Schooling, Child Labor, and the Returns to Healthcare in Tanzania

Achyuta R. Adhvaryu
Anant Nyshadham

ABSTRACT

We study the effects of accessing better healthcare on the schooling and labor supply decisions of sick children in Tanzania. Using variation in the cost of formal-sector healthcare to predict treatment choice, we show that accessing better healthcare decreases length of illness and changes children's allocation of time to school and work. Children attend school for more days per week—but not for more hours per day—as a result of accessing better healthcare. There are no significant effects on child labor, but the results suggest that time spent in physically strenuous activities such as farming and herding increases.

I. Introduction

Child labor is common in many developing countries (Edmonds 2008). The demand for, and acceptability of, child labor in these settings generates a tradeoff between time spent in school (an investment that reaps future benefits) and time spent at work (an investment with short-run benefits) in the child's time allocation problem. Further, a growing number of studies have demonstrated that child labor matters. For example, it has been shown that child labor reduces schooling (Kruger 2007) and has effects on health and wages which can persist into adulthood (Beegle, Dehejia, and Gatti 2009).[1]

Most studies in this literature seek to estimate the causal effect of one of these activities—that is child labor or schooling—on the other. They do so either by

Achyuta R. Adhvaryu is an assistant professor of public health at Yale University. Anant Nyshadham is a graduate student in the Department of Economics at Yale University. *The authors thank Prashant Bharadwaj, Michael Boozer, Eric Edmonds, Jason Fletcher, James Fenske, Fabian Lange, T. Paul Schultz, Chris Udry, and participants at the Yale Labor/Public Finance Lunch and the IZA Workshop on Child Labor for helpful comments. The data used in this article can be obtained beginning October 2012 through September 2015 from Achyuta Adhvaryu, 60 College Street, New Haven, CT 06510, (achyuta.adhvaryu@yale.edu web: http://www.yale.edu/adhvaryu).*
[Submitted October 2010; accepted May 2011]
ISSN 022-166X E-ISSN 1548-8004

1. See Edmonds (2008) for an excellent review of the related literature.

exploiting shocks to the returns to labor—for example the profitability of farm labor or local labor market conditions (Beegle Dehejia, and Gatti 2009)—or to the price or benefits of school (Ravallion and Wodon 2000). We pose a related question, which has received less attention than these others: How do productivity shocks that affect the *child* change the allocation of time across both schooling and labor? The answer is not obvious, given that labor productivity and productivity in school are likely positively correlated. For example, deworming campaigns (Miguel and Kremer 2004) or nutritional interventions (Martorell, Habicht, and Rivera 1995) may have beneficial effects on achievement in school *and* productivity on the farm. Understanding the degree to which time allocated to each sector responds is important in quantifying the returns to such interventions.

We study the effects of shocks to child productivity (in both school and work) by exploiting variation generated by acute illness and its corresponding treatment. Specifically, we document how accessing formal-sector healthcare speeds up recovery from acute illnesses and shifts the amount of time allocated to both school and work for children, using household survey data from an area in northwest Tanzania. The premise of our study is that access to better healthcare, if it improves health outcomes for sick children, should, via its effects on productivity, shift time allocations as well. This hypothesis seems plausible for three reasons.

First, a large body of literature demonstrates that better health improves attendance and performance in school for children in developing countries (Alderman et al. 2001). It stands to reason, then, that investments in better healthcare for acutely sick children may affect these children's educational outcomes. Second, since much of child labor involves agricultural work, which is often quite physically strenuous, the link between health and labor productivity is likely strong.[2] Third, there is a large amount of heterogeneity in quality across healthcare choices in our setting (Das, Hammer, and Leonard 2008). Choosing formal-sector healthcare over informal (or no) care for a child's illness could thus generate significant shifts in health and productivity via the large difference in quality across those two healthcare options.

The main difficulty in estimating these effects arises from a well-known self-selection problem. Certain individuals—for example, those with more severe illnesses, higher preferences for health, or greater access to financial resources—are more likely to select into better healthcare options. Thus, comparing outcomes across individuals who used different healthcare options will lead to a biased estimate of the impact of using better healthcare. Moreover, fixed effects estimators cannot adequately address this problem because they do not control for unobserved severity, which varies by illness episode (that is, within multiple observations of the same individual), and which other studies have shown is the most salient bias for acute shocks (Gowrisankaran and Town 1999; Cutler, Huckman, and Landrum 2004).

We overcome this self-selection problem by using an instrumental variables (IV) strategy that exploits exogenous variation in cost of formal-sector care. Following Adhvaryu and Nyshadham (2011a), we propose an interaction instrument. We interact a dummy variable for the presence of a formal-sector health facility in one's community with the number of days of rainfall in the month of the individual's

2. See Edmonds' (2008) handbook chapter on child labor for details on the breakdown of labor activities.

sickness, and exclude only this interaction from the second stage, while controlling for the main effects of facility "existence" and days of rainfall in the first and second stages of a two-stage instrumental variables estimator. We find that the instrument is sufficiently predictive in the first stage; it is also robust to a variety of additional controls, and passes various falsification tests, all of which are discussed in detail in Section V.

Using this strategy, we first verify that using formal-sector healthcare does induce a significantly speedier recovery for acutely sick school-aged children (ages 7–19 inclusive). We then employ individual time use data from the week preceding survey to estimate effects on hours (and days) spent in school and hours spent in various activities. The results show a large increase in school hours (the point estimate is about 27 hours) as a result of accessing formal-sector healthcare. This essentially amounts to a 100 percent increase at the mean, suggesting that formal-sector care works on the extensive margin. Consistent with this interpretation, we find that the probability of any school hours in the last week and the number of days the child attended both increase significantly, while the hours per day spent in school (conditional on having attended in the past week) do not change. Together, these results suggest that the main barrier to schooling for sick children, at least in terms of attendance, is actually getting to school. Of course, performance in school, which we cannot measure in these data, may also be affected by fluctuations in productivity induced by health shocks.

On child labor variables, we find no significant effects, though the effects on some labor activities are large but imprecisely estimated. Though we cannot interpret the coefficients with certainty, the pattern of labor adjustment suggests an increase in total hours worked as a result of receiving better healthcare, with most of the increase coming from farm hours (in particular, hours spent farming or herding, which are among the most strenuous agricultural activities).[3]

Our study is different from previous work in two ways. First, we focus on treatment for acute illness. Mitigating the effects of acute health shocks likely leads to a different time use response than nutritional or long-term care interventions, whose effects on health and time allocation have been studied to date (Thomas et al. 2006; Thirumurthy, Graff Zivin, and Goldstein 2009).

Second, we focus on the effects of using formal-sector healthcare. The only other study to our knowledge that studies such effects in a developing country context is Dow et al. (1997), which evaluates a healthcare price experiment in Indonesia. While Dow et al. (1997) successfully link shifts in the localized price of healthcare to changes in labor outcomes, their analysis, as a product of the experimental design, is reduced-form: The study is not able to estimate the structural effect of *choosing into* higher-quality healthcare among those on the margin. Our analysis is the first in the developing country setting to estimate the effects of choosing higher-quality healthcare on the health and labor supply outcomes of sick individuals. Similarly, the effects of nutrition on school enrollment for children have been investigated in the developing country context (Glewwe and Jacoby 1995; Alderman et al. 2001).

3. We refer the reader to the FAO report on energy expenditure by agricultural activity type for more details (FAO 2001).

Neither nutritional status nor school enrollment, however, are short-term state variables. We examine the effects of acute illness on "acute" schooling outcomes, which are most likely to respond to short-term fluctuations in health status.

We believe our study makes three contributions. First, no other paper, to our knowledge, examines the effects of individual productivity shocks on time allocations in both labor and schooling. As discussed above, most of the literature to date focuses on estimating the various impacts of child labor. Second, studying how individual productivity shocks affect the differential returns to schooling vis-a-vis labor helps us understand the interaction of health and the time allocation of children. Third, in settings where school attendance is not mandatory for children, the effects of formal-sector healthcare on children go beyond just health: We show that attendance in school can increase dramatically, which is an added return to investments in formal-sector healthcare infrastructure or transport.

The remainder of the paper is laid out as follows. Section II describes our dataset. Section III presents our identification strategy and discusses its validity. Section IV presents the empirical results. Section V presents a variety of robustness checks and falsification tests. Finally, Section VI concludes.

