previous research showed that untreated respondents within treatment communities experienced significant health and education gains (Miguel and Kremer 2004), complicating estimation of treatment effects on the treated.

3.3 Kenya Life Panel Survey (KLPS)

The Kenyan Life Panel Survey (KLPS-2) was collected during 2007-2009, and tracked a representative sample of approximately 7,500 respondents who had been enrolled in primary school grades 2-7 in the 75 PSDP schools at baseline in 1998.[9]

Survey enumerators traveled throughout Kenya and Uganda to interview those who had moved out of local areas.[10] As time progressed and the pace of locating respondents slowed, a representative (random) subsample containing approximately one quarter of still-unfound target respondents was drawn. Those sampled were tracked “intensively” (in terms of enumerator time and travel expenses) for the remaining months, while those not sampled were no longer actively tracked. We re-weight those chosen for the “intensive” sample by their added importance to maintain the representativeness of the sample. As a result, all figures reported here are

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

[10] Baird, Hamory and Miguel (2008) further discuss the respondent tracking approach.

“effective” tracking rates (ETR), calculated as a fraction of those found, or not found but searched for during intensive tracking, with weights adjusted appropriately. This is analogous to the approach in the Moving To Opportunity study (Kling *et al.* 2007, Orr *et al.* 2003). The effective tracking rate (ETR) is a function of the regular phase tracking rate (RTR) and intensive phase tracking rate (ITR) as follows:

$$ETR = RTR + (1 - RTR)*ITR \qquad \text{(eqn. 2)}$$

Overall, the RTR in KLPS-2 is 65.0% and the ITR is 62.1%, which implies that 86.7% of respondents were effectively located by the field team, with 82.7% surveyed while 4% were either deceased, refused to participate, or were found but were unable to be surveyed (Table 1, Panel B). The effective survey rate among those still alive is 86%. These are high tracking rates for any age group over a decade, and especially for a highly mobile group of adolescents and young adults, and they are on par with some well-known panel survey efforts in less developed countries, such as the Indonesia Family Life Survey (Thomas, Frankenberg and Smith 2001, Thomas *et al.* 2012).[11] Reassuringly, survey tracking rates are nearly identical and not significantly different in the treatment and control groups.

4. Deworming impacts on health, labor hours and consumption

This section presents the estimation strategy, and discusses deworming impacts on health, education, hours worked and meal consumption.

4.1 Estimation strategy

[11] Other successful longitudinal data collection efforts among African youth are Beegle *et al.* (2011) and Lam et al (2008). Pitt, Rosenzweig and Hassan (2011) document high tracking rates in Bangladesh.

The econometric approach relies on the PSDP's prospective experimental design, namely, the fact that the program exogenously provided individuals in treatment (Group 1 and 2) schools two to three additional years of deworming treatment. We also adopt the approach in Miguel and Kremer (2004) and estimate the cross-school externality effects of deworming. Exposure to spillovers is captured by the number of pupils attending deworming treatment schools within 6 kilometers; conditional on the total number of primary school pupils within 6 kilometers, the number of treatment pupils is also determined by the experimental design.

The dependent variable is an outcome (such as hours worked in the last week), $Y_{ij,2007\text{-}09}$, for individual i from school j, as measured in the 2007-09 KLPS-2 survey:

$$Y_{ij,2007\text{-}09} = a + bT_j + c_1N_j^T + c_2N_j + X_{ij,0}'d + e_{ij,2007\text{-}09} \qquad \text{(eqn. 3)}$$

The labor market outcome is a function of the assigned deworming program treatment status of the individual's primary school (T_j), and thus this is an intention to treat (ITT) estimator; the number of treatment school pupils (N_j^T) and the total number of primary school pupils within 6 km of the school (N_j); a vector $X_{ij,0}$ of baseline individual and school controls; and a disturbance term $e_{ij,2007\text{-}09}$, which is clustered at the school level. The $X_{ij,0}$ controls include school geographic and demographic characteristics used in the "list randomization" for the PSDP, the student gender and grade characteristics used for stratification in drawing the KLPS sample, the pre-program average school test score to capture school academic quality, the 2001 cost-sharing school indicator, as well as controls for the month and wave of the interview. The externality results are unchanged if we focus on the proportion of local pupils who were in treatment schools as the key spillover measure (i.e., N_j^T / N_j, results not shown), but we opt with the current specification for comparability to Miguel and Kremer (2004).[12] Miguel and Kremer (2004) also

[12] Several other econometric issues related to estimating externalities are discussed in Miguel and Kremer (2004).

separately estimate effects of the number of pupils between 0-3 km and 3-6 km, but since the analysis in the current paper does not generally find significant differences in impacts across these two ranges, we focus on 0-6 km for simplicity.

The main coefficients of interest are b, which captures gains accruing to deworming treatment schools, and c_1, which captures spillover effects of treatment for nearby schools. Bruhn and McKenzie (2009) argue for including variables used in the randomization procedure as controls in the analysis, which we do, although the coefficient estimates on the treatment indicator are robust to whether or not we do, as expected given the research design. Results are also robust to accounting for the cross-school spillovers. In fact, accounting for externalities tends to increase the b coefficient estimate; in other words, a failure to account for the program treatment "contamination" generated by spillovers dampens the "naïve" difference between treatment and control groups (and also leads the researcher to miss a second dimension of program gains, the spillovers themselves).

Given the large baseline differences in marital and labor market choices for men and women, we explore heterogeneity by gender throughout, and also estimate heterogeneous treatment effects by age and by the local disease environment for certain outcomes. Theoretically, the sign of the interaction of treatment with the local level of serious worm infections is ambiguous, and the effect of the program at higher levels of initial disease prevalence need not be monotonic. This is because areas with higher prevalence will typically have conditions more conducive to transmission of the disease; re-infection is thus likely to occur more quickly in these areas and hence the impact of treatment could potentially be smaller in these areas than in areas where it takes longer for re-infection to occur. Given this theoretical ambiguity, and the lack of strong evidence in the data that interaction terms or higher order polynomial externality terms are justified, we focus both in the theoretical model and in the econometric analysis on specifications in which T_j and N_j^T are additively separable.

The interpretation of coefficient estimates on the externality term (N_j^T) is complicated by the fact that those who benefit from cross-school spillovers (in terms of reduced infection intensity) themselves generate positive spillovers for others, as a result of the reduction of their own worm burden. While in the short-run (as in Miguel and Kremer 2004) the cross-school spillovers are likely to fairly accurately capture the magnitude of these externality impacts, over time the infection "feedback" effects generated across nearby communities would lead us to understate the magnitude of cross-school externality magnitudes as they converge to a common local infection rate, as predicted by standard epidemiological models such as those in Anderson and May (1991). This is a form of "contamination" of the externality "treatment". As a result, it is reasonable to interpret the c_1 estimate as a lower bound on the true magnitude of long-run cross-school externality effects.

4.2 Deworming Impacts on Health, Fertility and Education

In this section, we consider health, fertility and education outcomes in the follow-up survey. Of course, since health, education, marital choices, and income levels may all affect each other, the results in this section should not be interpreted strictly as all reflecting the direct health impact of deworming, but rather the result of a cumulative process of interaction among them. To be clear, we do not expect that deworming treatment as children would have a direct impact on respondents' worm loads as young adults a decade later, since worms' average lifespan in the human body is only one to three years (Anderson and May 1991, Bundy and Cooper 1989).

Our best summary measure of health status is a self-report in the follow-up survey. Many studies have found that self-reported health reliably predicts actual morbidity and mortality even when other known health risk factors are accounted for (Idler and Benyamini 1997, Haddock et al. 2006, Brook et al. 1984). Respondent reports that their health was "very good" rose by 4.1 percentage points (s.e. 1.8, significant at 95% confidence, Table 2, Panel A), on a base of 67.3% in the control group. While we cannot reject equal effects, gains are slightly larger for females than for males. There is some suggestive evidence of a positive externality effect, with an increase of 2.8 percentage points for each additional 1,000 treatment group pupils located within 6 km of the school, although the effect is not significant at traditional confidence levels.

There is no significant impact on current marital status or on respondent pregnancies through the 2007-09 survey for the full sample, although point estimates are negative. A noteworthy pattern is that female respondents are nearly twice as likely to be married as males (44.9% versus 25.5%) and have had nearly twice as many pregnancies as male respondents' partners (1.28 versus 0.71 on average), reflecting the fact that Kenyan women tend to marry and have children at younger ages than men. These family circumstances may also partially account for the lower average number of work hours reported by females, as discussed below.

We also examine miscarriage rates as a further proxy for adult health. Miscarriage rates are known to be highly sensitive to general maternal health and nutritional status (Hotez 2009). Among females, deworming significantly reduced miscarriage rates, by 2.8 percentage points (s.e. 1.3) on a base of 3.9 percent, a very large reduction (Table 2, Panel A). The externality effect is also large and negative at -1.5 percentage points (s.e. 0.7) per additional 1000 treatment pupils within 6 km, providing evidence of positive health spillover gains. The reduction in miscarriages for females is robust to controlling for age at pregnancy and respondent's socioeconomic background (results not shown).

The comparison across female and male respondents offers a useful "placebo check" in this case. Two possible (and not mutually exclusive) explanations for the health gains reported in Table 2, Panel A are, first, that they reflect the persistent health gains generated by child deworming investments, and second, that they are mainly a result of higher respondent living standards in adulthood (as documented below). While the female respondents potentially experience both of these channels, the miscarriage patterns among male respondents' partners should only reflect the adult living standards channel. The estimated effect of deworming treatment for males on their partners' miscarriage rate is essentially zero (estimate -0.001, s.e. 0.005), in contrast to the very large reduction in miscarriage risk reported among female respondents, suggesting that the persistent health gains generated by child deworming are important.

Deworming treatment also leads to some persistent gains in educational outcomes. The total increase in years of school enrollment in treatment relative to control schools is 0.279 years (s.e. 0.147, significant at 90% confidence – Table 2, Panel B). Given that the school enrollment data misses out on attendance impacts, which are sizeable, a plausible lower bound on the total increase in time spent in school induced by the deworming intervention is the 0.129 gain in school participation from 1998-2001 plus the school enrollment gains from 2002-2007 (multiplied by average attendance conditional on enrollment), which works out to nearly 0.3 additional years of schooling. Despite the sizeable gains in years of school enrollment, there are no significant impacts on total grades of schooling completed (0.153, s.e. 0.143). This is because the increased time in school is accompanied by increased grade repetition (0.060, s.e. 0.017, significant at 99%)..

There is suggestive evidence that deworming improved human capital: English vocabulary knowledge (collected in 2007-09) is slightly higher in the treatment group (0.076 standard deviations in a normalized distribution, s.e., 0.055), and the passing rate improved on the key primary school graduation exam, the Kenya Certificate of Primary Education (4.8 percentage points on a base of 50.5%, s.e. 3.1, Table 2, Panel B), although neither is significant. If we focus on the subsample of respondents who are no longer in school – a natural subsample of interest in the analysis of labor market impacts to follow – we find larger test score impacts. Note that there is no statistically significant difference in the proportion of treatment and control group students who are currently out-of-school at the time of the follow-up survey (both are at 75%), nor are there any significant differences across these groups along observable dimensions (appendix Table A1) nor is there differential selection into the out-of-school subsample along a observable characteristics across the treatment and control groups (appendix Table A2), including by gender, making this a reasonable comparison. The average gain on English vocabulary in the out-of-school subsample is 0.107 s.d. units (s.e. 0.052, significant at 95%), with large associated externality gains of 0.149 per 1,000 additional treatment pupils within 6 km (s.e. 0.047, significant at 99%). There is also an increase of 6.1 percentage points in the KCPE passing rate (on a base of 41.3%), and once again significant externalities (8.3 percentage points, significant at 99%). There is no clear gender pattern, with males showing somewhat larger gains in English vocabulary and females on passing the primary school graduation exam.

Beyond the moderate test score gains documented here, it is also possible that the increased amount of time the treatment group actually spent enrolled in school might yield labor market returns through dimensions other than human capital (as measured on tests). These might include improved social or other non-cognitive skills (Heckman, Stixrud, and Urzua 2006), such as greater ability to follow rules or show up regularly and on time. However, we do not have information on these individual attributes in our dataset.

4.3 Deworming Impacts on Work Hours and Meal Consumption

We first estimate deworming impacts on total hours worked in the last week. Among all KLPS respondents, deworming increased mean hours worked by 1.53 hours (s.e. 1.03 hours, Table 3, Panel A) on a control group mean of 18.4 hours. While this mean impact is not statistically significant at traditional confidence levels, a Kolmogorov-Smirnov test rejects equality of the hours worked distribution between the treatment and control groups at 90% confidence (p-value=0.09). Here and below we estimate impacts separately for males and females given the different marital and fertility patterns, family constraints, labor market opportunities and leisure options available by gender in Kenya. The hours worked effects are much larger among males, with an increase of 3.40 hours (s.e. 1.39) on a control group mean for males of 20.3 hours, a 16.7% increase significant at 95% confidence. Effects for all females are small and not statistically significant, at 0.29 hours (s.e. 1.34).

