



Published in final edited form as:

Am Econ J Econ Policy. 2015 August ; 7(3): 29–57. doi:10.1257/pol.20120262.

Returns to Treatment in the Formal Health Care Sector: Evidence from Tanzania*

Achyuta Adhvaryu[†] and Anant Nyshadham[‡]

Achyuta Adhvaryu: adhvayu@umich.edu; Anant Nyshadham: nyshadha@usc.edu

[†]University of Michigan

[‡]University of Southern California

Abstract

Improving access to the formal health care sector is a primary public health goal in many low-income countries. But the returns to this access are unclear, given that the quality of care at public health facilities is often considered inadequate. We exploit temporal and geographic variation in the cost of traveling to formal sector health facilities to show that treatment at these facilities improves short-term health outcomes for acutely ill children in Tanzania. Our results suggest that these improvements are driven in part by more timely receipt of and better adherence to antimalarial treatment.

Keywords

health care; adherence; child health; Africa

Expanding access to the formal health care sector is a primary goal for public health policy in many low-income countries. Yet the assumption underlying this priority—that access to the formal health care system improves the health (and ultimately the welfare) of marginalized populations—remains untested. Moreover, given that the quality of care at public health facilities in this setting is often considered inadequate (Jishnu Das, Jeffrey Hammer, and Kenneth Leonard 2008), it is unclear whether policies that enable or encourage access to formal sector health care could generate meaningful health improvements without additional public expenditure to increase quality (e.g., infrastructure development, supply chain improvements, and investment in human capital).

The main difficulty in estimating the returns to treatment in the formal health care sector is that individuals select into health care options (Maureen Cropper 1977; Thomas Selden 1993; Fwu-Ranq Chang 1996; Michael Grossman 2000). The determinants of treatment choice—such as preferences, information, and the severity of illness—likely also affect health outcomes. Since these factors are not all observed, they will, in general, bias the estimate of

*We thank Abhijit Banerjee, Prashant Bharadwaj, Michael Boozer, Betsy Bradley, Joe Doyle, James Fenske, Jason Fletcher, T. Paul Schultz, Tavneet Suri, Chris Udry, and participants at the Yale Development Lunch, PACDEV, and the NBER Summer Institute for helpful comments and suggestions. Thanks also to Mark Duggan and two anonymous referees for greatly improving the paper through the review process. Adhvaryu gratefully acknowledges funding from the NIH/NICHD (5K01HD071949). This paper formerly circulated under the title “Healthcare Choices, Information and Health Outcomes.” All errors are our own.

the impact of formal sector treatment. Several recent studies in the United States have exploited natural experiments to estimate the returns to various categories of health care net of these biases. This literature has examined the effects of hospital quality (John Geweke et al. 2003; Thomas Buchmueller et al. 2006), physician quality (Joseph Doyle, Steven Ewer, and Todd Wagner 2010), emergency care (Doyle 2012), and postnatal and postpartum care (Douglas Almond et al. 2010; Almond and Doyle 2011).

In this study, we examine the effects of treatment following acute illness at formal sector health facilities on short-term health outcomes for young children using nationally representative data from Tanzania. Our empirical strategy exploits temporal and geographic variation in the costs of traveling to formal sector health facilities. Since health insurance is not present and the government heavily subsidizes the pecuniary costs of care, the most salient cost is related to traveling to the care option (Paul Gertler, Luis Locay, and Warren Sanderson 1987; Germano Mwabu, James Mwanza, and Wilson Liambila 1995). This cost can be economically substantial, especially in remote areas.¹

We use variation in this cost generated by the interaction of distance (to the nearest health facility) and rainfall to predict the choice of health care following acute illness. The intuition behind our interaction instrument is simple: rainfall generates random variation in the cost (or disutility) of traveling a given distance. Heavier rain should discourage individuals who live farther away *more* than individuals living closer to the nearest health facility. To account for the many direct effects of rainfall and remoteness on both the choice of care and health outcomes, we control for the main effects of rainfall and distance in both stages of the two-stage least squares estimation.

We find that the instrument is strongly predictive in the first stage: consistent with our prediction, the negative effect of distance on formal health care usage is exacerbated in rainy months. This effect is robust to a variety of additional controls and passes various falsification tests.² Using a sample of young children who were sick with fever in the two weeks preceding survey, we focus on two main health outcome variables: the incidence of fever and malaria on the day of survey. We find that the instrumental variable (IV) estimates are several times as large as the ordinary least squares (OLS) estimates, consistent with self-selection based on the severity of illness. Our overall finding is that formal sector health care greatly reduces the incidence of fever and malaria among children who sought treatment for acute illness.

Given the large magnitudes of the estimated effects, next we ask *why* formal sector health care is more effective. Our results suggest that these improvements are driven in part by

¹In the rural part of our sample from Tanzania, where health facilities are relatively densely located compared to the rest of East Africa, the average distance to the nearest health facility is 4.67 kilometers, which is most often traveled by foot.

²Additional controls include the historical mean and standard deviation of rainfall in a given locality; distance to the nearest market (which is a measure of remoteness of the household) and its interaction with rainfall; and the geographic (region fixed effects), demographic and socioeconomic characteristics of sample households. We provide falsification tests using rainfall in past and future months, and find no effects of their interactions with distance on health care choice. Additionally, we show that the instrument does not predict selection into the sick sample (reported fever in 2 weeks prior to survey) nor does it predict sickness (positive test for malaria) at the time of survey among the reportedly non-sick sample. These checks alleviate concerns that the instrument is picking up on unobservable, systematic differences between more or less remote locales that experienced more or less rainfall in the month of survey.

more timely receipt of and better adherence to antimalarial treatment. In particular, we find that children using formal sector health care begin antimalarial treatment with less delay and are more likely to adhere to their antimalarial therapy regimens. Receipt of medications, both antimalarial and non-antimalarial, show only weak differences across formal vis-a-vis informal care. Taken together, these results suggest, at least in the Tanzanian context, that the outcome gradient across the formal and informal health care sectors is driven by differences in the receipt and appropriate usage of antimalarial treatment.