II. Data

A. Overview

This study uses survey data from the Kagera region of Tanzania, an area west of Lake Victoria, and bordering Rwanda, Burundi, and Uganda. Kagera is mostly rural and primarily engaged in producing bananas and coffee in the north, and rain-fed annual crops (maize, sorghum, and cotton) in the south. The Kagera Health and Development Survey (KHDS) was conducted by the World Bank and Muhimbili University College of Health Sciences. The sample consists of 816 households from 51 "clusters" (or communities) located in 49 villages covering all five districts of Kagera, interviewed up to four times, from Fall 1991 to January 1994, at 6–7-month intervals. The randomized sampling frame was based on the 1988 Tanzanian Census.[4] There was moderate attrition from the longitudinal sample. 9.6 percent of households sampled in Wave 1 were lost by Wave 4. However, to preserve balancing across health profiles in the sample, lost households were replaced with randomly selected households from a sample of predetermined replacement households stratified by sickness.[5]

4. A two-stage, randomized stratified sampling procedure was employed. In the first stage, census clusters (or communities) were stratified based on agro-climactic zone and mortality rates and then were randomly sampled. In the second stage, households within the clusters were stratified into "high-risk" and "low-risk" groups based on illness and death of household members in the 12 months before enumeration, and then were randomly sampled.

5. We should not be too concerned about the possibility of bias due to attrition in our results. Our sample is a constructed cross-section of child-year observations, selected in each year on the basis of the reporting of acute illness. That is, selection into our sample is a combination of selection into the resurveyed sample in each wave (nonattrition) and selection into the reporting of acute illness in each wave). While selection into our sample could be correlated with probability or severity of acute illness or health endowments or

KHDS is a socioeconomic survey following the model of previous World Bank Living Standards Measurement Surveys. The survey covers individual-, household-, and cluster-level data related to the economic livelihoods and health of individuals, and the characteristics of households and communities. In the following paragraphs, we outline the variables we use in our analyses.

B. Health Variables

In the health module of the KHDS, all household members are asked about chronic illnesses and acute illness episodes; care sought for these episodes; and current illness (at the time of survey).[6] As our main sample restriction, we use information on whether individuals were sick with an *acute illness*, that is, one which began 14 days or less before the date of survey. We also restrict our attention to school-aged children between seven and 19 years of age (inclusive), for whom time use data is collected.

Table 1 shows summary statistics for the sample of sick school-aged children. Across the four waves of the survey, the data report 1,954 child-year observations of sickness (that is, each time an individual aged 7–19 reports being sick he is counted as a child-year observation). This acutely ill subsample of school-aged children-year observations makes up roughly 27.5 percent of the total sample.

Of the sample of acutely ill child-year observations, roughly 37 percent report still being ill at the time of survey. About 22 percent of this sample sought formal-sector healthcare for their illness episode, where formal-sector healthcare is defined as care at a hospital, health center, or dispensary (which includes government, NGO, and private facilities). Of the subsample of formal-sector care users, 12.7 percent went to health facility, 28.6 percent went to a hospital, and the remaining percentage went to a dispensary.

Columns 2 and 3 of Table 1 show summary statistics for the subsamples of formal-sector care users and nonusers, respectively. Notice that the probability of still being ill at the time of survey is slightly higher among those who did *not* visit formal-sector care. Of course, this comparison of means is not evidence of a causal relationship between formal-sector care and speedy recovery from acute illness; however, it will motivate a more in-depth analysis below.

One interesting comparison we can make, however, is the difference in the means of symptoms reported across subsamples of formal-sector care users and nonusers. We see in Columns 2 and 3 that those who seek formal-sector care are more likely to report fever than those who seek informal care or no care at all. To the degree that fever proxies for severe illnesses such as malaria and systemic infections, this provides suggestive evidence of selection in formal-sector care use on the basis of severity of illness. Severity bias is a primary motivation of our empirical strategy discussed in Section III below.

preferences and also school and labor outcomes, we would only be concerned to the degree that our instrument affected this nonattrition and/or reporting of illness. We show below that our instrument is seemingly orthogonal to selection into our sample (see Table VII), which is a combination of nonattrition and selection into sickness.

6. In the case of individuals below the age of 15, the primary caretaker of the child is asked to answer on the child's behalf.

C. School and Labor Variables

The time use module of the KHDS collects detailed information on various types of productive activity for all individuals seven years of age and older. Individuals are asked how many hours in the past seven days they spent in each of a variety of activities. We construct a composite variable for total labor hours in the week preceding survey, as well as breakdowns into several important types of labor activities.

In particular, we first split total labor hours into farm hours and nonfarm hours. Then, we further split farm hours into time spent in the field (farming or herding), time spent processing agricultural and livestock products, and time spent in wage employment (mostly on the farms of other village members), and nonfarm hours into self-employment hours and home hours. Self-employment includes any nonfarm activities producing profit that accrues to the individual Self-employment is contrasted with working for someone else's business. It may include household enterprise, production or sale of market goods, or owning another type of small business (restaurant, hotel, etc.). Home hours include time spent in household chores and time spent collecting water and firewood.

We see in the summary statistics reported in Table 1 that sick school-aged children spend roughly 21 hours per week allocated to labor activities and 24 hours allocated to schooling. Only 82 percent of sick school-aged children spent any time in school in the week prior to survey. The average number of days in which sick children spent any time in school in the week prior to survey is 3.5. Conditional on attending some school in the week prior to survey, sick children spent on average roughly 6.7 hours per day in school. Interestingly, the means of schooling outcomes are quite similar across formal-sector users and nonusers.

Labor outcomes are quite similar across healthcare users and nonusers as well. Among school-aged sick children, farm labor makes up roughly a third of total labor hours and nonfarm makes up two-thirds. Sick children allocate 8.6 hours to farm labor and 12.7 hours to nonfarm labor, on average. Within farm labor, the majority of hours are spent in the field; while the vast majority of nonfarm labor hours are spent performing home chores.

This lack of variation in mean school and labor hours across subsamples of those who did and did not seek formal-sector care could be evidence of one of two possible cases: Either formal-sector care has no effect on labor supply and school attendance, and perhaps even no effect on health outcomes in this population, or the choice of healthcare is endogenous (for example, on the basis of unobserved severity) rendering a simple comparison of means across healthcare choice subsamples and even ordinary least squares (OLS) estimates useless in investigating the effects of healthcare choice on health, labor supply, and schooling outcomes. As mentioned above, the presence of severity bias in OLS estimates is well-established in the literature. Therefore, in what follows, we will propose and employ an empirical strategy which accounts for this bias (see Section III).

D. Other Individual-, Household-, and Cluster-level Variables

We use a variety of individual-, household-, and community-level demographic and socioeconomic characteristics in our regressions. The most important for the purposes of our analysis is the existence (or, to be precise, the lack of existence) of

Table 1
Summary Statistics

	Sick Children						Children from Nonsick Households	
	All		Formal Care		No Formal Care			
	Count		Count		Count		Count	
Child-year observations (illness in past two weeks, 7 < = age < = 19)	1,954		434		1,520		1,240	
Currently enrolled in school	1,224		290		934		657	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Health status and care								
Still ill	0.368	0.482	0.329	0.471	0.379	0.485		
Visited formal healthcare	0.222	0.416	1.000	0.000	0.000	0.000		
Symptoms								
Fever	0.362	0.481	0.530	0.500	0.314	0.464		
Cough	0.284	0.451	0.200	0.401	0.308	0.462		
Headache	0.265	0.441	0.339	0.474	0.244	0.430		
Schooling of school-aged children (aged 7–19, currently enrolled)								
Hours spent in school (in week before survey)	23.813	15.010	24.609	15.800	23.565	14.756	25.538	15.542
Any time in school (in week before survey)	0.825	0.380	0.824	0.381	0.825	0.380	0.833	0.374
Days spent in school (days with > 0 hours, week before survey)	3.579	1.954	3.459	2.016	3.617	1.934	3.799	1.965
Hours per day in school (in week before survey)	6.655	1.839	7.129	1.670	6.508	1.865	6.722	1.868