In contrast, there is no significant change in the proportion in the treatment group working at all (greater than zero hours in the past week), which is roughly 70% of those not still in school. There is thus a considerable degree of "non-activity" for a young adult population.[13] In the full sample, females are somewhat more likely to be classified as non-active which is likely related to the fact that more than three quarters of out-of-school females have had at least

[13] In the control group, 23% of females are non-active, compared to 18% of males, where we define "non-activity" as those not in school or employed in agriculture, wage labor or self-employment.

one pregnancy. However, note that some females are engaged in home production or time spent on child-rearing activities that were not collected in the survey and thus not classified as work here, and this issue is likely to lead us to understate total work hours for many female respondents. This possible mis-measurement of total work hours, which is likely to be particularly important for females, provides another reason to conduct the analysis separately for males and females.

In terms of externality effects for total hours worked, the estimated impact is quite large, at 1.71 hours (s.e. 1.43) per 1,000 additional treatment pupils within 6 km, but it is not statistically significant at traditional confidence levels. These imprecisely estimated externality effects are large in magnitude: an increase of 1,000 local treatment pupils is approximately one standard deviation in the local density of treatment school pupils (Table 1), and is equivalent to treating over 20% of the local primary school population.[14]

In the fiscal impact calibrations that follow (in section 5), we focus on these full sample results. However, it is also natural to focus on the out-of-school subsample when considering labor outcomes. As mentioned above, the proportion of respondents who are out-of-school is nearly identical across the treatment and control groups (Table 2, Panel B) and there is no evidence of differential selection along observables across the treatment and control groups (appendix tables A1 and A2), lending credibility to the analysis that follows. In the out-of-school subsample for both genders, the treatment effect on total hours worked is 2.74 hours per week (s.e. 1.30, significant at 95%), an increase of 12.5% on a control group base of 22.0 hours. Gains are once again larger for males, at 4.14 hours (s.e. 1.95, significant at 95% confidence), and effects are also relatively large for out-of-school females, at 2.13 (s.e. 1.47), though still not

[14] The results are very similar if we use the proportion of local pupils in treatment schools rather than the number as the externality term: the point estimate on this proportion externality term in the full sample analysis is 7.65 (s.e. 5.85) for total hours worked, and 0.311 (s.e. 0.093) for meals eaten (results not shown).

significant. These magnitudes are substantial: the average difference in weekly work hours in the United States versus France – two countries with starkly different labor market regulations – is roughly 5 hours (OECD 2011).

The next focus is on a living standards measure. We do not have complete data from a consumption module[15], but did collect data on the number of meals respondents consumed, and find large deworming impacts. Deworming treatment respondents miss 0.096 fewer meals per day (s.e. 0.028, 99% confidence, Table 3, Panel B) than the control group, and the externality impact is also large and positive (0.080, s.e. 0.023, 99% confidence). Mirroring the hours worked results, the increase in meals eaten is somewhat larger for males, at 0.127 (s.e. 0.041) than for females (0.051 meals, s.e. 0.043). Impacts are similar though slightly larger for the out-of-school subsample, with an overall increase of 0.103 meals (s.e. 0.029), and an externality effect of 0.101 (s.e. 0.032) per 1,000 additional treatment pupils within 6 km. It is worth noting that some additional calories may be required to allow for the increased physical effort at work (given the gains in labor hours), as suggested by Deaton and Drèze (2008).

An interesting methodological question is the extent to which the results would differ had the survey data collection not included efforts to track respondents living outside the original study district. Perhaps surprisingly, individuals found in the "intensive" tracking phases do not differ significantly over a range of mean observable characteristics (see appendix Table A4). We also cannot reject that treatment effect estimates are equal in the regular tracking subsample and the intensive tracking subsample for either outcome in Table 3 (results not shown). For total hours worked in the last week (in the full sample), we cannot reject that the treatment effect

[15] A consumption expenditure module was collected as a pilot for roughly 5% of the KLPS-2 sample during 2007-2009, for a total of 255 complete surveys. The estimated treatment effect for total per capita consumption is near zero and not statistically significant (-$13, s.e. $66, trimming the top 2%), but the confidence interval is large and includes substantial gains, since average per consumption consumption is $584.

estimate is the same if we exclude the intensive tracking subsample, but the estimated effect is smaller for the number of meals eaten if the intensive subsample is excluded (not shown).

Beyond gender, we explored heterogeneity along several other dimensions. We first find no evidence that the younger cohorts in the sample (those in grades 2 through 4 at baseline) showed larger gains in either work hours or meals eaten than the older cohorts initially in grades 5 through 7 (appendix table A5) [16]. We subsequently examined impacts in geographic zones within the sample with different levels of baseline worm infection rates, but do not find significantly different treatment effects (appendix table A5), nor when we break down baseline infection rates by geohelminths versus schistosomiasis (not shown). As discussed above, it is theoretically ambiguous whether treatment effects should be larger or smaller in areas with higher baseline infection rates (since areas with higher prevalence will typically have conditions more conducive to transmission of the disease and thus more rapid re-infection). Nor are there significant interaction effects between the treatment indicator and the externality term (not shown). The lack of significant interaction effects between the treatment indicator and baseline infection rates, and with the externality term, provides a rationale for the theoretical specification in section 2 (and that is calibrated in section 5 below), and for the core econometric specifications throughout the paper. Note that we use the zonal-level baseline infection rate, rather than individual-level data (which was not collected at baseline for the control group for ethical reasons); using zonal averages is likely to introduce some measurement error and attenuation bias, and thus this interaction effect may understate the true extent of differential impacts in high worm infection areas.

[16] Another noteworthy finding in appendix table A5 is the consistently negative coefficient estimates on the 2001 cost-sharing indicator variable for both hours worked and meals eaten. Recall that cost-sharing led to large drops in deworming take-up, so the fact that estimates are of opposite sign compared to the direct treatment and the externality terms is reassuring that, in fact, higher levels of deworming treatment do lead to improved hours worked and meal consumption outcomes.

Beyond the finding of statistically significant deworming externality effects on meals consumed, and suggestive externality evidence on hours worked, in Table 3, a further word is in order regarding deworming externalities. Including the two outcome measures in Table 3, we explore the impact of deworming on 42 distinct dependent variables and/or subsamples in Tables 2, 3, 5 and 6 in this paper. One simple way to assess the effect of externalities is to test whether they externality effect has the same "sign" as the direct deworming treatment estimate. To the extent that these are consistent, it would provide more confidence that the deworming program was actually generating externalities; in contrast, if the externality effects were mainly "noise", there would be no reason to expect the coefficient estimates on the externality term to be related to the direct deworming treatment effects. We find that in 36 of the 42 specifications, the sign of the treatment school indicator and the externality term are the same. This is extremely unlikely to occur by chance: in the case where the externality effect was pure "noise", then the likelihood of a "match" between the two terms would be distributed as a binomial distribution with p=0.5, 36 of 42 cases would have the same sign in fewer than one in 10,000 cases (available upon request).

5. The Fiscal Impact of Deworming Subsidies

The estimated impacts of deworming on total work hours, combined with earlier experimental estimates of the sensitivity of deworming take-up to price (Kremer and Miguel 2007), allow us to compute the fiscal impacts of deworming subsidies within the model framework developed in section 2 above. In summary, we find that the additional government revenues generated by increased work hours caused by deworming subsidies are far greater than the direct costs of the subsidies, suggesting that deworming subsidies are Pareto-improving in this case, and providing evidence for an "expenditure Laffer curve" for this health investment.

Recall from section 2 that higher levels of deworming subsidies are Pareto-improving if the increased cost of the higher subsidies is outweighed by the future increase in tax revenues. For more realistic projections about the future path of earnings and thus revenues, we allow for earnings to evolve over time, and use this modified version of equation 1:

$$s_1\theta(s_1) - s_2\theta(s_2) < \sum_{t=0}^{t=\infty} r^t w_t \left[\left(\theta(s_1) - \theta(s_2)\right)(x + x_{EXT})\right]\tau(s_2) \qquad \text{(eqn. 1a)}$$

To compute the left hand side of this expression, 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 p = \$1.42 per person treated. Under partial deworming subsidies, as implemented in the 2001 cost-sharing program in our sample, 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 4, Panel A, we compare these subsidy levels with the default case of no subsidies, namely, $s_2 = 0$.

To compute the right hand side, we use a combination of estimates from this paper, as well as other information on the Kenyan economy and public finances. The hours worked estimates (Table 3) indicate that treatment group males work 3.40 more hours per week, whereas the treatment effect estimate for women is near zero. In this exercise, we thus set the average hours work gain (x in the model in Section 2) equal to 3.40/2. The estimated increase in work hours due to epidemiological externalities is 1.71 hours/week per 1,000 neighboring pupils dewormed (Table 3), and we combine this information with the total density of treatment schools (Table 1) to determine x_{EXT}. Since this estimate is not statistically significant at conventional confidence levels, we first present the calibration assuming there is no epidemiological externality ($x_{EXT} = 0$), and then assuming the externality has the estimated magnitude. At the time of writing, the Government of Kenya pays 11.85% interest on its sovereign debt and inflation is approximately 2%.[17] As a result, we set the real cost of capital r, at 9.85%. We assume that the sample population begins working ten years after they first began receiving deworming and retires after 40 years of work.[18] We use the pattern of lifecycle earnings reported in Knight, Sabot and Hovey's (1992) study of Kenyan workers, and assume the initial starting wage w is \$0.16 per hour (Suri 2011)[19]. Kenyan taxes (mainly in the form of consumption taxes) absorb roughly 16.5% of GDP so we set $\tau(0)$ = 16.5%.[20]

For the full subsidy, the average cost per-person is \$1.07 (Table 4, Panel A). Assuming that $x_{EXT} = 0$, then $\sum_{t=10}^{t=50} r^t w_t \left[\left(\theta(p) - \theta(0)\right)x\right]$ = \$55.26, implying that individuals gain an average of \$46.10 in take-home pay and the NPV of government revenue increases by \$9.16 per person (Panel B), far greater than \$1.07. If deworming also generates positive externalities for others in the area – as suggested by the hours worked and especially the meals eaten results – the earnings gains are much larger, with the per-capita increase in government revenue rising to \$52.76 (Panel C).

The calculations above assume an annual discount rate of 9.85%. An alternative approach to assessing the attractiveness of deworming as an investment is to compute the internal rate of return (IRR) for a government policymaker. Focusing solely on government expenditures and revenues, the IRR is the interest rate for which $s\theta(s) = \sum_{t=10}^{t=50} r^t w_t \left[\left(\theta(s) - \theta(0)\right)(x +\right.$

[17] See http://www.centralbank.go.ke/securities/bonds/manualresults.aspx and World Bank Development Indicators.
[18] This ten year gap roughly corresponds to the time elapsed from the start of the PSDP until the KLPS2 follow-up survey (2007-09). This is a conservative assumption since some respondents began working before KLPS2.
[19] We use the weighted average of the statistically significant coefficient estimates in Knight, Sabot and Hovey's (1992) Table 2, rows 1 and 2. Suri (2011) presents wage information from the Tegemeo Agricultural Monitoring and Policy Analysis Project.
[20] From World Bank Development Indicators, Kenyan government expenditures are roughly 19.5% of GDP, and from http://blogs.worldbank.org/africacan/three-myths-about-aid-to-kenya about 15% of government expenditure is financed from donors, thus 0.195*0.85=0.165.

$x_{EXT})\big]\tau(0)$. The IRR in the case of no health spillovers is 24.7% per annum for full subsidies, and with health spillovers it rises to 42.0% (Table 4, Panel D). Note that the linear separability of the direct treatment effect and the externality implies that the IRR is nearly identical in both the partial subsidy and the full subsidy cases. These are very high rates of return by any standard, and are far higher than the current interest rate faced by the government of Kenya.

An additional approach is to compute the social internal rate of return for a perfectly benevolent social planner, by solving for the interest rate that equates the NPV of the full social cost and all earning gains: $p\theta(s) = \sum_{t=10}^{t=50} r^t w_t \left[\big(\theta(s) - \theta(0)\big)(x + x_{EXT})\right]$. The social IRR is even higher than the IRR in terms of government revenue and is very high by any standard: with no health spillovers it is 42.5% for full subsidies, and with health spillovers 64.6%.

We make several conservative assumptions to reach these conclusions about deworming's fiscal impacts and rate of return. For instance, we do not incorporate recent evidence about the extent of positive deworming externalities: Ozier (2010) finds that living in a deworming treatment community early in life (age 0 to 2) leads to improved cognitive and academic performance ten years later. Individuals in other age groups could also potentially have benefited from the health spillovers of treatment. To the extent that these human capital gains among other populations not in the baseline PSDP sample would also generate higher earnings and an improved quality of life more broadly, the estimated fiscal impacts presented in this section are lower bounds on the true returns to deworming.