Our study makes two main contributions. First, we provide the first assessment, to our knowledge, of these returns in a developing country setting, using methods which account for the bias induced by self-selection into health care options. Despite a large number of studies in developing country contexts on the effects of health interventions—e.g. nutritional supplements, preventive technology, and treatments³—on outcomes, little is known about the causal effects of choosing formal sector health care following acute illness.⁴ We add to this knowledge base by providing an estimate of these returns for young children.

Second, though disparities in quality of care (formal vis-a-vis informal health care options) have been well documented, the particular *mechanisms* through which improvements in health may be generated, in a causal sense, have not to our knowledge been explored. A long line of work in public health has highlighted the so-called “Last Mile Problem” in global health: even when effective health technologies exist, enabling access and acceptance in the population—going the last mile—is often difficult but has large returns.⁵ Our results suggest that access to antimalarial treatments is only weakly better in the formal health care sector but that patients receive more timely treatment and adhere more when accessing formal sector health care, and that this subtle difference may generate large health returns, at least for the young children in our context.

The remainder of the paper is structured as follows. Section 1 describes a model of health care choice and outcomes to motivate the empirical analysis. Section 2 describes the data we use and the construction of important variables. Section 3 explains the empirical strategy. Section 4 presents the results, and section 5 concludes.

1 Model

In this section, we develop a simple model which relates healthcare choice to health outcomes, to better understand why comparing the health of individuals who do and do not choose formal-sector healthcare produces a likely biased estimate of the effect of healthcare choice on outcomes. The model emphasizes the role of severity of illness, which simultaneously influences healthcare choice and outcomes and is unobserved to the researcher.

³See, for example, John Strauss and Duncan Thomas (1998), Edward Miguel and Michael Kremer (2004), Gustavo Bobonis et al. (2006), Thomas et al. (2006), and Harsha Thirumurthy et al. (2009).

⁴There are, of course, good studies documenting the variation in quality of care (Leonard and Masatu 2007) and the fact that patients bypass poor quality health facilities to reach higher quality ones despite the increase in distance (Klemick et al. 2009)

⁵See, e.g., Hans Hogerzeil (2004), WHO (2004), and Minmin Zhu et al. (2008).

1.1 Setup

We consider a utility-maximizing agent who falls sick at random and must make a healthcare choice. His realized health outcome is determined by the inherent severity of his illness s , and by his choice of healthcare $h \in \{0, 1\}$. We will think of $h = 1$ as the choice of formal-sector healthcare, and $h = 0$ as care outside the formal-sector (or no care at all). Severity, which is observed by the individual, is randomly drawn from a distribution $F(s)$.

There are two health outcomes (represented as the random variable D , number of days ill) which may ensue: $D = D^g$ (good), or $D = D^b$ (bad), where $D^g < D^b$.⁶ Healthcare choice and severity combine to determine the probability that the good health outcome occurs. Denote $\alpha(s, h)$ as the function which maps severity and healthcare choice into a probability: $\alpha(s, h) = \Pr(D = D^g | s, h) \in [0, 1]$. By definition of severity s , $\alpha(s, h)$ is decreasing in s for all h . That is, for both choices of healthcare, a higher severity level decreases the probability of realizing the good health outcome.

We impose the following assumption: $\alpha(s, 1) - \alpha(s, 0)$ is increasing in s . Formal-sector care is more comparatively effective at higher levels of severity: the differential benefit of visiting the health facility will be lower for low-severity illnesses than for high-severity illnesses. The intuition here is that self-treatment for the common cold, for example, is not likely to be very different from treatment in the formal-sector; however, treatment for a more severe illness like pneumonia will likely be very different at a facility as compared to treatment in one's own home.

There are two main caveats to making this assumption. First, for extremely severe illnesses, the difference between formal-sector care and informal care is likely to matter little (i.e. both $\alpha(s, 1)$ and $\alpha(s, 0)$ are likely to be close to 0). If this were true, then $\alpha(s, 1) - \alpha(s, 0)$ would be a non-monotonic function of s ; in particular, we would expect the difference in the effectiveness of care to increase up to a certain point in the severity distribution and then begin to decline. For our analysis, however, we are interested less in this extreme case. We focus on individuals in the portion of the severity distribution whose healthcare choices can be shifted through exogenous movements in the relative price of care. (Indeed, it is perhaps more policy-relevant to focus on this sub-group.) It is unlikely that these individuals are at either extreme of the severity distribution. Second, Leonard (2007) suggests, in a similar formulation, that this assumption may not hold for all illnesses. Nevertheless, the empirical application presented in this paper is specific to malaria-related symptoms and illness, which are unlikely to violate this assumption.

It is clear that for there to be a non-trivial tradeoff between quality and cost of healthcare, s must lie in a region of $F(s)$ such that $\alpha(s, 1) > \alpha(s, 0)$. That is, the probability of a good health outcome is larger at each level of severity if the agent chooses formal-sector care. This establishes the comparative effectiveness of formal-sector healthcare over other forms of care. Most evidence from both developed and developing countries suggests that the price

⁶Consistent with our empirical setting, we can interpret the D^b as the number of days a parent may have to spend caring for a young child with a prolonged bout of fever or malaria, and D^g as the number of days he or she would spend if the child recovered quickly.

of formal-sector care is substantially larger than informal care, which establishes the tradeoff in price.