Labor supply of school-aged children (aged 7–19, hours in week before survey)								
Total labor	21.278	17.845	20.438	19.113	21.517	17.465	22.715	19.195
Farm	8.623	10.581	8.057	10.890	8.784	10.489	9.574	12.330
Field	7.643	9.037	6.895	9.123	7.856	9.004	8.873	11.255
Processing	0.065	0.666	0.073	0.629	0.063	0.676	0.089	0.797
Employment	0.915	5.857	1.090	6.525	0.864	5.653	0.613	5.039
Nonfarm	12.655	12.322	12.381	13.318	12.733	12.026	13.141	12.263
Self-employment	0.841	7.546	0.912	8.634	0.821	7.209	0.376	3.637
Home	11.814	9.918	11.469	10.024	11.912	9.888	12.765	11.666
Costs of Healthcare (Instruments)								
Number of days of rain in month of survey	7.998	5.313	8.288	5.495	7.915	5.259	8.294	4.792
No health facility in community	0.622	0.485	0.415	0.493	0.681	0.466	0.669	0.471
Resources in Community								
Daily market	0.622	0.485	0.654	0.476	0.613	0.487	0.551	0.498
Periodic market	0.338	0.473	0.297	0.458	0.349	0.477	0.390	0.488
Drivable road	0.966	0.181	0.963	0.189	0.967	0.178	0.974	0.159
Public transport	0.267	0.443	0.316	0.465	0.253	0.435	0.317	0.465
Secondary school	0.084	0.277	0.108	0.311	0.077	0.267	0.101	0.301
Bank	0.107	0.309	0.099	0.299	0.109	0.312	0.103	0.304
Post office/telephone booth	0.136	0.343	0.166	0.372	0.128	0.334	0.144	0.352
Demographic characteristics								
Age	12.843	3.552	13.147	3.443	12.756	3.579	12.914	3.677
Household size	7.623	3.560	7.737	3.517	7.591	3.573	6.335	2.574
Female	0.506	0.500	0.521	0.500	0.502	0.500	0.489	0.500
Household assets (25 quantiles)	13.057	7.175	13.094	7.427	13.047	7.103	12.115	7.220

Notes: The sample, unless otherwise noted, is made up of individuals aged 7–19 (inclusive) who reported illnesses that began in the two weeks prior to survey. The schooling sample is further restricted to individuals who are currently enrolled in school.

a formal healthcare facility in the cluster.[7] As Table 1 reports, about 62 percent of sick individuals lived in communities without a formal-sector healthcare facility. Among those who did not seek formal-sector care, 68 percent lived in a community without a formal-sector care facility; among those who sought formal-sector care, the percentage is much lower at only 42 percent.

As we describe in Section III, we accordingly control for the direct effect of the lack of a health facility in the community, along with a variety of other variables related to the existence of resources in one's community (existence of a daily market, periodic market, driveable road, public transportation, secondary school, number of primary schools, bank, and post office/telephone). Table 1 shows that access to these resources in general appears to be (insignificantly) greater for those who chose formal-sector care.

Indeed, this fact is corroborated by the positive correlation between the existence of a health facility in one's community and the existence of the other above-mentioned resources. This, of course, invalidates the use of the existence of a health facility alone as an instrument for healthcare choice and motivates the inclusion of the potentially correlated existence of other resources in the community as controls as well.

We also control for the distance to various types of formal-sector care options if they are not in the individual's community; in particular, we include the distances to the nearest dispensary, health facility, and hospital (please note that if these options are in the individual's cluster, this variable equals 0).[8]

We include individual-level controls for the number of days before date of survey the individual's illness began (deciles and linear term in days). We also include fixed effects for the interaction of gender and years of completed schooling (quintiles), age (deciles and linear term in years), and year of survey. Household-level controls include household size (deciles); total assets owned by the household (25 quantiles of an asset index generated using principal components analysis); and a dummy for whether the survey took place during one of two rainy seasons.

In the last four rows of Table 1, we find no evidence of significant differences in demographic composition across healthcare choice subsamples. While the empirical strategy proposed below ought to be robust to such demographic differences, their absence is preliminary evidence of the relative importance of access to healthcare as a primary mover of healthcare choice and of healthcare choice as a primary determinant of health outcomes, at the least.

Column 4 of Table 1 shows corresponding statistics for the sample of nonsick school-aged children from households with no acutely ill members. We might expect that these children form an appropriate comparison group to children who have completely recovered from an acute illness. Of course, to the extent that children from nonsick households are systematically different (perhaps, with respect to access

7. We use the lack of existence of a formal health facility (instead of simple existence) because we expect the effect of existence on the probability of visiting a formal-sector health facility to be positive, whereas we expect rainfall to have an incrementally negative effect for individuals living far from a facility. Thus, we construct the slightly awkwardly termed "nonexistence" variable for ease of interpretation of the coefficient on the interaction instrument.

8. In our empirical specification, we control for quintiles of the distance to each option separately.

to health and other resources, health preferences or endowments, etc.), their labor and schooling outcomes may not entirely resemble those of recovered children.

Indeed, children from nonsick households tend to be younger, are more likely to be male, and belong to smaller and less wealthy households. Nonsick households also seem to be located in communities with less access to healthcare and daily markets; though access to other resources seems to be roughly the same for children from sick and nonsick households. Acutely ill children who seek formal-sector care spend more total hours on average in school during the week prior to survey than those who do not seek formal-sector care, though not quite as many hours as children from nonsick households. Formal-sector care users also spend less time in the field and in home chores than do children from nonsick households. This pattern could suggest that children from nonsick households have higher health endowments and are, therefore, generally more productive than the acutely ill sample.

E. Rainfall Data

We obtained monthly rainfall data from the Tanzania Meteorological Agency spanning from 1980 to 2004.[9] The data set includes the amount of rainfall (in millimeters) per month and total days with rainfall per month for 21 weather stations in Kagera region. The data set provides a matching file which reports the closest and second closest weather station to each cluster in the KHDS sample. Two measures of “closest” are available: a straight-line distance between each cluster and each rainfall station, and a distance measure that takes into account the location topology of the area.

We use the straight-line measure definition of “closest” and use the number of days of rainfall in the month the individual was sick as the primary measure of rainfall in our regressions. Further, we match the rainfall observation to the sick individual by taking the rainfall value in the month the individual was surveyed, in the cluster of the individual’s residence. If the rainfall value for this cluster-by-month observation is missing, we use the value at the second closest rainfall station to the cluster.

There appears to be, as shown in Table 1, significant variation in the number of days of rainfall across all samples. While the means across healthcare choice subsamples are only minimally different, it is interesting to note that the mean days of rainfall is slightly larger on the subsample of sick children who visited formal-sector care, corresponding to a role of severity in healthcare choice. That is, if more rainfall corresponds to more severe illness, and the more severely ill, in turn, are more likely to choose formal-sector care, we would expect to see a larger mean number of days of rainfall in the sample of sick individuals who ultimately chose to visit formal care. This, of course, invalidates the use of rainfall alone as an excludable instrument for healthcare choice and motivates its inclusion as a control.

We also control for the number of days of rainfall in the month *prior to* the individual’s sickness (as we discuss in Section III); the historical mean and historical standard deviation of the distribution of rainfall in the given month, computed over

9. The data set is downloadable from the EDI-Africa website: http://www.edi-africa.com/research/khds/introduction.htm.

all the years of available data for the month in question (quadratic terms of these variables are included as well); fixed effects for the closest rainfall station; deciles for the number of days of rainfall; deciles for the amount of rainfall (in millimeters) in the month the individual fell sick; and interactions of days of rainfall with the existence of resources variables defined in the previous subsection. The importance of these interaction controls will be discussed in Section III) below. For further details on the construction of rainfall variables, please see the Data Appendix.

III. Empirical Strategy

Our goal in this section is to propose and discuss the validity of an instrument for healthcare choice, and to discuss how we use the variation induced by the instrument to measure the effects of healthcare choices first on health outcomes and then on time spent in school and labor supply.

A. An Instrument for Healthcare Choice

Let O_{ij} denote an outcome for individual i in cluster j, let h_{ij} denote the individual's healthcare choice, and let X_{ij} denote a vector of individual-, household- and community-level characteristics. Consider the following empirical model:

$$(1) \quad O_{ij} = \beta h_{ij} + X_{ij}^{'}\gamma + \varepsilon_{ij}.$$

Measuring the relationship between healthcare choice and health outcomes (or schooling and labor outcomes, for that matter) as shown above in Equation 1 likely results in a biased estimate of the effect of h on O, due to unobserved determinants of outcomes in the error term ε that are correlated with healthcare choice. In particular, the severity of the health shock likely influences the care option chosen (that is, individuals with higher-severity illnesses will choose into higher-quality healthcare options) as well as the outcome (higher-severity illnesses will generate worse health, labor, and schooling outcomes).

To address these endogeneity concerns, we use an instrument for healthcare choice that exploits exogenous variation in the costs of formal-sector healthcare. The instrument builds on the methodology introduced in Adhvaryu and Nyshadham (2011a). A major point discussed in that paper is the fact that the largest costs of formal-sector care in developing countries are often those associated with the opportunity cost (or the direct costs) of travel to the care facility. This is particularly true in developing contexts, like that in which this study's empirical analysis is conducted, in which nominal fees for healthcare are heavily subsidized. Distance to the nearest facility (or alternatively, the presence of a formal-care facility in one's community) is thus a large determinant of healthcare choice in developing countries, through its effects on costs (Gertler, Locay, and Samuelson 1987; Mwabu, Mwanzia, and Liambila 1995; Mwabu 2009).