6. Deworming impacts on other labor market and productive activities

6.1 Work hours by employment sector

We document the impact of deworming on total hours worked in Table 3 above, but the break-down by economic sector, and along the extensive versus intensive margins, generates further insights. For the full sample, there is no significant change in working some labor hours (hours > 0), with a point estimate of 0.004 (s.e. 0.022, Table 5, Panel A). Both males and females in the control group participate in work at the same rate, 68%. This proportion increases markedly for males in the deworming treatment group by 5.0 percentage points (s.e. 2.6), but not for females, further indication that child health investments translate into different impacts by gender. Among those working at least some hours in the last week, the gains in total hours worked are once again concentrated among males, with an increase of 3.38 hours (s.e. 1.56).

The breakdown across the three main economic sectors – agriculture, wage employment, and non-agricultural self-employment – yields a more nuanced picture by gender.[21] The increase in labor hours for males is positive along the extensive and intensive margins in all three sectors (with statistically significant increases in total hours worked in both agriculture and self-employment, Table 5, Panel B). These patterns are shown graphically in Figure 1, where the marked shift in work hours among those working for wages (panel C) and in self-employment (panel D) are particularly apparent.

There is a shift out of agricultural work and into non-agricultural self-employment for women. Hours in agriculture (which includes work in crop cultivation as well as livestock) fall by 1.29 hours (s.e. 0.57, Table 5, Panel B) while hours worked in non-agricultural self-employment rise by 1.82 (s.e. 0.80, Panel D), and both of these effects are significant at 95% confidence. The different impacts for women may reflect the fact that it is costlier for women to increase their total work time than is the case for men, perhaps due to childcare constraints and a

[21] Appendix table A6 presents the proportions of respondents working in multiple economic activities, and shows that a sizeable share of respondents combine work in agriculture with some additional economic activities.

greater burden of home production and chores, leading them to substitute away from lower productivity agricultural activities and into small-scale businesses (and as we will show below, into more production of cash crops). Women tend to set up small retail businesses (76% of the female self-employed are in retail), while the male self-employed work in a wider range of self-employed activities, ranging from retail (44%), fishing (20%), small manufacturing (11%), and passenger transport (9%), several of which are likely to regularly take the respondents farther afield (e.g., fishing and transport). There is also evidence of externality impacts of deworming on hours worked in non-agricultural self-employment, with an increase of 1.15 hours per 1,000 treatment pupils within 6 km (s.e. 0.59, Table 5, Panel D).

One noteworthy pattern is the fact that there does not appear to be noticeable shifts into wage employment in the full sample or by gender, with the overall point estimate on the indicator for any wage work close to zero and not statistically significant (-1.4 percentage points, s.e. 1.6, Table 5, Panel C). We also cannot reject that the observable characteristics of wage earners are the same in the treatment and control groups (appendix table A3), nor is there differential selection along observables (including gender) into the wage earner sample across the treatment groups (appendix table A2). Taken together, this suggests that the comparison of labor productivity across wage earners in the treatment versus control groups is likely to be driven mainly by causal impacts of deworming rather than selection, and this motivates some of the analysis that follows.

In contrast, there are larger shifts between the agriculture and self-employed subsamples across the treatment and control groups, with substantial differences by gender (Table 5), making the comparison of productivity differences between the treatment and control groups within these sectors more difficult to interpret due to potential selectivity.[22]

6.2 Labor Productivity, earnings and profits

Just as we decompose the increase in overall hours into changes in hours in agriculture, hours working for wages, and non-agricultural self-employment, it is useful to separately estimate treatment impacts on output and productivity in each sector. Unlike hours worked, however, it is more challenging to measure labor productivity in a comparable fashion across economic sectors and activities. Measuring labor productivity is especially challenging in subsistence agriculture, which is the most common economic activity among our respondents.

Deworming treatment leads to pronounced shifts in the occupation of employment, out of relatively low-skilled and low-wage sectors into better paid and higher work intensity sectors (Table 6, Panel A). Deworming treatment respondents are three times more likely to work in manufacturing from a low base in the full sample of 0.005 (coefficient 0.010, s.e. 0.004). The gains among males are particularly pronounced at 0.018 (s.e. 0.007) from a male control group base of 0.007. Among wage earners, roughly 3% have manufacturing jobs, and the increase in manufacturing employment within this subsample is 7.2 percentage points (s.e. 2.4, appendix table A7). Survey responses indicate that the two most common types of manufacturing jobs in our sample are in food processing and textiles, with establishments ranging in size from small local corn flour mills in Busia district up to large blanket factories in Nairobi. On the flip side,

[22] More speculatively, the finding that treatment group respondents from academically stronger schools are significantly less likely to shift into both agriculture and self-employment (appendix table A2) – in other words, that there is some "negative" selection – suggests that if anything the estimated treatment effects on productivity within these sectors might be lower bounds on the actual causal effects.

casual labor employment falls significantly (-0.005, s.e. 0.002, Table 6), and there is suggestive evidence that domestic service work falls for females (-0.174, s.e. 0.110, appendix table A7).

Manufacturing jobs tend to be quite highly paid, with average real monthly earnings of 5,311 Shillings (roughly US$68), compared to casual labor (2,246 Shillings) and domestic services (3,047 Shillings, appendix table A8). Manufacturing jobs are also characterized by longer work weeks than average at 53 hours per week, in contrast to 43 hours for all wage earning jobs, indicating that these are high work intensity jobs. Workers in manufacturing jobs also tend to have relatively few work days missed due to poor health, at just 1.1 days (in the control group), compared to 1.4 days in the last month among all wage earning jobs. One explanation for this pattern that ties into our earlier labor supply findings is that health investments improve individuals' capacity to carry out physically demanding jobs, characterized by long work weeks and little tolerance of absenteeism, and thus allow them to access higher paid jobs such as those in manufacturing. Casual laborers typically do not have to commit to work a certain number of days in a week in advance, so the significant reduction in casual work is also consistent with the hypothesis that deworming helps people obtain jobs that require regular attendance.

There was extensive migration out of the study district, with 17.9% of respondents living in cities at the time of the follow-up survey (Table 6, panel A). However, there was no significant impact of deworming in urban migration (point estimate -0.001, s.e. 0.019), or in migration out of the study district overall (not shown). The fact that deworming led to increased manufacturing employment but not to urban migration might be surprising at first glance, but is perhaps most plausibly explained by the fact that work is but one of several leading reasons for respondent migration – the others being family reasons (i.e., marriage) and for schooling – as well as the existence of many small-scale manufacturing and other wage employment opportunities even in the largely rural study district.

Turning to direct measures of productivity, the impacts on wage earners' productivity are the most straightforward to measure. Here point estimates of the increase in earnings are larger than those of the increase in hours, consistent with the hypotheses that certain jobs require higher numbers of work hours, worked on a regular schedule, and that these jobs are better paying. It is also consistent with the idea that people adjust their work effort along intensive as well as extensive margins, as we find some evidence for wage gains. Treatment shifts the distribution of log wage earnings sharply to the right (Figure 2, Panel A).[23] In the regression analysis, we find that deworming treatment leads to higher log earnings (Table 6, Panel B), with a gain of 25.3 log points (s.e. 9.3, 99% confidence). Log wages computed as earnings per hour worked rise 16.5 log points (s.e. 11.7) in the deworming treatment group, indicating that nearly equal parts of the earnings gain works through increased hours worked (Table 3) and through greater productivity per hour worked.

The earnings result is robust to several alternative specifications. It changes little in response to trimming the top 1% of earners, so the result is not driven by outliers; to including a full set of gender-age fixed effects; and to including fixed effects for each of the "triplets" of Group 1, Group 2 and Group 3 schools from the list randomization (results not shown). Positive wage earnings impacts are similar in the larger group of respondents who have worked for wages at any point since 2007, where we use their most recent monthly earnings, with a mean impact of 0.211 (s.e. 0.072, Table 6, Panel B).

[23] Here and below we present real earnings measures that 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, however.

There is suggestive evidence for positive deworming externalities on earnings. While the coefficient estimate on the local density of treatment pupils is not significant at traditional confidence levels (19.9 log points, s.e. 16.8, in Table 6, Panel B), it reassuringly has the same sign as the main deworming treatment effect, and a substantial magnitude: an increase of 1,000 treatment school pupils, or roughly 20% of the local primary school population, would boost labor earnings by nearly 20 log points.

Recall that deworming does not seem to affect the likelihood that people become wage earners or the process by which observable characteristics influence the likelihood of becoming a wage earner. In Table 5, we found no evidence that deworming treatment respondents are more likely to be working for wages in the last month (Panel C, estimate -0.014, s.e. 0.016). We further cannot reject that the observable characteristics of wage earners, including academic performance measures, are the same in the treatment versus control groups (appendix Table A3), nor that there was differential selection across a range of other characteristics (appendix table A2). These factors point towards an interpretation of the difference in earnings between the deworming treatment and control groups primarily reflecting causal effects rather than selection.

A decomposition along the lines of Oaxaca (1973) – which uses mean earnings by occupation in the control group as a reference point – indicates that 74% of the increase in labor earnings for the treatment group can be accounted for by the occupational shifts documented in appendix table A7. We cannot determine if the shift in work hours led to a shift in occupations, or a shift in human capital led to a shift to occupations demanding more hours, and in any case it is not clear there is a meaningful distinction since hours and occupations are chosen jointly.

We next turn to productivity among those working in non-agricultural self-employment, and in agriculture. Profits among owners of non-agricultural businesses increase in the treatment group by similar percentage as the earnings increases among wage earners, but are estimated with less precision, partly because fewer people work in self-employment and partly because the underlying variance of reported profits is higher than that of reported wages (presumably due to a combination of stochastic variation and measurement error). The estimated deworming treatment effect on the monthly profits of the self-employed (as directly reported in the survey) is positive at 343 shillings (s.e. 306, Table 6, Panel C), although this 19% gain is not significant at traditional confidence levels.[24] Trimming the top 5% of self-reported profits results in a similarly sized but marginally significant treatment effect of 324 Shillings (s.e. 177, significant at 90% confidence). We also find suggestive evidence of impacts on the total number of employees hired (0.466 additional employees on a base of 0.188, s.e. 0.361).

We next construct a measure of total monthly non-agricultural earnings by summing wage earnings and self-employed profits, setting earnings or profits equal to zero for those not engaged in these activities (in order to avoid the possible selection into particular economic activities). We estimate a treatment effect of 99 Shillings (s.e. 104, Table 6 Panel D) on a base of 789 shillings in the control group, for a 12.5% increase; impacts are similar with a measure that trims that top 5% of profits (not shown). Most respondents either work solely in agriculture or are idle and thus have zero non-agricultural earnings, making this a particularly stringent test.

Unfortunately, we do not have a concrete measure of agricultural yield or output analogous to the wages or profits of those working in other sectors. In any case, measuring the on-farm productivity of an individual worker in the context of a farm where multiple household members (and sometimes hired labor) all contribute to different facets of the production process

[24] There are large, positive but not statistically significant impacts on a monthly profit measure based directly on revenues and expenses reported in the survey (not shown). We focus here on self-reported profits in the last month, which appear to be less noisy. De Mel, McKenzie and Woodruff (2009) have recently argued in favor of focusing on self-reported profits rather than computed profits in their work on small firms in less developed countries.

is difficult. We also lack sufficiently detailed information on farming choices to compute a reliable yield measure, and thus rely on several proxies for agricultural productivity. Deworming led more respondents working in agriculture to grow cash crops, with an increase of 2.0 percentage points (s.e. 0.8, significant at 95%, Table 6, panel E), and this effect is concentrated among female respondents (3.1 percentage points, s.e. 1.4). The increase in cash crop cultivation among women, as well as their shift in self-employment (Table 5), may reflect their desire to engage in higher productivity activities within their family and social constraints, which may complicate moves into manufacturing jobs or other lucrative opportunities.

We also find some suggestive evidence of increased adoption of intensified agriculture such as using fertilizer, hybrid seeds, or irrigation, with an increase of 3.2 percentage points (s.e. 2.6) on a base of 31.0 percent, although these effects are not statistically significant. Together with the finding that there was a doubling in the cultivation of cash crops, this suggests there were modest improvements in agricultural productivity. However, deworming did not significantly increase crop sales in the past year. This may be because respondents are consuming more of the grain they produced, as suggested by the increase in meals eaten.

In terms of living standards across sectors, there are large improvements in meals eaten for those working in all employment sectors, and mirroring the hours worked and productivity results, the gains are largest outside of agriculture (not shown).