1.2 Utility maximization

Individuals choose h in order to maximize their expected utility subject to a budget constraint. The utility maximization problem is the following:

$$\max_{h \in \{0,1\}} E(u(C)) - P(h) \quad \text{subject to } C \leq w(\Omega - D). \quad (1)$$

Here, C is consumption; $u(C)$ is the utility function; $P(h)$ is the price of healthcare at option h ; w is the wage; and Ω is the amount of time an individual would work if fully healthy. The health outcome enters utility through its effect on the amount of time an individual is able to work (and thus the amount he can consume).

Since there are two possible states ($D = D^g$ or $D = D^b$) with known probabilities ($\theta(s, h)$), when the budget constraint binds we can write the maximization problem as

$$\max_{h \in \{0,1\}} \theta(s, h)u(w(\Omega - D^g)) + (1 - \theta(s, h))u(w(\Omega - D^b)) - P(h). \quad (2)$$

Let us define $\bar{u} = u(w(\Omega - D^g))$ and $\underline{u} = u(w(\Omega - D^b))$ as the utilities in the good and bad state, respectively. Then, the expected utility of choosing $h = 1$ and $h = 0$, respectively, are:

$$\begin{aligned} U_1 &= \theta(s, 1)\bar{u} + (1 - \theta(s, 1))\underline{u} - P(1), \\ U_0 &= \theta(s, 0)\bar{u} + (1 - \theta(s, 0))\underline{u} - P(0). \end{aligned}$$

The individual will choose $h = 1$ if and only if $U_1 - U_0 > 0$. Denoting $\Delta P = P_1 - P_0$ and $\Delta u = \bar{u} - \underline{u}$, we can express this inequality as

$$\theta(s, 1) - \theta(s, 0) > \frac{\Delta P}{\Delta u}. \quad (3)$$

Since we have assumed that the left-hand side of the above inequality is increasing in s , it follows that the function (let us denote $g(s) = \theta(s, 1) - \theta(s, 0)$) has an inverse (g^{-1}). Thus, the utility maximization problem can be expressed as a simple cutoff rule:

$$\text{Choose } h=1 \text{ iff } s > g^{-1}\left(\frac{\Delta P}{\Delta u}\right) \equiv K. \quad (4)$$

The individual will thus choose to use formal-sector healthcare if the severity of his illness is greater than a cutoff, which is in turn a function of the model's parameters. The parameters enter in intuitive ways in the cutoff value. An increase in the relative price of formal-sector care (ΔP) increases the cutoff, which in turn decreases the probability that an individual with randomly chosen severity uses formal-sector care. On the other hand, an increase in the

relative return to formal-sector healthcare (u), for example due to an increase in the wage w , lowers the cutoff and thus increases the probability of choosing $h = 1$.

1.3 Empirical implications

We are primarily interested in estimating the effect of formal-sector health care usage on health outcomes. In doing this, we are essentially estimating the returns to formal-sector care in the health production function. In this section, we investigate why comparing the average outcomes of individuals who used formal-sector care with the outcomes of those who did not is an invalid strategy for estimating this effect. We then discuss the model implications for a valid identification strategy.

First, we calculate the true average treatment effect of formal-sector healthcare on health outcomes over the entire distribution of severity. Denote $f(s)$ as the pdf of $F(s)$. The difference between the expected outcome under $h = 1$ and $h = 0$, which we denote $E(O_1 - O_0)$, is

$$\int_{-\infty}^{\infty} \left(\left(\theta(s, 1)D^g + (1 - \theta(s, 1))D^b \right) - \left(\theta(s, 0)D^g + (1 - \theta(s, 0))D^b \right) \right) f(s) ds$$

We can rewrite this quantity as $(D^g - D^b) \int_{-\infty}^{\infty} (\theta(s, 1) - \theta(s, 0)) f(s) ds$. Since, $D^g < D^b$, and $\theta(s, 1) - \theta(s, 0) > 0$ for all s , we have that $E(O_1 - O_0)$ is negative, which indicates that the true average treatment effect of formal-sector healthcare is an improvement in health outcomes. This fact arises rather trivially from the second assumption we made on p .

Second, we calculate the difference in health outcomes between individuals who chose to use formal-sector care and those who did not. We know from the model's cutoff rule that individuals below the cutoff (K) in the distribution of severity will choose $h = 0$, while those with severity levels above it will choose $h = 1$. We can thus calculate the average outcome as:

$$\int_K^{\infty} \left(\left(\theta(s, 1)D^g + (1 - \theta(s, 1))D^b \right) f(s) ds - \int_{-\infty}^K \left(\theta(s, 0)D^g + (1 - \theta(s, 0))D^b \right) f(s) ds \right) \quad (5)$$

Again, we can rewrite this quantity as $(D^g - D^b) \left(\int_K^{\infty} \theta(s, 1) f(s) ds - \int_{-\infty}^K \theta(s, 0) f(s) ds \right)$. Thus, the effect of formal-sector healthcare calculated by simply comparing health outcomes for individuals who chose $h = 1$ and $h = 0$ has the following bias (calculated by subtracting the average treatment effect from this naively measured effect):

$$\int_K^{\infty} \theta(s, 0) f(s) ds - \int_{-\infty}^K \theta(s, 1) f(s) ds. \quad (6)$$

It is clear that this bias term is not in general equal to 0. The direction of the bias depends on the shape of the severity distribution. For example, if we assume that severity s is uniformly distributed between 0 and B , such that $0 < K < B$ it is easy to see that the bias will be negative (given the assumptions made on p):

$$\int_K^B \theta(s, 0) ds < \int_0^K \theta(s, 1) ds. \quad (7)$$

The results presented in this paper are consistent with this attenuating bias due to severity. That is, the probability of a good health outcome appears to be higher under formal-sector care, and the severity bias in the OLS estimates appears to be large and negative.