Given the importance of proximity to formal-sector care, one might argue that this variable would be a good candidate for an instrument for healthcare choice, particularly in developing country settings. However, it is likely, due to endogenous placement of facilities on the basis of a local population's health stock, that the existence of a facility in one's community and distance to the nearest facility are correlated with the error term in a second-stage regression with health or labor supply outcomes as dependent variables. Later in this section, we present some evidence that this is the case in our context, as well.

Following Adhvaryu and Nyshadham (2011a), we propose an interaction instrument. Specifically, we interact a dummy variable for the absence of a formal-sector health facility in one's community with the number of days of rainfall in the month of the individual's sickness, and exclude only this interaction from the second stage, while controlling for the main effects of facility "existence" and days of rainfall in the first and second stages of a two-stage instrumental variables estimator.

The two stages of analysis are specified as follows. Define $NoFac_j$ to be a dummy variable that equals 1 if *no* formal-sector health facility exists in cluster j, and R_{ij} to be the number of days of rainfall in cluster j at the time of individual i's sickness.[10] The two-step estimator is written as follows:

$$(2) \quad 1st\,stage{:}\; h_{ij} = \alpha_1(NoFac_{ij}xR_{ij}) + \alpha_2 NoFac_{ij} + \alpha_3 R_{ij} + X_{ij}^{'}\alpha_4 + \zeta_{ij}$$

$$(3) \quad 2nd\,stage{:}\; O_{ij} = \beta_1 h_{ij} + \beta_2 NoFac_{ij} + \beta_3 R_{ij} + X_{ij}^{'}\beta_4 + \varepsilon_{ij}$$

The intuition behind the instrument is simple. The main effects of facility nonexistence and days of rainfall are likely both negative; that is, not having a facility in one's community and being exposed to more rainfall should, for the purposes of travel costs, discourage formal-sector health facility usage for individuals seeking care. Moreover, heavier rains should discourage individuals who live farther away from a health facility *more* than individuals living in a community with a health facility.

Imagine one household who lives next to a facility, while another is located many villages away. In times of dry weather, clearly the household in the community with a health facility will be more likely to choose formal-sector care than the one farther away. But in times of heavy rains, the rain should incrementally deter the farther household even *more* than it does the one just next door.

It should be noted that the panel structure of the data allows for a fixed effects model to be conducted. We do, in fact, include fixed effects for the rain station—matched to each household's cluster in the data. The rain station fixed-effects partition the region into 18 climatically similar areas. Of course, no fixed-effects specification can purge the estimates of bias deriving from unobserved severity of the illness. Fixed effects, particularly at the individual or household level, can only account for bias due to unobserved static heterogeneous determinants of healthcare use and outcomes, such as health preferences or endowments.

Preferences, endowments, and severity of illness (or more precisely, selection into formal-sector care on the basis of these unobserved characteristics) are the most probable sources of bias in estimates of the effects of formal-sector care on health and other outcomes. Given that our instrument is orthogonal to all three sources of

10. We define the facility "existence" variable in the negative in order to make interpretation of the interaction coefficient easier; of course, changing this variable to reflect the existence of a health facility as opposed to the lack of existence has no effect on the estimation procedure or the results (barring changing the sign of the coefficients on the interaction term and the main effect of facility existence).

bias, the inclusion of individual or household fixed effects is not particularly beneficial. Furthermore, individual (or household) fixed effects would only identify the effects of healthcare choice on outcomes using variation over time in healthcare choice among individuals (or households) who report being acutely ill in multiple waves. These select individuals or households who report multiple illness spells are likely unrepresentative of the acutely ill sample as a whole.

B. Instrument Validity

Ideally, we would like variation in the instrument to be equivalent to experimental variation in the price of formal-sector care. That is, we would like to answer the question, "Holding all other prices constant, if we shift only the price of formal-sector care, how does the demand for this care change, and subsequently, how do these shifts affect health, labor supply, and schooling outcomes?" One crucial element of our argument is thus that the interaction instrument must induce price changes solely in the cost of formal-sector care, as opposed to shifting other prices that determine access to other resources, as well as directly influence consumption and labor allocations.

1. Controlling for General Remoteness

It is plausible that fluctuations in rainfall induce shifts in the prices of nonhealthcare goods and services differentially across communities with health facilities as compared with communities without. For example, suppose nonexistence of a formal-care facility was correlated with a community's general remoteness; that is, communities lacking health facilities lacked access to other important resources (commodity and labor markets, roads, schools, etc.). Since rainfall, through the interaction instrument, acts as a randomized amplifier of the costs of access to formal-sector care, rainfall would amplify the costs of access to these other resources as well. If this were true, the instrument would not be excludable.

To address this problem, we control for the existence of a variety of important resources, as well as the interactions of these variables with days of rainfall.[11] Controlling for these main effects and interactions ensures that the variation induced by the instrument is specific to the costs of formal-sector care.

2. Isolating Transitory Rainfall Variation

Crucial to the interpretation of the instrument is the hypothesis that rainfall in the month of sickness induces a temporary, randomized amplifying effect in the costs of travel to formal-sector care. Moreover, rainfall generates a larger temporary effect in places where no formal-sector facility exists. To isolate this temporary variation from persistent high rainfall (which is a common phenomenon in our context given that rainy seasons in Tanzania last for months at a time and cause seasonal variation in incomes and opportunity costs of time), we control for the days of rainfall in the

11. For example, we include existence of a daily market, drivable road, public transport, secondary school, and post office or telephone; for a full listing of the variables included, please refer to the note at the bottom of Table II.

month *prior to* the individual's sickness, as well as the interaction of this variable with the facility existence dummy.

3. Nonlinear Effects of Endogenous Distance

Finally, we allow for the possibility that distance enters the first and second stages nonlinearly. We do this to further preclude the possibility that the interaction instrument is only capturing a nonlinear effect of distance, rather than the interaction of distance with a randomized, transitory source of variation. To account for this concern, we include quintiles of the distribution of distance to the nearest health facility, hospital, and dispensary in all regressions.

IV. Results

In this section, having defended the validity of the proposed empirical strategy in this context, we report and discuss results from first- and second-stage regressions of health, school, and labor outcomes on formal-sector healthcare use.

A. Health Outcomes

In Table 2, we explore the effects of formal-sector care use on subsequent health. The first Column of Table 2 presents results from the first-stage regression conducted on the sample of all school-aged children who reported being acutely ill. In the first-stage specification, we regress a binary for whether the sick individual chose formal-sector care on the proposed instrument of the interaction of days of rainfall in month of survey and a dummy for the lack of a formal-sector healthcare facility in the individual's community. The results in Column 1 of Table 2 show a significant reduction in the probability of a sick individual choosing formal-sector care when the interaction instrument increases. The F-stat on the instrument coefficient is nearly 10, with a p-value of just above 0.002.

In Column 2 of Table 2, we present results from the second-stage instrumental variables regression of a binary for whether the individual was still ill at the time of survey on a binary for whether he visited a formal-sector healthcare facility. The results show a large and significant reduction in the probability of still being ill at the time of survey for those sick individuals driven exogenously to formal-sector care. For sick individuals on the margin, being exogenously driven to visit a health facility decreases the probability of still being ill at the time of survey by approximately 86 percentage points.

The magnitude of these results corresponds to the results found in Adhvaryu and Nyshadham (2011a), which applies a similar analysis to a nationally representative sample of children under five in Tanzania. Note that the marked attenuation in the OLS estimates reported in the third column of Table 4 is also consistent with estimates from previous studies and corresponds to bias due to self-selection into formal-sector care on the basis of severity.

Table 2
Effects of Healthcare Choice on Probability of Still Being Ill at Time of Survey

	First Stage	Second-stage IV	OLS
	Formal Healthcare	Still Ill	Still Ill
Formal healthcare		−0.857**	0.0167
		(0.431)	(0.0286)
Days of rainfall × no facility	−0.0193***		
	(0.00616)		
No facility	−0.222**	−0.274	−0.000315
	(0.0952)	(0.186)	(0.0914)
Days of rainfall	−0.00346	0.0822**	0.0812**
	(0.0307)	(0.0391)	(0.0316)
F-test: rain × distance = 0	9.763		
Prob > f	0.00201		
Observations	1,954	1,953	1,953
Mean of dependent variable	0.226	0.367	0.367

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Standard errors are clustered at the sampling cluster by year-month level. All specifications include main effects of days of rainfall and "No Facility," assets (25 quantiles), rain station, and year of survey fixed effects; as well as dummies for deciles of household size, unless otherwise stated. Specifications also include, unless otherwise stated, education by gender fixed effects; dummies for deciles of age (along with a polynomial up to the third degree in age), days of rainfall and levels of rainfall as well as number of days before survey the illness started (dummies for deciles as well as linear terms); and dummies for quintiles of distance to nearest hospital, healthcare facility, and dispensary. Dummies for the existence of a daily market, periodic market, drivable road, public transport, secondary school, the number of primary schools, bank and post office/telephone are included; along with interactions of these dummies with days of rainfall, unless otherwise noted. Other controls include rainy season fixed effects; historical means and standard deviations of both rainfall and quadratic terms of these; days of rainfall in month prior to survey and its interaction with "No Facility", unless other noted. Sample is restricted, unless otherwise noted, to all individuals, aged 7–19 (inclusive), with illnesses that began in the two weeks prior to survey.