7. Conclusion

Arthur Laffer (2004) famously suggested that tax cuts could potentially pay for themselves by endogenously generating more economic activity and thus more tax revenue. There has been considerable debate on whether this idea is relevant in practice, but there is no consensus that a case has been identified where Laffer's prediction holds. In this paper, we present evidence for the existence of an expenditure Laffer curve: certain expenditures, in particular child health investments, may generate sufficient future gains in labor market outcomes and government revenue to allow reductions in tax rates.[25]

The Kenya Primary School Deworming Program was experimentally phased-in across 75 rural schools between 1998 and 2001 in a region with high rates of intestinal worm infections, one of the world's most widespread diseases. As a result, the treatment group exogenously received an average of two to three more years of deworming treatment than the control group. A representative subset of the sample was followed up for roughly a decade through 2007-09 in the Kenya Life Panel Survey, with high survey tracking rates, and the labor market outcomes of the treatment and control groups are compared to assess impacts.

There were large increases in total hours worked for males as a result of deworming. There are sharp shifts in employment towards jobs that require full-time regular work, and have higher wages, notably towards manufacturing sector jobs (especially for males) and away from casual labor. As a result, among those working for wages average earnings rise by over 20%. These findings complement Bleakley's work on historical deworming programs in the U.S. South in the early 20th century, and the correspondence between the two sets of results – using distinct research designs and data – increases confidence in both findings.

The finding that shifts into different employment sectors account for the bulk of the earnings gains suggests that characteristics of the broader labor market – for instance, sufficient demand for manufacturing workers – may be critical for translating better health into higher living standards. Our finding of considerable labor market impacts (outside of the agricultural

[25] DeLong and Summers (2012) have recently made a different, Keynesian, argument that current government expenditures could pay for themselves during periods of inadequate aggregate demand.

sector) suggests that Kenyan labor markets are able to more flexibly allocate higher productivity workers to different tasks than is often believed. We cannot decompose how much of our labor market impacts are working through health versus education without imposing strong assumptions, but both "channels" appear to play a role.

The social returns to child deworming treatment appear high using an approach based on calibration the Grossman (1972) model, or an alternative social planner approach, where the latter generates an annualized social internal rate of return of 64.6%. In fact, using parameter estimates from our data and actual Kenyan public finance statistics, and under conservative assumptions, deworming generates far more in later government revenue (appropriately discounted) than it costs in upfront subsidies, making it a highly attractive public investment.

It goes without saying that deworming alone, and its associated increase in earnings, cannot make more than a small dent in the large gap in living standards between poor African countries like Kenya and the world's rich countries. Yet that obvious point does not make deworming any less attractive as a public policy option given its extraordinarily high rate of return, the possibility that deworming subsidies would be Pareto-improving, and the fact that boosting earnings by roughly a fifth would have major welfare impacts for households living near subsistence.

References

Abudu, F. (1986). "Work attitudes of Africans, with special reference to Nigeria", *International Studies of Management and Organization*, 16(2), 17-36.

Alderman, Harold, John Hoddinott, Bill Kinsey. (2006a). "Long-term consequences of early childhood malnutrition," *Oxford Economic Papers* 58, 450-474.

Alderman, H., J. Konde-Lule, I. Sebuliba, D. Bundy, A. Hall. (2006b). "Increased weight gain in preschool children due to mass albendazole treatment given during 'Child Health Days' in Uganda: A cluster randomized controlled trial", *British Medical Journal*, 333, 122-6.

Alderman, Harold. (2007). "Improving nutrition through community growth promotion: Longitudinal study of nutrition and early child development program in Uganda", *World Development*, 35(8), 1376-1389.
Alesina, Alberto, Paola Giuliano, and Nathan Nunn. (2012). "On the Origins of Gender Roles: Women and the Plough", unpublished working paper.
Almond, Douglas, and Janet Currie. (2010). "Human Capital Development Before Age Five", NBER Working Paper #15827.
Anderson, R.M., and R.M. May. (1991). *Infectious Diseases of Humans*. New York: Oxford University Press.
Ashraf, N., J. Berry, and J.M. Shapiro. (2010). "Can high prices stimulate product use? Evidence from a field experiment in Zambia", *American Economic Review*, 100(5), 2383-2413.
Atkinson, Anthony B. and Nicholas Stern. (1974). 'Pigou, Taxation and Public Goods', *Review of Economic Studies,* 41, 119–28.
Baird, Sarah, Joan Hamory, and Edward Miguel. (2008) "Attrition and Migration in the Kenya Life Panel Survey." University of California, Berkeley CIDER Working Paper.
Beegle, K., J. De Weerdt, S. Dercon. (2011). "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey", *Review of Economics and Statistics*, 93(3):1010-33.
Behrman, J.R., M.C. Calderon, S. Preston, J. Hoddinott, R. Martorell and A.D. Stein. (2009). "Nutritional Supplementation of Girls Influences the Growth of their Children: Prospective Study in Guatemala", *American Journal of Clinical Nutrition*, 90 (November), 1372-1379.
Bleakley, Hoyt. (2007). "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics*, 122(1):73-117.
Bleakley, Hoyt. (2010). "Health, Human Capital, and Development". *Annual Review of Economics,* 2, 283-310.
Brook, R.H., J.E. Ware, W.H. Rogers, E.B. Keeler, A.R. Davies, C.D. Sherbourne, G.A. Goldberg, K.N. Lohr, P. Camp, J.P. Newhouse. (1984). *The Effect of Coinsurance on the Health of Adults: Results from the RAND Health Insurance Experiment*. RAND: Santa Monica, CA.
Brooker, S., Miguel, E.A., Moulin, S., Luoba, A.I., Bundy, D.A.P. and Kremer, M. (2000a). "Epidemiology of Single and Multiple Species Helminth Infections among Schoolchildren in Busia District, Kenya", *East African Medical Journal*, 77, 157-161.
Brooker, Simon, Megan Rowlands, Laurence Haller, Lorenzo Savioli, and Donald Bundy. (2000b) "Towards an Atlas of Helminth Infection in Sub-Saharan Africa: The Use of Geographical Information Systems (GIS)", *Parasitology Today*, 16(7), 303-307.
Bruhn, M., D. McKenzie. (2009). "In pursuit of balance: Randomization in practice in development field experiments", *American Economic Journal: Applied Economics*, 1(4), 200-232.
Bundy, D.A.P. (1994). "The Global Burden of Intestinal Nematode Disease." *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 88, 259-261.
Bundy, D.A.P., and E.S. Cooper. (1989). "Trichuris", in *Tropical and Geographical Medicine*, eds. K.S. Warren and A.A.F. Mahmoud. New York: McGraw-Hill.
Bundy, D.A.P., M. Kremer, H. Bleakley, M.C.H. Jukes, E. Miguel. (2009). "Deworming and development: Asking the right questions, asking the questions right", *PLoS Neglected Tropical Diseases*, 3(1): e362. doi:10.1371/journal.pntd.0000362.

Case, Anne, and Christina Paxson. (2010). “Causes and Consequences of Early Life Health”, *Demography*, 47 (Supplement), S65-S85.
Clark, Greg (1994). “Factory Discipline” *Journal of Economic History* 54 (1):128-163
Cohen, Jessica and Pascaline Dupas. (2010). “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment,” *Quarterly Journal of Economics* 125 (1):1-45**.**
Costa, Dora (2000) “The Wage and the Length of the Work Day: From the 1890s to 1991.” *Journal of Labor Economics*. 18(1): 156-181.
Crimmins, Eileen M., and Caleb E. Finch. (2005). “Infection, inflammation, height, and longevity”, *Proceedings of the National Academy of Sciences*, 103(2): 498-503.
Currie, Janet, Eliana Garces, and Duncan Thomas. (2002). “Longer Term Effects of Head Start”, *American Economic Review*, 92(4), 999-1012.
Currie, Janet. (2009). “Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development?”, *Journal of Economic Literature*, 47(1), 87-122.
Delong, Bradford, and Lawrence Summers. (2012). “Fiscal policy in a depressed economy”, unpublished working paper, U.C. Berkeley.
de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2009). “Measuring microenterprise profits: Must we ask how the sausage is made?”, *Journal of Development Economics*, 88(1): 19-31.
de Silva, NR, Brooker S, Hotez PJ, Montresor A, Engels D, Savioli L. (2003). “Soil-transmitted helminth infections: updating the global picture” *Trends in Parasitology,* 19(12):547-51.
Deaton, Angus and Jean Drèze (2008). “Food and Nutrition in India: Facts and Interpretations,” *Economic and Political Weekly* 44 (7): 42-65
Diamond, Peter A., and James A. Mirlees. (1971). “Optimal Taxation and Public Production I: Production Efficiency and II: Tax Rules”, *American Economic Review*, 61: 8-27, 261-78.
Dupas, Pascaline. (2011). “Health Behavior in Developing Countries”, *Annual Review of Economics*.
Fafchamps, Marcel. (1993). “Sequential Labor Decisions Under Uncertainty: an Estimable Household Model of West African Farmers”, *Econometrica*, 61(5): 1173-1197.
Fernandez, R. and A. Fogli. (2009). “Culture: An empirical investigation of beliefs, work, and fertility”, *American Economic Journal: Macroeconomics*, 1(1):146-177.
Fox, M.P., S. Rosen, W.B. MacLeod, M. Wasunna, M. Bii, G. Foglia, and J. L. Simon. (2004). “The impact of HIV/AIDS on labour productivity in Kenya”, *Tropical Medicine and International Health*, 9(3), 318-324.
Glewwe, Paul, Hanan G. Jacoby, Elizabeth M. King (2001). “Early childhood nutrition and academic achievement: a longitudinal analysis.” *J. of Public Economics,* 81: 345-368.
Government of Kenya. (1986). Kenya Socio-Cultural Profiles: Busia District, ed. Gideon Were (Nairobi: Ministry of Planning and National Development).
Grossman, Michael. (1972). “On the Concept of Health Capital and the Demand for Health”, *Journal of Political Economy*, 80(2), 223-255.
Guyatt, H.L., Brooker, S., Kihamia, C.M., Hall, A. & Bundy, D.A. 2001, "Evaluation of efficacy of school-based anthelmintic treatments against anaemia in children in the United Republic of Tanzania", *Bulletin of the World Health Organization,* 79(8), 695-703.
Habyarimana, James, Bekezela Mbakile, and Cristian Pop-Eleches. (2010). “HIV/AIDS, ARV Treatment and Worker Absenteeism: Evidence from a Large African Firm”, *Journal of Human Resources*, 45(4).

Haddock, C.K., W.S.C. Poston, S.A. Pyle, R.C. Klesges, M.W. Vander Weg, A. Peterson, M. Debon. (2006). “The validity of self-rated health as a measure of health status among young military personnel: evidence from a cross-sectional survey”, *Health and Quality of Life Outcomes*, 4(57).

Hanna, Rema, and Paulina Oliva. (2011). “The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City”, *NBER Working Paper #17302*.

Heckman, James, Jora Stixrud, and Sergio Urzua. (2006). “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”, *Journal of Labor Economics*, 24(3): 411-482.

Hoddinott, John, John A. Maluccio, Jere R Behrman, Rafael Flores and Reynaldo Martorell. (2008). “The Impact of Nutrition During Early Childhood on Income, Hours Worked, and Wages of Guatemalan Adults”, *Lancet*, 371 (February), 411-416.

Hotez, Peter J. (2009). “Empowering Women and Improving Female Reproductive Health through Control of Neglected Tropical Diseases”, *PLoS Neglected Tropical Diseases*, 3(11): e559. doi:10.1371/journal.pntd.0000559.

Hotez, Peter J and Alan Fenwick. (2009) “Schistosomiasis in Africa: An Emerging Tragedy in Our New Global Health Decade”, PLOS Neglected Tropical Diseases, 3(9): e485.

Ichino, Andrea, and Enrico Moretti. (2009). “Biological Gender Differences, Absenteeism and the Earning Gap”, *American Economic Journal: Applied Economics*, 1(1).

Idler, Ellen L., and Yael Benyamini. (1997). “Self-rated health and mortality: A review of twenty-seven community studies”, *Journal of Health and Social Behavior*, 38(1), 21-37.

Jukes, Matthew C. H, Margaret Pinder, Elena L. Grigorenko, Helen Baños Smith, Gijs Walraven, Elisa Meier Bariau, Robert J. Sternberg, Lesley J. Drake et al. (2006). “Long-Term Impact of Malaria Chemoprophylaxis on Cognitive Abilities and Educational Attainment: Follow-Up of a Controlled Trial.” *PLoS Clinical Trials*, 1(4): e19.

Kaplow, Louis. (2008). *The Theory of Taxation and Public Economics,* Princeton: Princeton University Press.

Kaplow, Louis. (2009). “Taxing Leisure Complements”, *Economic Inquiry*, 48(4): 1065-1071.

Kaur, Supreet, Michael Kremer and Sendhil Mullainathan. (2012). “Self-Control at Work”, unpublished working paper.

Kirwan, P., Jackson, A.L., Asaolu, S.O., Molloy, S.F., Abiona, T.C., Bruce, M.C., Ranford-Cartwright, L., O'Neill, S.M. & Holland, C.V. 2010, "Impact of repeated four-monthly anthelmintic treatment on Plasmodium infection in preschool children: a double-blind placebo-controlled randomized trial", *BMC infectious diseases,* vol. 10, pp. 277.