Finally, we discuss how the model motivates our instrumental variables strategy, which we will develop further in Section 3. As we have shown, comparing outcomes for individuals who chose $h = 1$ and $h = 0$ generates a biased estimate of the effects of formal-sector healthcare. The ideal experiment for estimating average treatment effect without bias would be to randomize the choice of healthcare for individuals across the severity distribution, and then compare outcomes for individuals who were randomly treated with formal-sector care with those who were not.

In the absence of such an experiment, we consider the following strategy. If, on a random basis, individuals were to face different prices for formal-sector care, the price variation would generate exogenous shifts in the cutoff K , and thus create treatment and control groups based on whether individuals were randomly exposed to higher or lower prices for formal-sector care.

For example, suppose that randomly chosen individuals were exposed to a higher relative price of formal-sector care ($P' > P$). This price shift would generate variation in the cutoff ($K' > K$), and would thus exogenously drive a portion of individuals (specifically, those who have severity levels between K and K') to choose $h = 0$ instead of $h = 1$, which is what they would have chosen if the cutoff had not changed. Thus, we can compare the health outcomes of individuals who would have chosen formal-sector healthcare but were incentivized on a random basis not to by the change in price with those of the individuals who were not exposed to this random price incentive.

Note that this type of experiment elicits not an average treatment effect but rather a local average treatment effect (LATE), since the price experiment will generate variation in choice only in the region of the severity distribution between K and K' . We can express the LATE estimate as

$$(D^g - D^b) \int_K^{K'} (\theta(s, 1) - \theta(s, 0)) f(s) ds. \quad (8)$$

The LATE estimate is the average treatment effect restricted to the region of the severity distribution between K and K' . In the following section, we discuss the particular instrument used to generate variation in the price of formal-sector healthcare, using the model to better understand threats to the validity of the instrument, and creating tests to evaluate empirically the extent to which these issues are present in our data.

Figure 1 depicts the price variation we propose to use as an instrument in the context of the utility function presented above. Normalizing prices to the price of the informal option, a

relative increase in the price of formal-sector care results in a shift downwards of $u_1(s)$, the expected utility of choosing $h = 1$ as a function of s . For the sake of simplicity in the graphical representation, we have assumed that $\theta(s, h)$, and hence $u_h(s)$, is linear in s with $s \in [0, 100]$. This is, of course, not necessary for the predictions of the model shown above to hold. The solid red line corresponds to $u_1(s)$ before an increase in the relative price of formal-sector care, and the dotted red line below it represents the same utility after the price change. The blue line, denoting the $u_0(s)$ is held constant, and thus the new cutoff value, where the new $u_1(s)$ and $u_0(s)$ meet, is to the right of the old in the severity distribution. We refer to the old cutoff as K (the solid vertical black line in Figures 1 and 2) and the new cutoff as K' (the dotted vertical black line in both figures).

Figure 2 depicts $\theta(s, 1)$ and $\theta(s, 0)$ under the linear functional form assumption along with the same shift in the severity cutoff as in Figure 1. Here we can see that the average treatment effect estimate from the OLS regression compares the average of $\theta(s, 1)$ over $s \in (K, 100]$ with the average of $\theta(s, 0)$ over $s \in [0, K]$. On the other hand, the LATE estimate from the second stage IV regression compares the average of $\theta(s, 1)$ over $s \in [K, K']$ with the average of $\theta(s, 0)$ over $s \in [K, K']$.

2 Context and Data

2.1 Context

Health care markets in Tanzania are highly informal. Formal health facilities (defined here as hospitals, regional health centers, and drug dispensaries) provide care for less than half of all illness episodes. In rural areas, where the density of formal sector facilities is low, this fraction is smaller. Most patients seek care and purchase medication from (largely unregulated) pharmacies, drug shops, kiosks, traditional healers, as well as family and friends. Most interaction with the health system is following episodes of acute illness—very little preventative health care is available/utilized in most areas of Tanzania. Among those seeking curative care, the most common symptom is fever; in children, this is due to the frequent incidence of malaria, pneumonia, and respiratory tract illnesses (Klemick et al. 2009).

Travel cost (utility cost and/or opportunity cost of time) is the most significant determinant of cost of care. This makes formal health facilities effectively more expensive than self-treatment (Paul Gertler, Luis Locay, and Warren Sanderson 1987). Health care costs are financed out-of-pocket; formal insurance markets essentially do not exist. Infants and young children, pregnant women, the elderly, and severely ill patients tend to choose formal care following acute illness.⁷ Care in the formal sector is delivered by health workers with varied training and experience. Fully trained doctors are few and far between, particularly in rural areas. Diagnosis and treatment are often low quality, both due to the lack of training of workers and the frequent absence of adequate diagnostic tools and medicine stocks (Heather Klemick, Kenneth Leonard, and Melkiory Masutu 2009).

⁷See, e.g., Cohen, Dupas, and Schaner (2013).