B. Schooling Outcomes

Now that we have established the fundamental links, first, between costs of formal-sector care and healthcare choice and, second, between higher-quality care and improved health outcomes, we turn our attention to the effects of formal-sector care on time spent in school in the week prior to survey. Specifically, if formal-sector care reduces the length of illness or in particular the probability that the individual is still ill on any subsequent day, then—to the extent that the illness had reduced the individual's total endowment of time, increased the effort necessary for or disutility from traveling to school, or reduced his marginal productivity in school—we should expect to see effects of formal healthcare on time spent in school.

In Table 3, we report estimates of the effects of formal-sector care use on hours, days, and hours per day spent in school in the week prior to survey as well as a

Table 3
Effects of Healthcare Choice on Hours and Days Spent in School Among Sick School-Aged Children

	First Stage	Second-Stage IV			
	Formal Healthcare	School Hours Last Week	Any School Last Week	Days Last Week	Hours per Day Last Week
Formal healthcare		27.45*	0.695*	3.626*	1.168
		(14.75)	(0.419)	(2.087)	(1.439)
Days of rainfall × no facility	−0.0266***				
	(0.00832)				
No facility	−0.206	19.84***	0.637***	2.620**	1.009
	(0.143)	(7.512)	(0.223)	(1.096)	(0.628)
Days of rainfall	−0.0272	−1.154	−0.00638	−0.256	0.102
	(0.0359)	(2.035)	(0.0558)	(0.292)	(0.162)
F-test: rain × distance = 0	10.20				
Prob > *F*	0.00161				
Observations	1,224	1,224	1,224	1,224	1,010
Mean of dependent variable	0.242	23.51	0.825	3.559	6.604

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). See Table II for additional comments. Sample in these specifications further restricted to children currently enrolled in school.

binary for whether the child went to school at all. The sample on which the schooling analysis is conducted is further restricted to acutely ill, school-aged children who also were enrolled in school. Column 1 of Table 3 reports first-stage regression results for the sample of school-aged children who reported acute illness and also reported being enrolled in school at the time of survey. Once again, the interaction instrument is strongly predictive of formal-sector care use. The F-statistic on the instrument coefficient is larger than 10 with a corresponding p-value of less than .002.

Columns 2–5 of Table 3 present results from second-stage regressions. In Column 2, we report estimates of the effects of formal-sector care use on total hours spent in school in the week prior to survey. We see a large and significant effect on hours spent in school. Sick children who are driven exogenously to seek formal-sector care spend more than 27 hours more in school in the week prior to survey than those who are not. Compared to a mean of roughly 24 hours per week, these results amount to an entire week's worth of schooling gained through the use of higher-quality care.

Column 3 of Table 3 shows estimates of the effects of formal-sector care on a binary for whether the child attended any school in the week prior to survey. Column 4 reports results from a second-stage IV regression of the number of days in the week prior to survey in which the child attended at least one hour of schooling on formal-sector care use. We again see large and significant effects of formal-sector care on time spent in school. Sick children who are driven exogenously to seek formal-sector care are nearly 70 percentage points more likely to attend some school in the week prior to survey, and attend roughly 3.6 days more than nonusers.

Interestingly, among sick children who attended at least one hour in one day of school, we find no significant effects on the number of hours per day spent in school, as reported in Column 5. These results suggest that the greatest cost of sickness as relates to schooling is the disutility of attending school at all on days in which the child is acutely ill. That is, a speedier and more complete recovery from acute illness due to formal-sector care use induces a child to attend more days of school, but not significantly more hours in the days he attends.

To the degree that traveling to and from school requires more physical effort than staying in school once there, these results provide suggestive evidence of the effects of acute illness on the capacity for physical effort. That having been said, because we do not have access to school performance data, we cannot comment on the effects of acute illness or subsequent healthcare use on productivity in school or returns to time spent in school.

C. Child Labor Supply

Now that we have seen significant effects of formal-sector care on time spent in school, we turn our attention to the other side of schoolwork decision. In Table 4, we present estimates of effects on total labor supply and allocations of time to various productive activities among school-aged children. Here the sample is identical to that used in the health outcomes regressions discussed above; that is, it is no longer restricted to children who are enrolled in school.

Table 4
Second-Stage IV: Effects of Healthcare Choice on Labor Supply of Sick School-Aged Children

	Total Labor Hours	Nonfarm Labor		
		Total	Self-Employment	Home
Formal healthcare	14.27 (17.30)	2.527 (8.751)	−0.00917 (5.026)	2.536 (7.289)
Observations	1,954	1,954	1,954	1,954
Mean of dependent variable	21.43	12.68	0.838	11.84

	Farm Labor			
	Total	Field	Processing	Employment
Formal healthcare	11.74 (11.73)	10.95 (9.920)	−0.490 (0.392)	1.280 (4.151)
Observations	1,954	1,954	1,954	1,954
Mean of dependent variable	8.748	7.792	0.0617	0.894

Notes: Robust standard errors in parentheses (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$). See Table II for additional comments. Sample is restricted, as noted in Table II, to children aged 7–19 (inclusive) who reported being ill in the two weeks prior to survey.

In the top panel of Table 4, we show results from second-stage IV regressions of total labor hours, nonfarm labor hours, and allocations of nonfarm hours to self-employment and home chores. We find large but insignificant effects on total labor hours in Column 1 of the top panel of Table 4. The point estimate of the effect of formal-sector care on total labor hours is roughly 14 hours as compared to mean labor supply of roughly 21.5 hours. Point estimates of the effects on total nonfarm labor hours and allocations to self-employment and home chores are much smaller and still insignificant. Overall, we find little evidence of an effect on nonfarm labor hours.

In the bottom panel of Table 4, we report results from second-stage IV regressions of farm labor hours and allocations to field and processing activities and wage labor. The point estimate on farm labor hours is again large, but insignificant (nearly 12 hours, as compared with a mean of fewer than nine hours). The majority of the effect on farm labor comes from an effect on hours spent in the field.

Of course, absent data on individual productivity, such as yield per unit time or piece rate wages, we cannot comment with certainty on the effect of healthcare on labor productivity. However, to the degree that farm labor (or more specifically, hours spent in the field) require more physical effort than nonfarm labor (or even hours spent processing agricultural and livestock products), these results provide suggestive evidence of the effects of acute illness and subsequent healthcare choice on productivity in particularly high effort activities.

V. Falsification, Checks, and Robustness

In Tables V–X, we present a series of falsification checks, instrument checks, and evidence of the robustness of the major results discussed above.

A. School Hours and Labor Supply of Nonsick Children

Table 5 reports results from reduced form regressions of school and labor outcomes on the interaction instrument and the usual set of controls using the sample of nonsick school-aged children from households with no acutely ill members. If the interaction instrument is truly excludable from second-stage specifications (that is, if the instrument does not affect outcomes except through its affect on the use of formal-sector care), then we should expect no effect of the interaction instrument on school and labor outcomes among nonsick children from nonsick households.

Notice we have excluded nonsick children from households with sick members from the sample. To the extent that intrahousehold labor, consumption, or leisure allocations are affected by the acute illnesses of all members of the households, we might expect healthcare choices of sick members to affect school and labor outcomes of nonsick members irrespective of the excludability of the interaction instrument. Therefore, we only run these regressions on the sample of nonsick children from entirely nonsick households.

Across all of the schooling outcomes and farm and nonfarm labor, we see no effect of the interaction instrument. These results provide additional evidence of the validity of the interaction of facility existence and days of rainfall in the month of illness as an excludable instrument for healthcare choice. Notice, however, that both facility existence and days of rainfall along are strongly predictive school and labor outcomes, further emphasizing their inappropriateness as candidate instruments for formal-sector care use.

B. Longer-term Schooling Outcomes

The validity of the interaction of facility existence and days of rainfall in the month of illness as an excludable instrument rests upon the assumption that this interaction term only affects outcomes through its affects on healthcare choices. If the interaction instrument only affects schooling and labor outcomes through its affect on healthcare choices, it ought not to predict schooling choices and outcomes made before the acute illness episode.