Knight, J.B., R.H Sabot, and D.C. Hovey. (1992). “Is the rate of return on primary schooling really 26 per cent?”, *Journal of African Economies*, 1(2), 192-205.

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

Kremer, Michael, and Alaka Holla. (2009). “Pricing and Access: Lessons from Randomized Evaluations in Education and Health”, in *What Works in Development: Thinking Big and Thinking Small*, eds. W. Easterly and J. Cohen. Washington DC: Brookings Institution Press.

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

Laffer, A. (2004) *The Laffer Curve, Past, Present, and Future*. Heritage Foundation.

Lam, D., C. Ardington, N. Branson, A. Case, A. Menendez, M. Leibbrandt, J. Seekings, M. Sparks. (2008). *The Cape Area Panel Study: Overview and Technical Documentation Waves 1-2-3-4 (2002-2006)*. Working paper, Univ. of Michigan, Univ. of Cape Town, Princeton University.

Lewis, W. Arthur. (1954). "Economic Development with Unlimited Supplies of Labor", *The Manchester School*, 22(2), 139-191.

Maccini, Sharon, and Dean Yang. (2009). "Under the Weather: Health, Schooling and Economic Consequences of Early-life Rainfall", *American Economic Review*, 99(3): 1006-1026.

Maluccio, J.A., J. Hoddinott, J.R. Behrman, R. Martorell, A. Quisumbing, A.D Stein. (2009). "The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults." *Economic Journal*, 199 (537) 734-763.

Miguel, Edward and Michael Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(1), 159-217.

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

OECD Average annual hours actually worked per worker, 2011, found at http://stats.oecd.org/Index.aspx?DataSetCode=ANHRS

Orr, Larry, Judith D. Ferris, Robin Jacob, Erik Beecroft, Abt. Associates Inc., Lisa Sanabonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, Jeffrey R. Kling. (2003). *Moving to Opportunity: Interim Impacts Evaluation.* Washington D.C.: U.S. Dept. of HUD.

Ozier, Owen. (2010). "Exploiting Externalities to Estimate the Long-Term Benefits of Early Childhood Deworming", unpublished working paper, U.C. Berkeley.

Pitt, Mark M., Mark R. Rosenzweig, and Nazmul Hassan. (2011). "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy", unpublished working paper.

Prescott, Edward (2004), "Why do Americans work so much more than Europeans?", *Federal Reserve Bank of Minneapolis Quarterly Review*, 28 (1), 2-13.

Pullan RL, Gething PW, Smith JL, Mwandawiro CS, Sturrock HJ, Gitonga CW, Hay SI, Brooker S. (2011). "Spatial modelling of soil-transmitted helminth infections in kenya: a disease control planning tool", *PLoS Neglected Tropical Diseases*, Feb 8; 5(2):e958.

Schapiro, Louis. (1919). "The Physical and Economic Benefits of Treatment for Hookworm Disease", *Journal of the American Medical Association*, Nov. 15, 1919, 73(20): 1507-09.

Schultz, T. Paul. (2005). "Productive Benefits of Health: Evidence from Low-Income Countries", Yale Economic Growth Center Discussion Paper #903.

Schultz, T. Paul and Aysit Tansel. (1997). "Wage and labor supply effects of illness in Cote d'Ivoire and Ghana: Instrumental variable estimates for days disabled", *Journal of Development Economics*, 53(2): 251-286.

Smith, James P. (2009). "The Impact of Childhood Health on Adult Labor Market Outcomes", *Review of Economics and Statistics*, 91(3), 478-489.

Stephenson, L.S., Latham, M.C., Adams, E.J., Kinoti, S.N. & Pertet, A. 1993, "Physical fitness, growth and appetite of Kenyan school boys with hookworm, Trichuris trichiura and Ascaris lumbricoides infections are improved four months after a single dose of albendazole", *The Journal of nutrition,* vol. 123, no. 6, pp. 1036-1046.

Stoltzfus, R.J., H.M. Chwaya, J.M. Tielsch, K.J. Schulze, M. Albonico, and L. Savioli. (1997). "Epidemiology of iron deficiency anemia in Zanzibari schoolchildren: the importance of hookworms", *American Journal of Clinical Nutrition*, 65(1): 153-159.

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

Taylor-Robinson D, Jones A, Garner P. (2007) "Deworming drugs for treating soil-transmitted intestinal worms in children: effects on growth and school performance." *Cochrane Database of Systematic Reviews* 2007, 4, CD000371.doi:10.1002/14651858.CD000371.pub31.

Thomas, Duncan, Elizabeth Frankenberg, James P. Smith. (2001). "Lost but not forgotten: Attrition in the Indonesia Family Life Survey," *J. of Human Resources*, 36(3), 556-592.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, Wayan Suriastini. (2012). "Cutting the costs of attrition: Results from the Indonesia Family Life Survey", *Journal of Development Economics*, 9(1): 108-123.

Thomas, Duncan, et al. (2008). "Causal effect of health on social and economic prosperity: Experimental evidence", unpublished working paper, Duke University.

Weisbrod, Burton A., Ralph L. Andreano, Robert E. Baldwin, Erwin H. Epstein, Allen C. Kelley, and Thomas W. Helminiak. (1973). *Disease and economic development: The impact of parasitic diseases in St. Lucia*. University of Wisconsin Press: Madison.

World Bank. (2006). *Repositioning nutrition as central to development: a strategy for large-scale action*. World Bank, Washington DC.

Table 1: Baseline (1998) summary statistics and PSDP randomization checks, and KLPS (2007-09) survey attrition patterns

	Treatment – Control (s.e.)			Control group mean (s.d.)		
Panel A: Baseline summary statistics	All	Male	Female	All	Male	Female
Age (1998)	-0.04	-0.16	0.08	12.0	12.2	11.7
	(0.11)	(0.17)	(0.12)	(2.6)	(2.7)	(2.5)
Grade (1998)	-0.03	-0.07	0.02	4.25	4.26	4.24
	(0.05)	(0.07)	(0.08)	(1.66)	(1.67)	(1.65)
Female	-0.004			0.473		
	(0.019)					
School average test score (1996)	-0.013	-0.038	0.014	0.038	0.042	0.032
	(0.109)	(0.108)	(0.114)	(0.406)	(0.403)	(0.408)
Primary school located in Budalangi division	-0.017	-0.030	-0.002	0.381	0.387	0.374
	(0.137)	(0.141)	(0.136)			
Population of primary school	58	49	68	436	445	426
	(54)	(51)	(57)	(146)	(145)	(146)
Total treatment (Group 1,2) primary school students within 6 km	-296	-290	-302	3,381	3,375	3,388
	(260)	(271)	(255)	(1,022)	(1,022)	(1,024)
Total primary school students within 6 km	-34	1	-74	4,732	4,717	4,749
	(389)	(399)	(386)	(1,555)	(1,553)	(1,558)
Years of assigned deworming treatment, 1998-2003	2.41***	2.45***	2.37***	1.68	1.68	1.67
	(0.08)	(0.10)	(0.09)	(1.23)	(1.24)	(1.23)
Panel B: Sample attrition, KLPS (2007-09)						
Found [a]	-0.007	0.007	-0.021	0.867	0.878	0.854
	(0.017)	(0.022)	(0.025)			
Surveyed	-0.003	0.016	-0.023	0.827	0.834	0.820
	(0.018)	(0.023)	(0.025)			
Not surveyed, dead	0.004	0.003	0.006	0.014	0.016	0.012
	(0.004)	(0.005)	(0.005)			
Not surveyed, refused	-0.003	-0.005	0.000	0.017	0.017	0.016
	(0.005)	(0.007)	(0.007)			

Notes: Panel A data is from the PSDP, and includes individuals surveyed in KLPS2. N=5,084 observations, with 2,595 males and 2,489 females (except for age, N=5,072). Years of assigned deworming treatment is calculated using the treatment group of the respondent's school and grade, but is not adjusted for the treatment ineligibility of females over age 13 or 2001 cost-sharing. Respondents who "age out" of primary school are no longer considered assigned to treatment. School average test scores are from the 1996 Busia mock exam, and are converted to normalized individual standard deviation units. Panel B includes all individuals surveyed, refused participation, deceased, found but unable to survey, and not found but sought in intensive tracking, for 5,569 respondents (3,686 treatment and 1,883 control; 2,827 males and 2,742 females). Observations are weighted to maintain initial population proportions. The "Treatment – Control" differences are derived from a linear regression on a constant and the treatment indicator. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence. [a] "Found" includes pupils surveyed, refused, deceased, and found but unable to survey.

Table 2: Impacts on health, fertility, and education outcomes

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Health and fertility, KLPS (2007-09)	All	Male	Female	All	All	Male	Female
Self-reported health "very good"	0.041**	0.024	0.053**	0.028	0.673	0.713	0.629
	(0.018)	(0.025)	(0.025)	(0.022)	(0.469)	(0.452)	(0.483)
					5,070	*2,585*	*2,485*
Currently married indicator	-0.019	0.012	-0.046	0.012	0.347	0.255	0.449
	(0.023)	(0.025)	(0.030)	(0.023)	*5,082*	*2,594*	*2,488*
Number of pregnancies	-0.093	-0.044	-0.138	-0.044	0.98	0.71	1.28
(for females – themselves; for males – their partners)	(0.066)	(0.062)	(0.095)	(0.065)	(1.29)	(1.20)	(1.31)
					5,072	*2,589*	*2,483*
Miscarriage indicator (observation at pregnancy level)	-0.015*	-0.001	-0.028**	-0.015**	0.030	0.015	0.039
(for females – themselves; for males – their partners)	(0.008)	(0.005)	(0.013)	(0.007)	(0.171)	(0.123)	(0.194)
					5,022	*1,622*	*3,238*
Panel B: Education							
Total years enrolled in school, 1998-2007	0.279*	0.112	0.366**	0.138	6.69	7.05	6.29
	(0.147)	(0.169)	(0.172)	(0.149)	(2.97)	(2.93)	(2.96)
					5,037	*2,567*	*2,470*
Indicator for repetition of at least one grade (1998-2007)	0.060***	0.064**	0.056*	0.010	0.672	0.669	0.676
	(0.017)	(0.025)	(0.029)	(0.023)	(0.470)	(0.471)	(0.468)
					5,084	*2,595*	*2,489*
Grades of schooling attained	0.153	-0.036	0.276	0.070	8.72	9.06	8.34
	(0.143)	(0.147)	(0.166)	(0.146)	(2.21)	(2.28)	(2.07)
					5,084	*2,595*	*2,489*
English vocabulary test score (normalized), 2007-09	0.076	0.061	0.081	0.067	0.000	0.115	-0.129
	(0.055)	(0.059)	(0.073)	(0.053)	(1.000)	(1.021)	(0.957)
					5,084	*2,595*	*2,489*
Passed primary school leaving exam during 1998-2007	0.048	0.001	0.095**	0.032	0.505	0.590	0.409
	(0.031)	(0.029)	(0.041)	(0.029)	(0.500)	(0.492)	(0.492)
					4,974	*2,541*	*2,433*
Indicator for out-of-school (at 2007-09 survey)	-0.003	0.029	-0.029	0.045*	0.75	0.70	0.80
	(0.022)	(0.030)	(0.025)	(0.026)	(0.43)	(0.46)	(0.40)
					5,058	*2,582*	*2,476*
Out-of-school sample							

English vocabulary test score (normalized), 2007-09	0.107**	0.149*	0.059	0.149***	-0.232	-0.151	-0.310
	(0.052)	(0.076)	(0.068)	(0.047)	(0.972)	(1.016)	(0.922)
					3,873	*1,869*	*2,004*
Passed primary school leaving exam during 1998-2007	0.061*	0.037	0.079**	0.083***	0.413	0.477	0.350
	(0.032)	(0.035)	(0.042)	(0.028)	(0.493)	(0.500)	(0.477)
					3,775	*1,822*	*1,953*

Table 2 notes: The sample includes all individuals surveyed in KLPS-2, and the rows underneath "*Out-of-school sample*" further condition on not being enrolled in school at the time of survey. Self-reported health "very good" takes on a value of one if the answer to the question "Would you describe your general health as somewhat good, very good, or not good?" is "very good", and zero otherwise. Each entry is from a separate OLS regression except the miscarriage outcome, which are marginal probit specifications in which each observation is a pregnancy.

Regression notes: All observations are weighted to maintain initial population proportions. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence. The externality term is the total treatment group (Group 1, Group 2) pupils within 6 km (in '000s), demeaned. All regressions 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.