2.2 Data

We use the 2007–2008 Tanzania HIV/AIDS and Malaria Indicator Survey (THMIS), which is part of the Demographic and Health Surveys (DHS). The THMIS used a two-stage sampling frame. In the first stage, sample points (clusters) were selected based on enumeration areas designated by the 2002 Tanzanian Population and Housing Census; 475 clusters were selected. A household census within each cluster was then used to randomly select approximately 16 households from each cluster to be surveyed. Weighting factors are included in the data, so when weighted the sample is nationally representative. In sampled households, all men and women ages 15–49 were interviewed, and blood samples for malaria and anemia were collected for children under five years old (excluding children less than 6 months old)⁸.

2.3 Child-level variables

The majority of our analysis is at the child level. Mothers were asked questions regarding the health of their children. We deal primarily with the sample of sick children, that is, children who were sick with fever or cough in the two weeks preceding the date of survey. Respondents who answered “yes” to this question were asked where they had sought care for the sick child, when they had sought care if they had, and whether the child was still sick (at the time of survey).

We construct variables corresponding to selection into the sick sample, healthcare choice, and health outcomes using answers to the survey questions mentioned above. If the respondent answered “yes” to the question about their child having a fever in the two weeks before survey, they are included in our sample of sick children. We then construct a binary variable h corresponding to the use of formal-sector healthcare for the child; $h = 1$ if the respondent brought the sick child to formal-sector healthcare, and $h = 0$ otherwise. Formal sector care is defined as a visit to a government or private hospital or health center.

We use two main health outcome variables. The first is a fever indicator variable, which equals 1 if the respondent reports that the sick child still has fever on the day of survey, and 0 otherwise. The second is an indicator variable which equals 1 if the child tests positive for malaria when surveyed, and 0 otherwise. We also construct and use binary variables for whether or not the sick child received any medicine; whether or not he received malaria medicine; days delayed before receiving malaria medicine; and whether or not he adhered to treatment regimen corresponding to these malaria medications. More detail on construction of the variables is provided in the Appendix.

2.4 Mother- and household-level variables

In subsequent analysis, we use the DHS’s survey of mothers and women of child-bearing age to understand impacts of formal-sector care usage on mothers’ health care-related information. We construct indices corresponding to the amount of health-related information the mothers have, by aggregating yes-or-no questions on the definitions of diseases, disease

⁸The Paracheck-Pf rapid diagnostic test used to detect malaria was found to be very reliable when measured against the current gold standard microscopy test for malaria in 5 districts in Tanzania (L.E.G. Mboera et al. 2006).

transmission, and treatment. For example, if there were 6 questions about transmission of various diseases, the respondent was awarded 1 point for every correct answer, deducted 1 point for every wrong answer, and given a 0 for a response of “don’t know” if applicable. The scores were then summed across the 6 questions yielding the disease transmission information score for that respondent. The same was done for a set of questions about the existence of various diseases and a set of questions about medical treatments for various diseases. A composite index equaling the sum across all of these information measures was also constructed. More detail on the construction of these indices is provided in the Appendix.⁹

We obtain data on the distance to the nearest health facility and distance to the nearest market (both in kilometers) from the household questionnaire (a module of questions for household heads).¹⁰ In addition, we include the following control variables from the child-level, mother-level and household-level questionnaires in our regressions: indicator variables for age of the child (in years); region fixed effects; wealth index category fixed effects; mother’s educational attainment in years; mother’s age at marriage; year in which mother was married; indicators for the month of survey; household size; number of living children in household; number of children under age of 5 in household; an indicator variable for urban clusters; gender of the child; and household’s altitude.

2.5 Rainfall and temperature data

We use the restricted version of the THMIS, which contains data on the latitude and longitude coordinates of each of the sampled clusters. We use these data to match clusters to rainfall and temperature data from the University of Delaware’s Center for Climatic Research. The rainfall dataset is called “Terrestrial Precipitation: 1900–2008 Gridded Monthly Time Series (1900 – 2008) (Version 2.01),” and the temperature dataset is called “Terrestrial Air Temperature: 1900–2008 Gridded Monthly Time Series (1900 – 2008) (Version 2.01).”

The rainfall and temperature measures for a latitude-longitude node (on a 0.5° latitude by 0.5° longitude grid) combine data from 20 nearby weather stations using an interpolation algorithm based on the spherical version of Shepard’s distance-weighting method. We matched the rainfall and temperature data to clusters by calculating the closest grid point to the latitude-longitude coordinates of each cluster.¹¹ The rainfall quantity we use for analysis is the one which corresponds to the month in which the individual was surveyed. Since the survey took place over a period of 4 months, each individual is matched based on the latitude-longitude coordinates of his household’s cluster, and on the month of survey.¹²

⁹We also constructed principal component indices using the health-related information questions (composite, as well as separately for diseases, disease transmission, and treatment). The results are similar to the alternate constructions discussed above, and thus are not reported in the paper, but are available upon request.

¹⁰Note as a caveat that distances are self-reported, and thus likely have a degree of error which may be correlated with household characteristics (for example, one might imagine that the extent of reporting error is correlated with the household head’s educational attainment). Nevertheless, in the absence of objective measures of distances, many studies use the self-reported measure (see, e.g., Gertler et al. 1987).

¹¹We thank Seema Jayachandran for the Stata code that performs this calculation, which is based on the Haversine formula.

In all our regressions, we control for the historical mean and standard deviation of rainfall and temperature in the month of survey at the closest grid point to the latitude-longitude coordinates of the household's cluster. We calculate these historical variables by averaging over the rainfall quantities in the last 50 years in the particular month in question. So, for example, if an individual in cluster 1 was surveyed in January, the value corresponding to the historical mean of rainfall would be the average rainfall in January in his cluster over the last 50 years. The historical standard deviation would be the standard deviation from this mean in the last 50 years.