Also, one necessary condition for the excludability of the instrument is that rainfall in the month of survey must be independent of rainfall in previous months and variations must be unexpected. In particular, our controls for polynomials in historical means and standard deviations of rainfall and days of rainfall in the month prior to survey should sufficiently purge the interaction term of its correlation with rainfall and outcomes in months other than the month of survey.

Table 5

School Hours and Labor Supply of Nonsick Children (Falsification)

	School Hours Last Week	Any School Last Week	Days Last Week	Hours per Day Last Week	Farm Labor	Nonfarm Labor
Days of rainfall × no facility	0.356	−0.0153	0.0330	0.0789	−0.404	0.232
	(0.453)	(0.0118)	(0.0717)	(0.0488)	(0.331)	(0.245)
No facility	6.293	0.304**	0.603	0.0111	14.18***	−0.383
	(5.297)	(0.135)	(0.821)	(0.521)	(3.659)	(3.548)
Days of rainfall	−7.536**	−0.203***	−1.006**	−0.448*	7.207***	−0.404
	(2.995)	(0.0762)	(0.455)	(0.267)	(1.822)	(1.337)
Observations	657	657	657	547	1,240	1,240
Mean of dependent variable	25.68	0.840	3.839	6.695	9.373	13.15

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). See Table II for additional comments. Samples are restricted to all individuals aged 7–19 (inclusive) who did not report being ill with an illness beginning two weeks prior to survey and live in households without any other sick children. The sample on which the regression reported in Columns 1–4 are run is further restricted to individuals aged 7–19 (inclusive) who also reported being currently enrolled in school.

Table 6
Effects of Healthcare Choice on Past Schooling Decisions (Falsification)

	IV		OLS	
	Currently Enrolled	Attended School Past Six Months	Currently Enrolled	Attended School Past Six Months
Formal healthcare	0.146	0.0321		
	(0.269)	(0.285)		
Days of rainfall × no facility			−0.00281	−0.000618
			(0.00530)	(0.00554)
No facility	0.00337	−0.0250	−0.0291	−0.0322
	(0.100)	(0.107)	(0.0564)	(0.0609)
Days of rainfall	−0.0101	−0.00955	−0.0106	−0.00966
	(0.0238)	(0.0246)	(0.0228)	(0.0246)
Observations	1,954	1,954	1,954	1,954
Mean of dependent variable	0.626	0.821	0.626	0.821

Notes: Robust standard errors in parentheses (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$). See Table II for additional comments.

In Columns 1 and 2 of Table 6, we report second-stage IV results of the effects of formal healthcare on the probability of being currently enrolled and attending school at all in six months prior to survey, respectively. Columns 3 and 4 report results from analogous reduced form regressions of these outcomes on the interaction instrument. We see no evidence of an effect of formal healthcare, as instrumented by the interaction of facility existence and days of rainfall, nor of an effect of the instrument directly on past schooling outcomes.

Columns 3 and 4 are also effectively selection equations, whose coefficient estimates can be used to test for whether selection into various definitions of enrollment varies with the instrument. The results indicate that the instrument does not significantly shift the selection probability for either sample.

C. Selection into Sickness and Symptoms

If the instrument affects the probability of reporting an acute illness or the severity and symptoms of illnesses reported, it might have an effect on outcomes outside of its effect on healthcare choice. In particular, due to the self-reported nature of the data on illness, we might be worried that heterogeneity in the threshold level of illness severity required for an individual to report acute illness might be correlated with unobserved characteristics and, in turn, with health, schooling and labor outcomes.

This could potentially be a source of bias in our estimates; however, we would only be concerned to the degree that the instrument covaries with this threshold. In

Table 7
Selection into Sickness and Symptoms

	Acute Illness Beginning < = 14 Before Survey	Fever Symptoms	Cough Symptoms	Headache Symptoms
Days of rainfall × no facility	0.00167	−0.0112	−0.00176	−0.00886
	(0.00421)	(0.00711)	(0.00702)	(0.00621)
No facility	−0.0713	0.162	0.0742	−0.104
	(0.0565)	(0.109)	(0.0891)	(0.0749)
Days of rainfall	0.0125	−0.0273	0.0475	0.0215
	(0.0226)	(0.0454)	(0.0370)	(0.0286)
Observations	7,094	1,954	1,954	1,954
Mean of dependent variable	0.276	0.367	0.280	0.266

Notes: Robust standard errors in parentheses (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$). See Table II for additional comments. Sample in Column 1 not restricted by acute illness; analysis conducted on all children aged 7–19 (inclusive).

order to check for this selection into reporting acute illness, we first regress the incidence of sickness (using a binary variable for reporting an illness which began in the past 14 days) on the full set of regressors including the interaction instrument in a linear probability OLS regression. The results from this regression are reported in Column 1 of Table 7, and verify that the instrument does not predict selection into the sample.

In Columns 2–4, we report results from the regressions of binaries for the reporting of fever symptoms, cough, and headaches, respectively, on the instrument and the usual set of controls. Again, we find no significant effects of the instrument on the probability of reporting specific symptoms. These results provide further evidence that the instrument is orthogonal to illness severity and heterogeneous severity thresholds for reporting illness.

Finally, in Table 8 we report first- and second-stage regression results from specifications identical to those reported in Tables II and III, but with the symptom binaries as additional controls. The results appear quite robust to the inclusion of these additional controls. That is, there does not seem to be any evidence that selection into sickness or symptoms is driving the results.

D. Instrument Checks

The main reason we use an interaction instrument is that it improves on using facility existence alone as an instrument for healthcare choice, since, as mentioned earlier, endogenous allocation of health facilities to communities on the basis of the community's health stock would render invalid facility existence as an instrument. Here, we present evidence that this important distinction is relevant in our context.

Table 8
Robustness to Inclusion of Symptom Controls

	First Stage		Second-stage IV						
	Whole Sample	Enrolled Sample	Health	Schooling				Labor	
	Formal Healthcare		Still Ill	School Hours Last Week	Any School Last Week	Days Last Week	Hours per Day Last Week	Farm Labor	Nonfarm Labor
Formal healthcare			−0.860*	33.76*	0.841*	4.520*	1.344	13.63	3.089
			(0.483)	(18.10)	(0.506)	(2.569)	(2.001)	(13.06)	(9.623)
Days of rainfall × no facility	−0.0173***	−0.0224***							
	(0.00600)	(0.00816)							
Fever symptoms	0.142***	0.151***	0.0451	−5.913*	−0.139	−0.826*	−0.107	−2.840	−0.807
	(0.0229)	(0.0268)	(0.0739)	(3.048)	(0.0861)	(0.429)	(0.319)	(1.925)	(1.350)
Cough symptoms	−0.0613***	−0.0633**	0.0294	1.335	0.0299	0.185	0.0698	1.211	1.346
	(0.0223)	(0.0275)	(0.0403)	(1.577)	(0.0410)	(0.216)	(0.138)	(0.880)	(0.820)
Headache symptoms	0.0474*	0.0651**	−0.0551	−2.748	−0.0613	−0.400*	−0.0369	−0.750	−0.467
	(0.0243)	(0.0297)	(0.0401)	(1.693)	(0.0449)	(0.238)	(0.155)	(0.908)	(0.786)
F-test: rain × distance = 0	8.359	7.549							
Prob > F	0.00420	0.00651							
Observations	1,954	1,224	1,953	1,224	1,224	1,224	1,010	1,954	1,954
Mean of dependent variable	0.226	0.242	0.367	23.51	0.825	3.559	6.604	8.748	12.68

Notes: Robust standard errors in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). For Columns 1, 3, 6–7, refer to Table II for additional comments. For Columns 2, 4–5, refer to Table 3 for additional comments. Samples are restricted to all individuals aged 7–19 (inclusive) who reported being ill in the two weeks prior to survey. The sample on which the regression reported in Columns 2, 4–5 are run is further restricted to individuals aged 7–19 (inclusive) who also reported being enrolled in school.

First, we regress indicators for various chronic illnesses on the facility existence variable alone (along with the full set of controls used across all specifications). The results, reported in Columns 1–4 of Table 9, show that we fail to accept that facility existence is not correlated with measures of chronic illness in the nonacutely ill population. In particular, in Columns 2 and 3 we see that facility existence is a positive predictor of chronic rash at the 10 percent level of significance and a negative predictor of chronic fever at the 5 percent level.

Second, we include the interaction term (along with the main effects and the full set of controls), and verify that the interaction instrument, in contrast, is not significantly correlated with chronic illness measures. These results are reported in Columns 5–8 of Table 9. The results indicate that the instrument is not, in fact, a predictor of measures of chronic illness at conventional levels of statistical significance.