Table 3: Deworming impacts on hours worked and meals eaten

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Hours worked across all sectors in last week	All	Male	Female	All	All	Male	Female
Full sample	1.53	3.40**	0.29	1.71	18.4	20.3	16.3
	(1.03)	(1.39)	(1.34)	(1.43)	(23.1)	(24.6)	(21.1)
					5,084	*2,595*	*2,489*
Out-of-school sample	2.74**	4.14**	2.13	2.04	22.0	25.9	18.3
	(1.30)	(1.95)	(1.47)	(1.74)	(24.8)	(26.5)	(22.4)
					3,873	*1,869*	*2,004*
Panel B: Number of meals eaten yesterday							
Full Sample	0.096***	0.127***	0.051	0.080***	2.16	2.10	2.23
	(0.028)	(0.041)	(0.043)	(0.023)	(0.64)	(0.65)	(0.62)
					5,083	*2,595*	*2,488*
Out-of-school sample	0.103***	0.157***	0.041	0.101***	2.16	2.08	2.25
	(0.029)	(0.046)	(0.044)	(0.032)	(0.64)	(0.66)	(0.62)
					3,872	*1,869*	*2,003*

Table 3 notes: Hours worked across "all sectors" includes work in agriculture, wage employment and self-employment. Each entry is from a separate OLS regression. For details on the regressions, see the "Regression notes" for Table 2.

Table 4: Fiscal Impacts of Deworming Subsidies

Panel A: Calibration parameters	No Subsidy	Partial Subsidy	Full Subsidy	*Notes*
Size of subsidy: s	$0.00	$1.15	$1.42	From Deworm the World; Kremer and Miguel (2007)
Take-up rate: $\theta(s)$	5%	19%	75%	From Kremer and Miguel (2007)
Average per-person cost: $s\theta(s)$	$0.00	$0.22	$1.07	= Subsidy x Take-up rate
Mean per person increase in work hours/week: x	0.09	0.32	1.28	Men: increase 3.4 hours/week; women: no change (Table 2). Multiply hours increase by take-up.
Mean increase in work hours/week from externality: x_{EXT}	0.40	1.54	6.07	Increase equivalent to increase in deworming take-up (within 6 km) of respondent's school from 0% to $\theta(s)$
Panel B: No health spillovers				
NPV increase in per-person earnings (relative to no subsidy)	-	$11.05	$55.26	9.85% Annual (real) discount rate in Kenya
NPV increase in per-person government revenue	-	$1.83	$9.16	NPV earnings x 16.5% tax rate under no subsidy
Panel C: With health spillovers				
NPV increase in per-person earnings (relative to no subsidy)	-	$63.66	$318.31	9.85% Annual (real) discount rate in Kenya
NPV increase in per-person government revenue	-	$10.55	$52.76	NPV earnings x 16.5% tax rate under no subsidy
Panel D: Internal rate of return				
Government internal rate of return, no health spillovers	-	24.5%	24.7%	Subsidies from Panel A, revenue from Panel B
Government internal rate of return, with health spillovers	-	41.7%	42.0%	Subsidies from Panel A, revenue from Panel C
Social internal rate of return, no health spillovers	-	39.9%	42.5%	Costs from Panel A, earnings from Panel B
Social internal rate of return, with health spillovers	-	61.4%	64.6%	Costs from Panel A, earnings from Panel C

Table 4 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 mean starting hourly wage rate (w) of $0.16 from Suri (2011), which is conservative (somewhat lower than the average in KLPS2). Information about Kenyan public finance comes from the Kenyan Central Bank website and the World Bank Development Indicators. The NPV of per-person lifetime earnings in the no subsidy case and no health spillovers is $858.44, and with spillovers is $877.23. We assume that earnings start 10 years after treatment and continue for 40 years. Life cycle earnings profiles for Kenya are from Knight, Sabot and Hovey (1992), Table 2. Full calculations are available by request.

Table 5: Hours Worked Decomposition

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Total Hours in All Sectors	All	Male	Female	All	All	Male	Female
Hours Worked	1.53	3.40**	0.29	1.71	18.4	20.3	16.3
	(1.03)	(1.39)	(1.34)	(1.43)	(23.1)	(24.6)	(21.1)
Indicator for hours > 0	0.004	0.050*	-0.036	0.002	0.68	0.68	0.68
	(0.022)	(0.026)	(0.031)	(0.022)	(0.47)	(0.47)	(0.47)
Hours worked, among those with hours > 0	2.12*	3.38**	1.50	2.52	27.0	29.8	24.0
	(1.20)	(1.56)	(1.51)	(1.76)	(23.4)	(24.5)	(21.8)
					3,579	*1,898*	*1,681*
Panel B: Agriculture							
Hours Worked	-0.08	0.99*	-1.29**	-0.16	8.3	7.8	8.8
	(0.43)	(0.55)	(0.57)	(0.62)	(11.4)	(11.6)	(11.2)
Indicator for hours > 0	-0.010	0.019	-0.040	0.024	0.55	0.53	0.58
	(0.025)	(0.028)	(0.032)	(0.027)	(0.50)	(0.50)	(0.49)
Hours worked, among those with hours > 0	-0.058	1.23	-1.34	-1.05	14.9	14.8	15.1
	(0.621)	(0.85)	(0.84)	(0.97)	(11.7)	(12.3)	(11.1)
					2,916	*1,454*	*1,462*
Panel C: Wage Employment							
Hours Worked	0.09	0.99	-0.24	0.72	6.9	8.8	4.8
	(0.82)	(1.29)	(1.06)	(1.00)	(18.5)	(20.0)	(16.5)
Indicator for hours > 0	-0.014	-0.006	-0.009	0.003	0.15	0.20	0.09
	(0.016)	(0.027)	(0.018)	(0.017)	(0.36)	(0.40)	(0.29)
Hours worked, among those with hours > 0	4.65*	5.59	-1.20	4.93	46.5	43.7	53.6
	(2.80)	(3.38)	(3.82)	(3.20)	(21.7)	(21.7)	(20.2)
					625	*470*	*155*
Panel D: Self-Employment (non-agricultural)							
Hours Worked	1.52***	1.41**	1.82**	1.15*	3.3	3.8	2.7
	(0.55)	(0.70)	(0.80)	(0.59)	(12.8)	(13.7)	(11.7)
Indicator for hours > 0	0.023*	0.022	0.028	0.005	0.09	0.09	0.08
	(0.011)	(0.015)	(0.017)	(0.014)	(0.28)	(0.29)	(0.27)
Hours worked, among those with hours > 0	6.05**	6.19	5.86**	7.41**	38.1	40.2	35.6
	(2.99)	(4.67)	(2.92)	(2.95)	(24.0)	(23.1)	(25.1)
					542	*288*	*254*

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

Table 6: Deworming impacts on other productive activities

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean (s.d.); *Number of Observations*		
Panel A: Occupational and geographic mobility, full sample	All	Male	Female	All	All	Male	Female
Manufacturing job	0.010***	0.018**	0.005	0.007	0.005	0.007	0.003
	(0.004)	(0.007)	(0.004)	(0.005)	*5,084*	*2,595*	*2,489*
Construction/casual labor job	-0.005**	-0.003	-0.006	-0.002	0.005	0.004	0.006
	(0.002)	(0.003)	(0.004)	(0.002)	*5,084*	*2,595*	*2,489*
Lives in an urban area	-0.001	0.000	-0.003	0.016	0.179	0.161	0.198
	(0.019)	(0.030)	(0.022)	(0.024)	*5,075*	*2,587*	*2,488*
Panel B: Wage earners							
Ln(Total labor earnings), past month	0.253***	0.217*	0.156	0.199	7.86	7.98	7.55
	(0.093)	(0.117)	(0.187)	(0.168)	(0.88)	(0.87)	(0.84)
					710	*542*	*168*
Ln(Wage = Total labor earnings / hours), past month	0.165	0.122	0.274	0.012	2.82	3.02	2.31
	(0.117)	(0.155)	(0.212)	(0.160)	(0.96)	(0.94)	(0.83)
					625	*470*	*155*
Ln(Total labor earnings), most recent month worked since 2007	0.211***	0.196*	0.192*	0.170	7.88	8.02	7.59
	(0.072)	(0.101)	(0.108)	(0.116)	(0.91)	(0.89)	(0.90)
					1,175	*819*	*356*
Panel C: Self-employed							
Total self-employed profits (self-reported) past month	343	103	185	-151	1,766	2,135	1,265
	(306)	(476)	(266)	(320)	(2,619)	(3,235)	(1,261)
					585	*313*	*272*
Total self-employed profits past month, top 5% trimmed	324*	274	41	24	1,221	1,184	1,265
	(177)	(309)	(221)	(230)	(1,151)	(1,056)	(1,261)
					553	*284*	*269*
Total employees hired (excluding self)	0.466	0.244	0.627	0.044	0.188	0.253	0.097
	(0.361)	(0.406)	(1.294)	(0.492)	(0.624)	(0.614)	(0.630)
					633	*343*	*290*
Panel D: Wage earners or self-employed							
Total labor earnings + self-employed profits, past month (=0 for non-earners)	99	116	85	9	789	1,169	367
	(104)	(186)	(76)	(141)	(2,236)	(2,829)	(1,151)
					5.037	*2,567*	*2,470*
Panel E: Agriculture							
Grows cash crop	0.020**	0.008	0.031**	0.005	0.010	0.011	0.009
	(0.008)	(0.009)	(0.014)	(0.007)	*3,757*	*1,895*	*1,862*

Uses "improved" practices (fertilizer, seed, irrigation)	0.032	0.048	0.018	0.005	0.310	0.301	0.319
	(0.026)	(0.033)	(0.034)	(0.024)	*3,766*	*1,897*	*1,869*
Total value (KSh) of crop sales past year	-81	-153	47	-460**	576	606	542
	(148)	(231)	(185)	(206)	(2,458)	(2,566)	(2,336)
					3,758	*1,894*	*1,864*

Notes: Panel A includes all individuals surveyed in KLPS2. Panel B includes those who report positive labor earnings. Ln(Wage) is missing for those with zero reported earnings. Panel C restricts to those with positive self-employed profits; this is not restrictive as no respondent reports negative profits and only 5% report zero profits. "Agricultural work" in Panel E includes both farming and pastoral activities. Each entry is from a separate OLS regression, except for "total employees hired" which utilizes a negative binomial regression. For details on the regressions, see the "Regression notes" for Table 2.

Figure 1: Distribution of hours worked in last week among males, treatment versus control (if working 10 to 80 hours in sector)
Panel A (top-left): Across all sectors; Panel B (top-right): Agricultural sector only;
Panel C (bottom-left): Wage earners only; Panel D (bottom-right): Self-employed only.