2.6 Summary Statistics

Table 1 presents means, standard deviations, and number of observations for select variables of interest to be used in the analysis below for 1) the sample of children whose mothers reported them as having been ill in the two weeks prior to survey; 2) sick children who received care at formal-sector health facilities; and 3) sick children who did not receive formal-sector care. We also conduct *t*-tests for differences in means of these variables across children who visited formal care facilities and those who did not. We report the *t*-statistics and corresponding *p*-values in the last two columns of Table 1.

We see that 16 percent of the total sample of children under the age of 5 were reportedly ill with fever and/or cough, leaving a restricted sample of 1200 children on which to perform the proposed analysis. Roughly 27 percent of this restricted sample still reported having fever at the time of survey and roughly 23 percent tested positive for malaria. It is important to note that these means are across the entirety of the restricted sample and that the relevant means against which to compare any local average treatment effect estimates from two-stage least squares regressions are those for the population on the margin. It is possible that these means are significantly different.

Columns 3–6 of Table 1 present means and standard deviations of the variables of interest across sick children who did and did not receive formal-sector healthcare. We discuss these comparisons for the different sets of variables below.¹³

Children who did and did not receive formal-sector care appear similar in terms of probability of anemia (both subsamples have 77 percent likelihood of severe anemia); we might expect this, given that anemia is an indicator for longer-term health status. On the other hand, the proportions of children with malaria and fever (at the time of survey) are substantially different. For both short-term health indicators, children who received formal-sector care appear better off with nearly half the likelihood of both fever and malaria (32 and 35 percent likelihood of fever and malaria, respectively, for children who do not seek care vis-a-vis 17 and 21 percent likelihood among care-seeking children). These statistically

¹²The fact that the rainfall grid is created using only 20 stations likely implies that there is a great deal of measurement error in rainfall at the community level. Unfortunately, we cannot do much to correct for this problem. If measurement error is classical, we would expect attenuation of the rainfall impacts. We might imagine, though, that the most remote communities have the largest measure error, because the rainfall stations associated with the grid points nearest to them are actually quite far away. If so, we would be leaning heavily on the quality of the interpolation algorithm to account for the density of rainfall stations around grid points.

¹³We bear in mind in examining these differences in means that children in the two categories of care are, of course, likely different on many margins, not all of which can be observed (e.g. severity of illness, or the child's family's preferences for health). Thus we cannot interpret these differences causally. We develop a methodology for causal identification of effects in section 3.

significant differences are evident in the t-statistics and p-values presented in the final two columns of Table 1.

Across healthcare choices, sick children appear generally similar in terms of gender, but differ in terms of age, household size and rural residence. Those not receiving formal care are slightly older (1.75 versus 1.49 years of age), more likely to live in rural areas (.87 versus .78 percent likelihood), and come from larger households (7.84 versus 6.8 household members).

Those who received formal sector care received more medicines on average, and the difference is apparent for number of antimalarials (.70 versus .34) as well as non-antimalarials (1.06 versus .78). Adherence, on the other hand, appears to be similar across healthcare choice groups (roughly 60 percent across all medicines and subsamples).

The information indices are constructed from responses to numerous questions regarding the existence, transmission, and treatment of various diseases. The information on disease index was constructed to range from -4 to 4 and has a mean value of just over 3 . The information on transmission index was constructed to range from -13 to 13 and has a mean of just over 8 . The information on treatment index was constructed to range from -3 to 3 and has a mean of roughly 1.9 . The composite ranges from -20 to 20 . It has a mean of just over 14 . More information on the creation of these indices is provided in the Appendix.

We see that the mean composite index (measured at the mother-level) is slightly larger for those who visited formal-sector care, and that the separate indices related to diseases, transmission and treatment are all slightly larger for formal-sector care users as well. These differences are fairly small as compared to the means and statistically insignificant except for treatment-related information.¹⁴

3 Empirical Strategy

In this section, we describe our instrumental variables (IV) strategy and present evidence related to the intuition behind and validity of our instrument. Our primary aim is to obtain unbiased estimates of the effects of accessing formal-sector care on health outcomes. As detailed in the previous section, ordinary least squares (OLS) estimates of this effect will likely be biased by the severity of illness (among other unobserved factors), which is omitted from the regression and a determinant of both health care choice and health outcomes.

We propose instrumenting for health care choice using variation in one component of the relative price of health care in the formal sector. The most salient costs to health care in developing countries are those incurred through travel (Gertler et al. 1987). Since care at health facilities and hospitals is often free or heavily subsidized by the government, the most relevant factors contributing to costs are the time and general disutility associated with traveling to the source of care.

¹⁴Larger values of the index correspond to more knowledge in the particular dimension being measured.

In countries with a low density of health facilities, and especially in sparsely populated rural areas, visiting a health facility for treatment often involves walking or riding public transportation for hours; individuals must incur the opportunity cost of this time and the disutility of strenuous travel if they choose to visit a formal health care provider. On the other hand, informal health care options—such as self-treatment with medicines obtained from a drug store or kiosk, a village health worker or a traditional healer—are often much more accessible in terms of distance.

Gertler et al. (1987) and others have used the distance to the nearest health facility as a proxy for the price of care. In our setting, using this variable as an instrument for the choice of care is likely invalid. Endogenous placement of health facilities—for example, allocating facilities to areas with very poor population health—is a primary concern. Moreover, more remote areas are less likely to have health facilities nearby; the manifold direct effects of living in remote area on health would thus invalidate the exclusion restriction.

We improve upon the use of variation in distance by interacting this variable with rainfall at the time the child fell sick, while controlling for the main effects of distance and rainfall in the first and second stages of the two-stage least squares estimator. That is, only the interaction is excluded; the main effects are allowed to have direct effects on both health care choices and health outcomes.