E. First-Stage Robustness

Table 10 shows the robustness of first-stage results to different sets of controls. Columns 1 and 4 present results from the preferred specifications of first-stage regressions, identical to the results shown in the first columns of Tables 1 and 2. In Columns 2 and 5 of Table 10, we report first-stage regression results from specifications identical to those corresponding to the results reported in Columns 1 and 4, respectively, but with a different sets of controls. The specifications reported in Columns 2 and 5 exclude dummies for presence of resources in the community and interactions of these dummies with days of rainfall in the month of sickness.

The instrument has a significant, negative effect on the probability of choosing formal-sector care in both specifications shown in Columns 2 and 5, though the F-statistic is roughly half of what it is in the preferred specification. These results show that once we isolate variation in the instrument to variation in the "cost" of access to healthcare by purging it of variation in the costs of access to other resources, the instrument is more strongly predictive of healthcare choice. Nevertheless, the point estimates are of similar sign and magnitude across both sets of controls, suggesting a general robustness of the first-stage relationship between the instrument and formal-sector care use.

In Columns 3 and 6 of Table 10, we present estimates of only the main effects of days of rainfall and lack of a health facility in the community on the binary for whether the sick individual chose formal-sector care. In this specification, the interactions between the dummies for the presence of resources in the community (including a health facility) and days of rainfall in month of sickness are not included. As is expected, the main effect of the lack of a formal-sector care facility in the community is negative on the probability of choosing formal-sector care.

Dummies for deciles in the days of rainfall and amount of rainfall in the month of sickness, however, are included to sufficiently control for nonlinear effects of rainfall on healthcare choice which might correlate with unobserved measures of remoteness. The inclusion of these terms renders the interpretation of the estimates of the effects of the linear term in days of rainfall difficult. As we see, though we would expect the effects of rainfall to be significant and negative, we find an insig-

Table 9
Instrument Checks

	"No Facility" Invalid as Instrument				Exogeneity of Instrument			
	Chronic Illness	Chronic Weight Loss	Chronic Rash	Chronic Fever	Chronic Illness	Chronic Weight Loss	Chronic Rash	Chronic Fever
Days of rainfall × no facility					0.00676	4.35e-05	0.00364	0.00195
					(0.00652)	(0.00520)	(0.00474)	(0.00373)
No facility	0.0134	0.0118	−0.0664*	0.0473**	−0.105	−0.0168	−0.0650	0.0403
	(0.0565)	(0.0446)	(0.0355)	(0.0231)	(0.0918)	(0.0657)	(0.0647)	(0.0390)
Days of rainfall	−0.0264	0.0251	0.0217	−0.0200*	−0.0481	−0.000829	−0.0186	−0.00357
	(0.0195)	(0.0197)	(0.0190)	(0.0106)	(0.0362)	(0.0414)	(0.0362)	(0.0186)
F-test: instrument 0	0.0559	0.0706	3.497	4.199	1.073	7.01e-05	0.592	0.273
Prob > *F*	0.813	0.791	0.0630	0.0418	0.302	0.993	0.443	0.602
Observations	1,117	1,122	1,122	1,122	1,117	1,122	1,122	1,122
Mean of dependent variable	0.0862	0.0484	0.0471	0.0270	0.0862	0.0484	0.0471	0.0270

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). See Table II for additional comments. All specifications reported in this table exclude dummies for deciles of how long ago the illness started as well as the linear term in days. The specifications reported in Columns 1–4 also exclude all interactions between rainfall variables and dummies for resources in the community. "Instrument" refers to the No Facility dummy in Columns 1–4, and to the interaction of this dummy and days of rainfall in the month of survey in Columns 5–9. The samples on which the regressions reported in this table are run is restricted to all individuals aged 7–19 (inclusive) years who reported *not* being ill with an illness that started in the two weeks prior to survey.

Table 10
First-Stage Robustness and Main Effects

	Whole Sample			Currently Enrolled in School		
	First Stage Robustness		Main Effects	First Stage Robustness		Main Effects
	All Controls	No Resource Controls, Nor Interactions	No Rain Interactions	All Controls	No Resource Controls, Nor Interactions	No Rain Interactions
Days of rainfall × no facility	−0.0193***	−0.0134**		−0.0266***	−0.0151**	
	(0.00616)	(0.00598)		(0.00832)	(0.00717)	
No facility	−0.222**	−0.161**	−0.316***	−0.206	−0.245**	−0.365***
	(0.0952)	(0.0777)	(0.0748)	(0.143)	(0.110)	(0.102)
Days of rainfall	−0.00346	0.0303	0.0154	−0.0272	0.0252	0.00785
	(0.0307)	(0.0217)	(0.0220)	(0.0359)	(0.0268)	(0.0287)
F-test: rain × distance = 0	9.763	4.997		10.20	4.445	
Prob > *F*	0.00201	0.0263		0.00161	0.0362	
Observations	1,954	1,954	1,954	1,224	1,224	1,224
Mean of dependent variable	0.226	0.226	0.226	0.242	0.242	0.242

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). See Tables II and III for additional comments. The specification reported in Columns 2 and 5 exclude resource dummies and the interactions of resource dummies with rainfall variables and rainfall in the month prior to survey. The specification reported in Column 3 and 6 also exclude the interaction instrument.

nificant positive effect on the probability of choosing formal-sector care. This could be entirely due to the presence of the decile dummies in rainfall in the specification.

F. Robustness to Noncontemporaneous Rain

Finally, in Table 11, we report primary first- and second-stage results from specifications with added controls for noncontemporaneous days of rainfall and interactions with the "No Facility" dummy. Specifically, we accumulate rainfall in the six months after the month of illness and the six months prior to month of illness and include both these variables and their interactions with the dummy for facility existence in the individual's community. We see in Table 11 that the results are quite robust to the inclusion of these additional controls. This provides additional evidence that the interaction instrument is, indeed, only picking up transitory and unexpected variation in the cost of accessing formal-sector care in the month of illness.

VI. Conclusion

We show that acute health shocks and corresponding healthcare investments affect children's time allocation decisions. Accurate measurement of the benefits of improved access to better healthcare should not only include first-order effects on health outcomes, but also, where appropriate, these second-order effects on child labor and schooling outcomes. The empirical results presented here first verify that better healthcare for children does in fact lead to speedier recovery from acute illness. Second, we show that better healthcare generates significant improvements in school attendance. Third, we find no significant effects on child labor supply, though the results suggest an increase in labor hours in farming and herding as a result of accessing better healthcare.

We expect that our conclusions are relevant to other developing countries given the similar characteristics of labor markets (especially as related to the prevalence of child labor) and healthcare systems between our setting and others, particularly those in sub-Saharan Africa. Nevertheless, more research is needed to understand how shocks to health other than acute illnesses (for example, injuries or chronic conditions) and corresponding healthcare investments affect child time allocation, and which types of health infrastructure or access improvements result in the greatest improvements in children's outcomes.

Furthermore, in developing country settings in which extended households make resource allocation decisions jointly, especially when the household serves as both a consumptive and productive unit (as it does in rural agricultural households), the intrahousehold allocation of labor (and reallocation of labor in response to acute health shocks) will play a large role in the time allocation decisions of both sick and nonsick household members, adults and children alike. Exploring the effects of acute illness and corresponding investments in quality healthcare in the context of a household resource allocation problem is thus an important area of further research.[12]

12. In a working paper (Adhvaryu and Nyshadham 2011b), we make some progress in exactly this direction of inquiry.

Table 11
Robustness to Inclusion of Noncontemporaneous Rain Controls

	First Stage		Second-stage IV						
	Whole Sample	Enrolled Sample	Health	Schooling				Labor	
	Formal Healthcare	Still Ill	School Hours Last Week	Any School Last Week	Days Last Week	Hours per Day Last Week	Farm Labor	Nonfarm Labor	
Formal healthcare			−0.665 (0.535)	37.66** (17.91)	1.013** (0.513)	4.848* (2.524)	1.304 (1.786)	23.84 (14.86)	16.38 (12.28)
Days of rainfall × no facility	−0.0150*** (0.00569)	−0.0243*** (0.00785)							
Cumulative days of rainfall six months after × no facility	−0.00430*** (0.00155)	−0.000375 (0.00186)	0.000722 (0.00366)	0.0111 (0.114)	−0.00104 (0.00334)	−0.00427 (0.0162)	0.00554 (0.00625)	0.0524 (0.0800)	0.0322 (0.0792)
Cumulative days of rainfall six months before × no facility	0.000722 (0.00152)	0.00158 (0.00178)	0.00347* (0.00189)	−0.0538 (0.110)	−0.00284 (0.00311)	−0.00728 (0.0153)	0.00424 (0.00786)	0.114** (0.0550)	0.0177 (0.0451)
Cumulative days of rainfall six months after	−6.80e-05 (0.00167)	−0.00244 (0.00222)	−0.00300 (0.00238)	0.120 (0.123)	0.00639* (0.00355)	0.0225 (0.0174)	−0.00462 (0.0106)	0.0193 (0.0660)	0.0629 (0.0552)