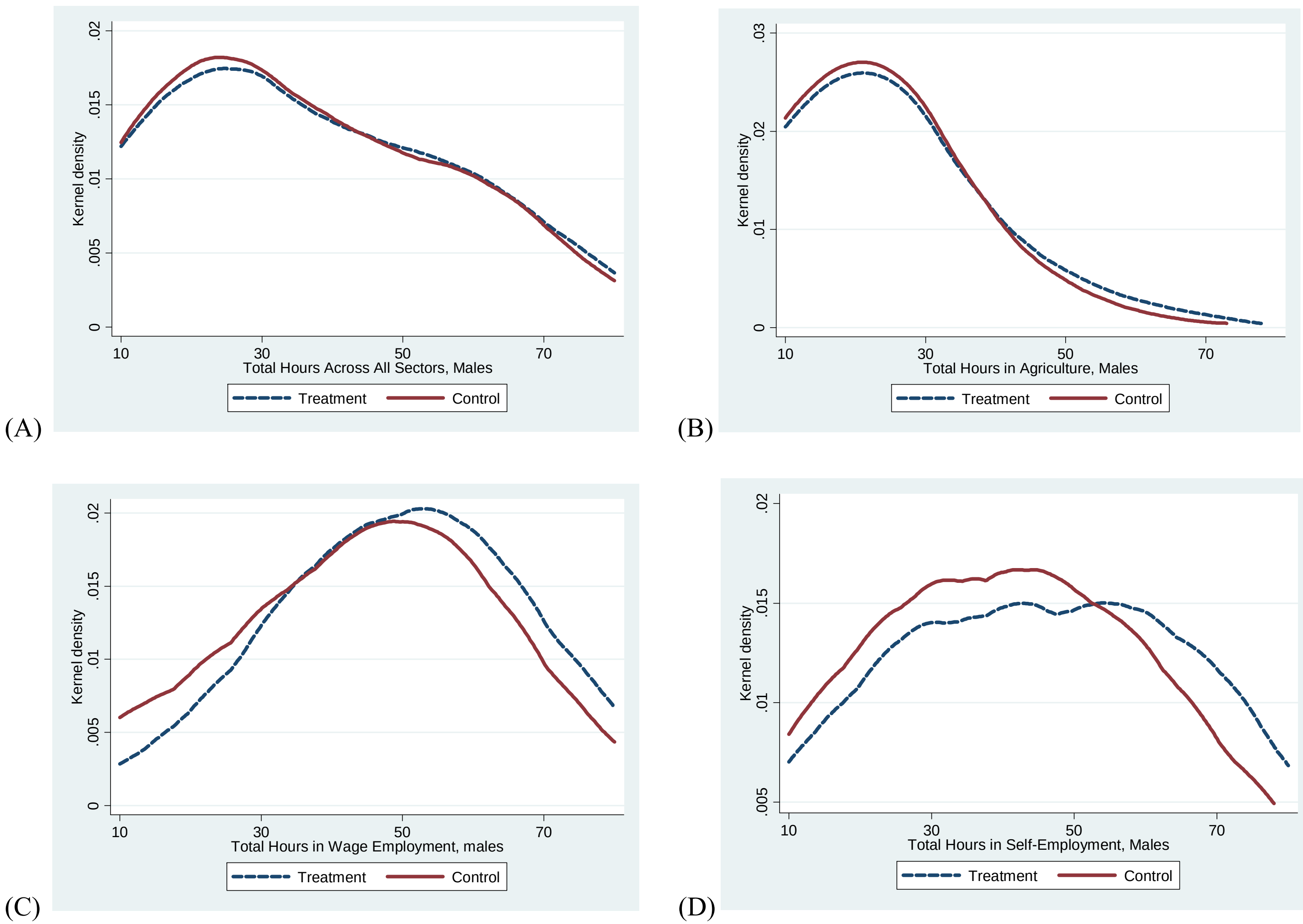
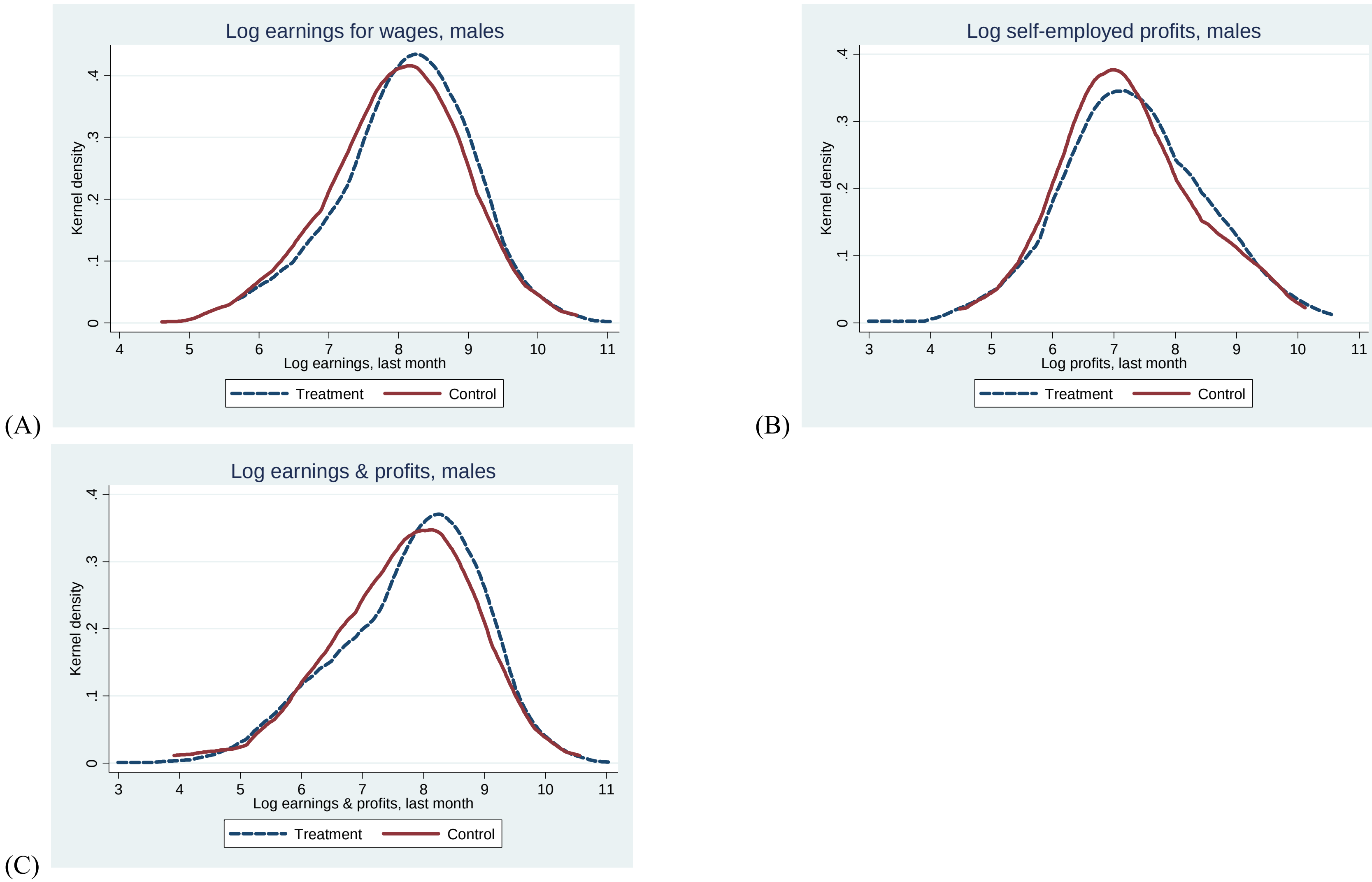


Figure 2: Distribution of log labor earnings and non-agricultural self-employment profits in the last month among males, treatment versus control (among those with positive labor earnings or profits)
Panel A (top-left): Log labor earnings; Panel B (top-right): Log self-employed profits (non-agricultural);
Panel C (bottom-left):Log labor earnings plus self-employed profits.



Supplementary Appendix A: Research Design Appendix (not intended for publication)

A.1 Selection of Primary Schools for the PSDP Sample:
There were a total of 92 primary schools in the study area of Budalangi and Funyula divisions, across eight geographic zones, in January 1998. Seventy-five of these 92 schools were selected to participate in PSDP. The 17 excluded schools include: town schools that were quite different from other local schools in terms of student socioeconomic background; single-sex schools; a few schools located on islands in Lake Victoria (posing severe transportation difficulties); and those few schools that had in the past already received deworming and other health treatments under an earlier small-scale ICS (NGO) program.

In particular, four primary schools in Funyula Town were excluded due to large perceived income differences between their student populations and those in other local schools. In particular, Moody Awori Primary School, Namboboto Boys Primary School, and Namboboto Girls School charged schools fees well in excess of neighboring primary schools, and thus attracted the local "elite". Nangina Girls Primary School is a private boarding school, and charged even higher fees, and was similarly excluded.

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

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

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

A.2 Prospective Experimental Procedure:
Miguel and Kremer (2004) contains a partial description of the prospective experimental "list randomization" procedure, and we expand on it here. Schools were first stratified by geographical area (division, then zone)[26], and the zones were listed alphabetically (within each division), and then within each zone they were listed in increasing order of student enrolment in the school. Table 1 shows there is no significant difference between average school populations in the treatment and control groups.

While the original plan had been to stratify by participation in other NGO programs, the actual randomization was not carried out this way. Schools participating in the intensive CSP/SHP program were dropped from the sample (as detailed above), while 27 primary schools with less intensive NGO programs were retained in the sample. These 27 schools were receiving assistance in the form of either free classroom textbooks, grants for school committees, or teacher training and bonuses. It is worth emphasizing that the randomized evaluations of these various interventions did

[26] There are two divisions (Budalangi and Funyula) containing a total of eight zones (Agenga/Nanguba, Bunyala Central, Bunyala North, Bunyala South, Bwiri, Funyula, Namboboto, Nambuku).

not find statistically significant average project impacts on a wide range of educational outcomes.[27] The schools that benefited from these previous programs were found in all eight geographic zones; the distribution of the 27 schools across the eight zones is: Agenga/Nanguba (5 schools), Bunyala Central (1), Bunyala North (4), Bunyala South (2), Bwiri (4), Funyula (5), Namboboto (1), Nambuku (5). The results in the current paper are robust to including controls for inclusion in these other NGO programs (results not shown).

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

Deaton (2010) raises concerns about the list randomization approach, in the case where the first school listed in the first randomization "triplet" is different than other schools (in our case, it has lower than average school enrolment); the same concerns would apply to several other well-known recent field experiments in development economics, most notably Chattopadhyay and Duflo's 2004 paper "Women as policymakers: Evidence from a randomized policy experiment in India" in *Econometrica*.[28] However, this is not a major threat to our empirical approach. Following Bruhn and McKenzie (2009) we include all variables used in the randomization procedure (such as baseline school enrolment) as explanatory variables in our regression specifications, thus controlling for any direct effect of school size, and partially controlling for unmeasured characteristics correlated with school size. Coefficient estimates on the deworming treatment indicator are largely unchanged whether or not additional explanatory variables are included, suggesting that any bias is likely to be very small. The difference in average school enrollment between the treatment and control groups is small and not statistically significant (Table 1). Moreover, even if the first school in the first randomization triplet were an outlier along some unobserved dimension (which seems unlikely), given our sample size of 75 schools and 25 randomization triplets, and the fact that school size is not systematically related to treatment group assignment for the other 24 randomization triplets (as discussed above), approximately 96% of any hypothesized bias would be eliminated. Taken together, the prospective experimental design we exploit in the current paper is likely to yield reliable causal inference.

[27] See Glewwe, Paul, Michael Kremer, and Sylvie Moulin. (2009). "Many Children Left Behind? Textbooks and Test Scores in Kenya", *American Economic Journal: Applied Economics*, 1(1): 112-135.

[28] The references are Deaton, Angus. (2010). "Instruments, Randomization and Learning about Development", *Journal of Economic Literature*, 48, 424-455, and Chattopadhyay, Raghabendra, and Esther Duflo. (2004). "Women as policymakers: Evidence from a randomized policy experiment in India", *Econometrica*, 75(2), 1409-1443.

Supplementary Appendix Table A1: Baseline (1998) summary statistics across treatment groups, out-of-school sample

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

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

Supplementary Appendix Table A2: Selection into out-of-school and employment types

	Out of school		In Agriculture		In Wage Employment		In Self-Employment	
Treatment	-0.006	-0.009	-0.010	-0.010	-0.016	0.027	0.014	-0.135***
	(0.022)	(0.076)	(0.026)	(0.083)	(0.018)	(0.056)	(0.012)	(0.042)
Female	0.089***	0.108***	0.043**	0.059*	-0.151***	-0.130***	-0.004	-0.025**
	(0.014)	(0.023)	(0.017)	(0.035)	(0.013)	(0.021)	(0.011)	(0.011)
Grade	0.087***	0.099***	-0.023***	-0.022**	0.036***	0.040***	0.025***	0.023***
	(0.005)	(0.008)	(0.006)	(0.011)	(0.004)	(0.006)	(0.003)	(0.004)
School average test score (1996)	-0.090***	-0.079***	-0.059*	0.025	-0.022	-0.087	-0.007	0.023
	(0.031)	(0.024)	(0.032)	(0.041)	(0.019)	(0.064)	(0.013)	(0.017)
Population of primary school	-0.005	-0.194**	0.032	-0.032	-0.015	-0.054	-0.025	-0.292***
	(0.040)	(0.083)	(0.046)	(0.080)	(0.033)	(0.096)	(0.028)	(0.054)
Cost sharing school (2001)	0.039*	0.036*	-0.024	-0.025	0.006	0.006	-0.001	-0.007
	(0.021)	(0.021)	(0.021)	(0.019)	(0.013)	(0.011)	(0.009)	(0.008)
Primary school located in Budalangi division	-0.009	0.001	-0.049	0.067	0.031	-0.014	0.031	0.066*
	(0.033)	(0.027)	(0.045)	(0.046)	(0.030)	(0.065)	(0.023)	(0.034)
Total treatment (Group 1,2) primary school students	0.036	0.015	0.030	0.100**	-0.007	0.022	0.003	0.050***
within 6 km	(0.028)	(0.050)	(0.028)	(0.041)	(0.020)	(0.055)	(0.011)	(0.018)
Total primary school students within 6 km	-0.048**	-0.036	-0.011	-0.063**	0.010	-0.007	-0.006	-0.031**
	(0.020)	(0.033)	(0.021)	(0.032)	(0.014)	(0.037)	(0.008)	(0.013)
Female * Treatment		-0.028		-0.020		-0.033		0.030
		(0.029)		(0.039)		(0.026)		(0.019)
Grade * Treatment		-0.018*		-0.001		-0.005		0.002
		(0.010)		(0.013)		(0.007)		(0.006)
School average test score * Treatment		-0.016		-0.106**		0.073		-0.037*
		(0.043)		(0.048)		(0.066)		(0.021)
Population of primary school * Treatment		0.222**		0.099		0.033		0.336***
		(0.095)		(0.094)		(0.100)		(0.057)
Budalangi division * Treatment		0.011		-0.120*		-0.020		-0.066**
		(0.052)		(0.061)		(0.058)		(0.028)
Total treatment school students within 6 km * Treatment		0.027		-0.088*		-0.030		-0.054**
		(0.059)		(0.051)		(0.059)		(0.022)
Total primary school students within 6 km * Treatment		-0.015		0.064		0.016		0.024
		(0.039)		(0.039)		(0.039)		(0.015)
R^2	0.137	0.141	0.025	0.033	0.074	0.081	0.025	0.032
Observations	5,058	5,058	5,043	5,043	5,081	5,081	5,083	5,083
Mean in the control group	0.748	0.748	0.555	0.555	0.166	0.166	0.100	0.100

Notes: The explanatory variables are from the PSDP, the outcome variables from KLPS2, and the analysis includes all individuals surveyed in KLPS2. The outcomes are indicator variables, and the employment variables take on a value of one if the respondent worked positive hours in the activity. F-tests of the joint significance of the treatment indicator and all treatment interaction terms give p-values of 0.155 for out-of-school, <0.001 for in agriculture, 0.087 for in wage employment, and <0.001 for in self-employment. Significant at 90% (*), 95% (**), 99% (***) confidence. All observations are weighted to maintain initial population proportions. All variables are 1998 values unless otherwise noted. The average school test score is from the 1996 Busia District mock exam, and has been converted to units of normalized individual standard deviations. Total treatment primary school students within 6 km and total primary school students within 6 km have been demeaned and are in 000's. Zone of 1998 primary school fixed effects are also included (as in the specifications in the main tables). Standard errors are clustered by school.