Let O_{ij} denote a health outcome for (sick) child i in cluster j ; let h_{ij} denote the health care choice made for the child; and let \mathbf{X}_{ij} denote a vector of child-, mother- and household-level characteristics. Denote the distance to the closest health facility as d_{ij} , and the quantity of rainfall in cluster j at the time of the child's illness as R_{ij} . Using this notation, the two-stage IV is specified as:¹⁵

$$\text{1st stage: } h_{ij} = \alpha_1 (d_{ij} \times R_{ij}) + \alpha_2 d_{ij} + \alpha_3 R_{ij} + \mathbf{X}_{ij}' \alpha_4 + \zeta_{ij} \quad (9)$$

$$\text{2nd stage: } O_{ij} = \beta_1 h_{ij} + \beta_2 d_{ij} + \beta_3 R_{ij} + \mathbf{X}_{ij}' \beta_4 + \varepsilon_{ij} \quad (10)$$

3.1 Intuition

The intuition behind the instrument is the following. The main effect of distance should be negative; that is, traveling a greater distance should discourage the usage of formal sector facilities for individuals seeking care.¹⁶ We posit that heavier rain should discourage individuals who live farther away *more* than individuals living closer to the nearest facility. Imagine for example that one household is located just next door to a facility, while another is located 10 kilometers away. In times of dry weather, clearly the household next door will be more likely to choose health facility care than the one farther away. But in times of heavy

¹⁵We use linear probability models throughout our analysis.

¹⁶We verify that this is the case, by omitting the interaction instrument from the first stage regression and estimating the main effects of nearest health facility distance and rainfall on health care choice. The results of this estimation are reported in Table A1, column 1. The estimated coefficients on distance and rainfall are both strongly negative.

rains, the rain should incrementally deter the farther household *more* than the one just next door.

Note also that in our baseline specification, we control for the historical average of rainfall in the individual's cluster. We do this to address the possibility that individuals living in wetter places might adjust their behavior (i.e., use different road networks or modes of transportation) in accordance with the prevalence of rainfall in their locality. We thus control for this "expected" amount of rainfall, and effectively use the "innovation" in rainfall realized in a particular month. The interaction instrument coefficient changes little, however, with and without the historical rainfall controls.

In section A in the appendix, we detail a variety of ways in which we address threats to this identification strategy. In particular, we address 1) the possibility that rainfall might have differential effects across remote and non-remote locales; 2) non-random selection into the self-reporting of recent illness; 3) the possibility that the instrument directly drives changes in health; 4) the potential role of past and future rainfall on contemporaneous health care-seeking behavior; 5) robustness to interactions of rainfall with a variety of covariates; and 6) robustness to non-linear distance terms. Details are provided in the appendix; in general we find strong support for the validity of the interaction instrument strategy in our setting.

3.2 Preliminary Motivating Evidence

Before employing this strategy below, we validate the intuition behind the instrument and its impact on health-seeking. We begin by showing that there is sufficient geographic variation in rainfall each month to drive variation in health care choice. We plot (in blue) the mean rainfall in each month of the survey period and the 95% confidence intervals around these means. For sake of comparison, we also plot the 50-year historical means and standard deviations of rainfall for each month. These plots are depicted in Figure 3.

We see in Figure 3 that there is, indeed, a great deal of geographic variation in rainfall each month, and that rainfall in the survey period conforms reasonably well to the historical distribution. Next, we check that the rainfall distributions, both during the survey period and historically, do not vary systematically by distance to health facility. We repeat the exercise from Figure 3 for subsamples of households with above and below median distance to nearest health facility; these plots are presented in the Appendix (Figure A6).

Next, we explore the degree to which health care choice varies by distance to health facility. We plot the fraction of households seeking care as a smoothed function of the distance between the household and the nearest formal sector care facility. Figure 4 shows that, indeed, health-seeking is a monotonically decreasing function of distance to facility.

Finally, this leads us to explore the degree to which the slope of the health-seeking function is steeper at higher rainfall realizations (relative to the historical mean of rainfall). This is the core intuition behind the proposed instrument. We depict the identifying variation in the data in two ways. First, we repeat the exercise from Figure 4 for subsamples of households experiencing rainfall during the survey period above and below the historical mean rainfall for that month. These two curves are depicted in Figure 5 and show clearly that higher

rainfall exacerbates the degree to which distance discourages health-seeking. This relationship is demonstrated through a contour plot in Figure A7 in the Appendix.

4 Results

4.1 First Stage Results

Table 2 presents results from the first stage regressions of healthcare choice on the interaction of rainfall and distance to nearest health facility for child level and mother level samples, respectively. All standard errors, here and in the results presented below, are clustered at the sampling cluster level. The first stage effects are negative, significantly different from zero and robust to the inclusion of various controls.¹⁷ Along with various demographic controls, we include region fixed effects representing the 26 regions in the data to ensure that the instrument is not picking up broad geographic variation across regions.

The specifications also control for demographic and geographic characteristics such as age and gender of child; household wealth; mother's education; mother's age at marriage; year in which mother was married; region, altitude, and size of household; number of living children and number of children under the age of 5 in the household and a dummy for whether the household is located in a rural or urban area.¹⁸

To the extent that rainfall and its subsequent effects on cost of travel are predictable, the ability of the instrument to predict healthcare choice in response to acute health shocks may be impaired. In order to account for this predictability of rainfall and the possibility that transportation infrastructure adapts to the predictable component of rainfall, we control for historical means and standard deviations of rainfall for the month of survey. Also, so as to ensure that the instrument is in fact reflecting specifically an increased cost of travel rather than a general impact of extreme weather on behavior, we control for average temperature in month of survey as well as historical mean and historical standard deviation. Table A1 columns 2–4 present results from the first stage specifications with sets of controls added incrementally.