(continued)

Table 11 *(continued)*

	First Stage		Second-stage IV						
	Whole Sample	Enrolled Sample	Health	Schooling				Labor	
	Formal Healthcare		Still Ill	School Hours Last Week	Any School Last Week	Days Last Week	Hours per Day Last Week	Farm Labor	Nonfarm Labor
Cumulative days of rainfall six months after before	−0.00227	−0.00564***	−0.00560**	0.0972	0.00553	0.0146	0.000447	−0.00582	0.0187
	(0.00169)	(0.00215)	(0.00285)	(0.136)	(0.00367)	(0.0188)	(0.0143)	(0.0792)	(0.0662)
F-test: rain × distance = 0	6.989	9.599							
Prob > *F*	0.00880	0.00222							
Observations	1,768	1,100	1,767	1,100	1,100	1,100	919	1,768	1,768
Mean of dependent Variable	0.226	0.242	0.367	23.51	0.825	3.559	6.604	8.748	12.68

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). For Columns 1, 3, 6–7, refer to Table 2 for additional comments. For Columns 2, 4–5, refer to Table 3 for additional comments. Samples are restricted to all individuals aged 7–19 (inclusive) who reported being ill in the two weeks prior to survey. The sample on which the regression reported in Columns 2, 4–5 are run is further restricted to individuals aged 7–19 (inclusive) who also reported being enrolled in school.

References

Adhvaryu, Achyuta, and Anant Nyshadham. 2011a. “Healthcare Choices, Information, and Health Outcomes.” *Economic Growth Center Discussion Paper No. 994*. URL available: http://www.econ.yale.edu/growth_pdf/cdp994.pdf

———. 2011b. “Labor Complementarities and Health in the Agricultural Household.” *Economic Growth Center Discussion Paper No. 996*. URL available: http://www.econ.yale.edu/growth_pdf/cdp996.pdf

Alderman, Harold, Jere Behrman, Victor Lavy, and Rekha Menon. 2001. “Child Health and School Enrollment: A Longitudinal Analysis.” *Journal of Human Resources* 36(1):185–205.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti. 2009. “Why Should We Care About Child Labor? The Education, Labor Market, and Health Consequences of Child Labor.” *Journal of Human Resources* 44(4):871–89.

Cutler, David, Robert Huckman, and Mary Beth Landrum. 2004. “The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery.” *American Economic Review (Papers and Proceedings)* 94(2):342–6.

Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. “The Quality of Medical Advice in Low-Income Countries.” *Journal of Economic Perspectives* 22(2):93–114.

Dow, William, Paul Gertler, Robert F. Schoeni, John Strauss, and Duncan Thomas. 1997. “Health Care Prices, Health and Labor Outcomes: Experimental Evidence.” *RAND Labor and Population Program working paper* series 97-01, no. DRU-1588-NIA.

Edmonds, Eric. 2008. “Child Labor.” *Handbook of Development Economics* 4:3607–709.

Food and Agriculture Organization, Food and Nutrition Technical Report Series. 2001. “Human Energy Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation.” Rome, 17–24 October. ISBN: 92-5-105212-3. URL available: ftp://ftp.fao.org/docrep/fao/007/y5686e/y5686e00.pdf

Gertler, Paul, Luis Locay, and Warren Sanderson. 1987. “Are User Fees Regressive? The Welfare Implications of Health Care Financing Proposals in Peru.” *Journal of Econometrics* 36:6788.

Glewwe, Paul, and Hanan Jacoby. 1995. “An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrtition.” *Review of Economics and Statistics* 77(1):156–69.

Gowrisankaran, Gautam, and Robert Town. 1999. “Estimating the Quality of Care in Hospitals Using Instrumental Variables.” *Journal of Health Economics* 18:747–67.

Kruger, Diana. 2007. “Coffee Production Effects on Child Labor and Schooling in Rural Brazil.” *Journal of Development Economics*82(2):448–63.

Martorell, Reynaldo, Jean-Pierre Habicht, and Juan A. Rivera. 1995. “History and Design of the INCAP Longitudinal Study (1969–77) and Its Follow-up (1988–89).” *Journal of Nutrition* 125(4):1027S-41S.

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

Mwabu, Germano, James Mwanzia, and Wilson Liambila. 1995. “User Charges in Government Health Facilities in Kenya: Effect on Attendance and Revenue.” *Health Policy and Planning* 10(2):164–70.

Mwabu, Germano. 2009. “The Production of Child Health in Kenya: A Structural Model of Birth Weight.” *Journal of African Economies* 18(2):212–60.

Ravallion, Martin, and Quentin Wodon. 2000. “Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy.” *Economic Journal* 110(462): C158-C175.

Thirumurthy, Harsha, Joshua Graff Zivin, and Markus Goldstein. 2009. "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya." *Journal of Human Resources* 43(3):511–52.
Thomas, Duncan, Elizabeth Frankenburg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Jaswadi, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, and Siswanto Wilopo. 2006. "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence." *California Center for Population Research, Working Paper,* CCPR-070-06.

Appendix 1

Construction of Variables

The following list describes the construction of variables used in analysis:

- *sick* = 1 if the individual was sick with an illness that began 14 days or fewer prior to the date of survey, *sick* = 0 otherwise.
- *h* = 1 if sick individual visited hospital, health center or dispensary (government, NGO or private); *h* = 0 otherwise
- *raindays* equals the number of days of rainfall at the rainfall station closest to the individual's sample cluster, in the month and year that the individual was surveyed
- *histmean* of rainfall is the number of days of rainfall in the month of survey averaged over all years in which rainfall data are recorded for that cluster in the particular month
- *histsd* is calculated as the standard deviation of the historical distribution of days of rainfall in the month of survey, across all years in which rainfall data are recorded for that cluster in the particular month
- *histmeansq* and *histsdsq* are smooth polynomials to the second degree in historical mean days of rainfall and historical standard deviation of days of rainfall, respectively
- *raindayslast* equals the number of days of rainfall at the rainfall station closest to the individual's sample cluster, in the month *before* that in which the individual was surveyed of the same year
- *decraindays* and *decrainfall* are categorical variables reporting which decile of the rain days and rainfall distributions, respectively, the rain in the survey month falls; fixed effects for each decile are included in all specifications
- *noexist* is a binary variable which takes value *noexist* = 1 if neither hospital, health center, nor dispensary of (government, NGO or private) exists in the community, and *noexist* = 0 otherwise (Note: for waves in which these data were missing, the values were filled first using the minimum from the waves

in which the data were not missing for that cluster, and second using the minimum of nonmissing values from clusters matched to the same rain station in the same wave; that is, if a facility of these types ever existed in that cluster or in very proximate clusters before or after the year in which the data are missing, we assumed it existed during this wave as well)

- For the following facilities/attributes (x), we calculate distances as $dist(x) = 0$ if the facility/attribute exists in the individual's village; $dist(x)$ equals the distance to the nearest such facility/attribute outside the individual's village if one does not exist in the village (Note: for waves in which these data were missing, the values were filled first using the mean from the waves in which the data were not missing, and second using nonmissing data from clusters matched to the same rain station in the same wave)
  - Hospital
  - Health center
  - Dispensary
  - Daily market
  - Periodic market
  - Driveable road
  - Public transportation
  - Secondary school
  - Bank
  - Post office/telephone booth
- Categorical variables for the quintiles of the distributions of the above defined distances to hospital, health center, and dispensary were created and included in all specifications
- $dechhsize$ is a categorical variable for deciles of the distribution the number of members of the household
- $age1$, $age2$, and $age3$ are smooth polynomials up to the third degree in the age of the respondent
- $decage$ is a categorical variable for deciles of the age distribution
- $assets$ is a categorical variable measuring 25 quantiles of the distribution of the value of all assets of the household; fixed effects for these categorical values are included in all specifications
- $kid = 1$ if $7 < = age < = 19$
- $educ$ is a categorical variable for the individual's completed level education; fixed effects for these categorical values are included in all specifications along with interactions with a dummy for female

- *rainyseason* = 1 if the individual was surveyed during one of the two rainy seasons (March-May and October-December)
- *illstart* is a continuous variable measuring the number of days prior to survey the illness began
- *decstart* is a categorical variable for deciles of *illstart*
- *station* is a unique identifier for the rain station closest to the cluster in which the individual's household is located