Supplementary Appendix Table A3: Baseline (1998) summary statistics across treatment groups, wage-earner sample

	Treatment – Control (s.e.)			Control group mean (s.d.)		
	All	Male	Female	All	Male	Female
Age (1998)	-0.28	-0.13	-1.09**	13.4	13.6	12.9
	(0.27)	(0.32)	(0.42)	(2.5)	(2.7)	(1.9)
Grade (1998)	-0.05	-0.03	-0.16	4.91	4.93	4.85
	(0.14)	(0.17)	(0.31)	(1.57)	(1.59)	(1.52)
Female	-0.071			0.280		
	(0.045)					
School average test score (1996)	-0.050	-0.020	-0.122	0.023	-0.010	0.111
	(0.106)	(0.103)	(0.138)	(0.391)	(0.357)	(0.460)
Primary school located in Budalangi division	0.052	0.026	0.115	0.378	0.405	0.310
	(0.144)	(0.149)	(0.156)			
Population of primary school	78	72	94	425	432	407
	(56)	(54)	(69)	(136)	(141)	(120)
Total treatment (Group 1,2) primary school students within 6 km	-268	-324	-63	3,382	3,390	3,363
	(282)	(254)	(420)	(1,064)	(987)	(1,250)
Total primary school students within 6 km	0	-75	250	4,730	4,759	4,655
	(420)	(382)	(633)	(1,598)	(1,495)	(1,846)
Years of assigned deworming treatment, 1998-2003	2.32***	2.28***	2.46***	1.23	1.23	1.24
	(0.14)	(0.17)	(0.24)	(1.23)	(1.25)	(1.16)

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

Supplementary Appendix Table A4: Baseline (1998) summary statistics and attrition checks

	Full KLPS Sample	Found: Regular Tracking	Found: Intensive Tracking	Not Found	Found (Regular and Intensive) – Not Found	Found (regular) – Found (intensive)
Age (1998)	12.4	12.4	12.5	12.7	-0.37***	0.18
	(2.2)	(2.2)	(2.2)	(2.1)	(0.09)	(0.14)
Grade (1998)	4.26	4.24	4.24	4.32	-0.105	0.003
	(1.69)	(1.68)	(1.70)	(1.70)	(0.063)	(0.097)
Female	0.486	0.461	0.495	0.535	-0.072***	0.033
	(0.500)	(0.499)	(0.501)	(0.499)	(0.016)	(0.031)
School average test score (1996)	0.043	0.035	0.023	0.066	-0.026	-0.012
	(0.439)	(0.434)	(0.416)	(0.453)	(0.021)	(0.025)
Primary school located in Budalangi division	0.380	0.361	0.389	0.420	-0.067***	0.029
	(0.486)	(0.480)	(0.488)	(0.494)	(0.023)	(0.037)
Population of primary school	484	480	465	496	-20**	-16
	(221)	(223)	(178)	(222)	(8)	(23)
Total treatment (Group 1,2) primary school students within 6 km	3171	3182	3174	3149	30	-7
	(910)	(915)	(918)	(900)	(36)	(55)
Total primary school students within 6 km	4678	4713	4691	4602	93	-21
	(1340)	(1342)	(1335)	(1334)	(62)	(79)
Years of assigned deworming treatment during 1998-2003	3.29	3.32	3.25	3.22	0.069	-0.077
	(1.83)	(1.82)	(1.83)	(1.85)	(0.090)	(0.106)
Assignment to the deworming treatment group	0.675	0.681	0.665	0.664	0.006	-0.016
	(0.468)	(0.466)	(0.473)	(0.472)	(0.020)	(0.030)
Group 1 school	0.357	0.355	0.354	0.362	-0.015	-0.001
	(0.479)	(0.479)	(0.479)	(0.481)	(0.025)	(0.033)
Group 2 school	0.318	0.326	0.311	0.302	0.021	-0.015
	(0.466)	(0.469)	(0.463)	(0.459)	(0.021)	(0.033)
Number of observations [a]	7,530	4,891	421	2,218	7,530	5,312

Notes: The regression results in the Found (Regular and Intensive) – Not Found column reweights appropriately for intensive tracking. [a] The number of observations is correct except for the Age (1998) variable, which has somewhat more missing data.

Supplementary Appendix Table A5: Heterogeneous deworming impacts, full sample

	Hours worked last 7 days, all sectors				Number of meals eaten yesterday			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Deworming Treatment indicator	1.527	3.419**	1.420	1.664	0.096***	0.127***	0.069*	0.071**
	(1.027)	(1.473)	(1.172)	(1.163)	(0.028)	(0.041)	(0.035)	(0.031)
Female	-6.629***	-3.960**	-6.605***	-6.623***	0.078***	0.122**	0.078***	0.079***
	(0.942)	(1.710)	(0.927)	(0.948)	(0.026)	(0.052)	(0.026)	(0.026)
Female * Treatment		-3.942*				-0.064		
		(2.001)				(0.059)		
Grades 5-7 in 1998			7.665***				-0.006	
			(1.621)				(0.030)	
Grades 5-7 * Treatment			0.127				0.060	
			(1.975)				(0.041)	
Moderate-heavy worm infection rate at the zonal level (1998), demeaned				3.457				-0.429***
				(6.844)				(0.149)
Moderate-heavy infection rate * Treatment				3.270				0.027
				(7.482)				(0.211)
Deworming treatment pupils within 6 km (in ‘000s), demeaned	1.705	1.735	1.884	1.883*	0.080***	0.081***	0.082***	-0.002
	(1.425)	(1.415)	(1.424)	(0.978)	(0.023)	(0.023)	(0.024)	(0.033)
Total primary school students within 6 km (in ‘000s), demeaned	-0.989	-1.014	-1.160	-1.201*	-0.070***	-0.070***	-0.071***	-0.007
	(1.124)	(1.116)	(1.128)	(0.677)	(0.018)	(0.018)	(0.018)	(0.023)
Cost-sharing school (2001) indicator	-1.493*	-1.540*	-1.521*	-1.489*	-0.069**	-0.070**	-0.070**	-0.062*
	(0.846)	(0.832)	(0.860)	(0.887)	(0.031)	(0.031)	(0.031)	(0.033)
R^2	0.061	0.062	0.055	0.059	0.035	0.035	0.032	0.029
Observations	5,084	5,084	5,084	5,084	5,083	5,083	5,083	5,083
Mean (s.d.) in control group	18.4	18.4	18.4	18.4	2.16	2.16	2.16	2.16
	(23.1)	(23.1)	(23.1)	(23.1)	(0.64)	(0.64)	(0.64)	(0.64)

Notes: The sample used in columns (1)-(8) include all individuals surveyed in the KLPS2 with data for the relevant dependent variable. All observations are weighted to maintain initial population proportions. Additional controls include baseline grade fixed effects, geographic zone fixed effects, the mean pre-program school test score, baseline school population, survey wave indicator, and month of interview fixed effects. Standard errors are clustered by school. Significant at 90% (*), 95% (**), 99% (***) confidence.

Supplementary Appendix Table A6: Proportion of Individuals Working in Multiple Sectors

	In school	Work in agriculture	Work for wages	Work in non-agricultural self-employment
Panel A: Total				
In school	*0.252*	0.132	0.008	0.000
Work in agriculture		*0.555*	0.055	0.049
Work for wages			*0.166*	0.004
Work in self-employment				*0.100*
Panel B: Males				
In school	*0.303*	0.159	0.012	0.000
Work in agriculture		*0.528*	0.083	0.052
Work for wages			*0.227*	0.005
Work in self-employment				*0.110*
Panel C: Females				
In school	*0.195*	0.102	0.003	0.000
Work in agriculture		*0.586*	0.024	0.045
Work for wages			*0.098*	0.002
Work in self-employment				*0.089*

Notes: This table explores the proportion of individuals in the control group (Group 3) working in different economic sectors. Individuals are considered to be "working in agriculture" if they devoted positive hours to agriculture in the week preceding the survey. Individuals are considered "working for wages" or "in self-employment" if they received labor earnings or had positive self-employed profits (respectively) in the month preceding the survey, and devoted positive hours to that activity in the week preceding the survey. The diagonal entries (in italics) present the total proportion of control group respondents working in that sector.

Supplementary Appendix Table A7: Deworming impacts on occupation, within the wage earner subsample

	Coefficient estimate (s.e.) on deworming treatment indicator			Coeff. est. (s.e.) externality term	Control group mean; *Number of Observations*		
	All	Male	Female	All	All	Male	Female
Agriculture	-0.015	-0.007	-0.016	-0.017	0.021	0.008	0.052
	(0.013)	(0.015)	(0.032)	(0.022)	*706*	*540*	*166*
Casual/Construction laborer	-0.038**	-0.020	-0.072	-0.020	0.029	0.018	0.059
	(0.018)	(0.015)	(0.050)	(0.017)	*706*	*540*	*166*
Fishing	-0.023	0.022	-0.047	-0.135	0.192	0.242	0.064
	(0.060)	(0.066)	(0.030)	(0.084)	*706*	*540*	*166*
Manufacturing	0.072***	0.090***	0.059	0.041	0.030	0.031	0.028
	(0.024)	(0.033)	(0.048)	(0.031)	*706*	*540*	*166*
Retail and wholesale trade	0.005	-0.027	0.096	0.047	0.182	0.190	0.160
	(0.046)	(0.051)	(0.083)	(0.046)	*706*	*540*	*166*
Services (all)	0.032	0.003	-0.026	0.037	0.423	0.341	0.633
	(0.054)	(0.054)	(0.096)	(0.075)	*706*	*540*	*166*
Domestic	-0.012	0.018	-0.174	-0.026	0.117	0.030	0.340
	(0.032)	(0.020)	(0.110)	(0.038)	*706*	*540*	*166*
Restaurants, cafes, etc.	-0.029	-0.019	-0.040	0.024	0.061	0.042	0.110
	(0.023)	(0.026)	(0.042)	(0.034)	*706*	*540*	*166*
Trade contractors	-0.005	-0.018	-0.004	0.060	0.093	0.128	0.004
	(0.028)	(0.040)	(0.009)	(0.044)	*706*	*540*	*166*

Notes: The sample includes all individuals surveyed in the KLPS2 who report working for pay (with earnings greater than zero) at the time of the survey. Each entry is from a separate OLS regression. For details on the regressions, see the “Regression notes” for Table 2.

Supplementary Appendix Table A8: Average characteristics of occupations within wage employment

	Mean (s.d.) in Control Group		
	Hours per week worked in sector	Days of work lost to poor health*	Earnings in sector, past month (KSh)
Agriculture	13	2.1	618
	(12)	(1.9)	(258)
Casual/Construction laborer	51	0.4	2,246
	(31)	(1.0)	(1,576)
Fishing	37	2.1	3,114
	(25)	(4.2)	(1,729)
Manufacturing	53	1.1	5,311
	(24)	(1.8)	(3,373)
Retail and wholesale trade	40	0.9	2,462
	(27)	(2.0)	(2,349)
Services (all)	49	1.3	4,741
	(22)	(2.6)	(5,016)
Domestic	61	1.5	3,047
	(17)	(2.5)	(1,754)
Restaurants, cafes, etc.	53	1.2	4,194
	(21)	(2.5)	(3,567)
Trade contractors	27	0.8	3,172
	(22)	(2.5)	(2,170)

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

Supplementary Appendix Figure A1: Project Timeline of the Primary School Deworming Program (PSDP) and the Kenya Life Panel Survey (KLPS)