4.2 Health Outcomes

Table 3 presents the main results from regressions of healthcare choice on the incidence of fever and malaria among the sample of children who reported being ill with fever in the two weeks prior to survey. The first two columns in Table 3 report (endogenous) OLS regression estimates. We find that children who sought care in the formal sector are roughly 9 percentage points less likely to still have fever or malaria at the time of survey. Columns 3 and 4 of Table 3 report IV estimates. Effects on fever and malaria are negative and precisely estimated, around 62 and 40 percentage points, respectively.¹⁹ The magnitudes of the estimated effects of care in the formal sector are quite large, particularly compared to the

¹⁷We make a note here regarding the interpretation of the coefficient estimate on the main effect of distance. This coefficient estimate is slightly positive, because most of the negative effect of distance is absorbed in the interaction of distance with rainfall. At the mean rainfall level we observe in the sample, the main effect of distance is indeed negative. This fact is reinforced by looking at the estimate of the main effect of distance when no interaction term is included – this estimate is reported in column 1 of Table A1, and is negative and significantly different from 0.

¹⁸We lose about 10 percent of observations due to missing values in control variables, thus the final sample sizes in these regressions are 1081 and 963 for the child and mother samples, respectively.

averages for fever and malaria prevalence in our sample. Of course, it bears mentioning that given that these impacts are not bound tightly around the point estimates, a fairly wide range of smaller impacts is possible.

What explains these magnitudes? We focus on two potential explanations. First, given that many fevers in young children in this context are malarial, treatment with effective antimalarials, particularly with ACT, should essentially reduce the probability of fever to 0.²⁰ On the other hand, inappropriate treatment (with ineffective antimalarials or with antipyretics alone) should have a small or transient effect on fever.²¹ Thus, we might expect that the treatment effect on children with malarial fevers would be large.

Second, as is the case with all LATE estimates, the appropriate mean by which to scale the size of the estimated effect is the mean of the dependent variable for the population *on the margin* of adoption. The characteristics of this sub-sample may be different than the sample of children on the whole. In Table 4, we explore heterogeneity by 1) the relative wealth of the child's household and 2) the child's gender in the effects of the instrument on healthcare choices in the first stage. The former dummy is constructed to equal 1 if the child's household is in the first or second quintile of the wealth distribution, for a wealth index generated using principal components analysis on household asset ownership data.²²

The top panel of Table 4 shows that the coefficient on the interaction instrument is larger for children in lower-wealth households and for male children (though not significantly so across the fully stratified models). The bottom panel of Table 4 reports means of the dependent variables for the portion of this subsample that does not seek care. These means might more closely reflect the scope for improvement in health outcomes through formal care. A comparison of the magnitudes of the coefficients of interest from the re reported in columns 2 and 4 to these means suggests that being exogenously driven to a formal sector health facility leads to nearly a full recovery from acute illness.

4.3 Mechanisms of Impact

Given the large estimates of effects on fever and malaria, we explore potential mechanisms through which care in the formal health care sector might improve these outcomes. While there are several possible mechanisms, we focus on three here due to data limitations.

4.3.1 Receipt of Medication—We begin by estimating the effects of formal sector care on the number and type of medications taken by acutely ill children. Improving access to care in the formal sector may simply entail enabling access to appropriate medications. DHS contains self-reported medications bought or received for all children who reported being sick with fever in the two weeks preceding survey.

¹⁹The fact that the OLS estimates are smaller than the IV is consistent with severity bias. The large difference in magnitudes between OLS and IV estimates is also consistent with related work. Table A8 reports a comparison of the estimates of severity bias found in this paper with those found in other work.

²⁰It bears mention, however, that overdiagnosis of fever as malaria is very common in Tanzania and similar African settings, though the probability of misdiagnosis is increasing with age, so equating fever and malaria in young children in highly endemic malarial settings generates less egregious diagnostic error (see, e.g., Achyuta Adhvaryu (2014).

²¹See, for example, Jessica Cohen, Pascaline Dupas, and Simone Schaner (2013).

²²The first quintile is the lowest relative wealth, and the fifth is the highest.

We look first at the quantity margin, namely, the total number of medications and the number of medications conditional on buying or receiving at least one medication. These results are reported in columns 1 and 2 of Table 5. The coefficient estimates on formal sector care (0.57 and 0.22, respectively) are positive and fairly large—about 38 and 13 percent of their respective means—but both are statistically insignificant. In columns 3 and 4, we examine the type of medications received, aggregating malaria and non-malaria medications. Again, we find a similar pattern: the coefficient estimates on formal sector care are positive, large compared to their respective means, but not statistically significant. Thus, if there was an increase in the number of medications, we unfortunately cannot detect it with confidence.

Finally, we look at one dimension of the quality margin: access to artemisinin-based combination therapy (ACT). ACT became the front-line antimalarial therapy in Tanzania in 2006 (a year before the DHS survey data we use were collected). The new therapy was much more effective than existing antimalarials at treating *falciparum* malaria, by far the most common type of malaria in Tanzania (Arrow et al. 2004).

ACT was initially available only in formal-sector health facilities (Adhvaryu 2014) (though leakage to the private sector was common). Indeed, in our data, 32 percent of children visiting health facilities with recent fever received ACT, while less than 4 percent of children who sought informal care purchased ACT. Of course, these differences are not causally attributed to formal sector usage. We thus ask: did formal sector care enable differential access to ACTs? The evidence from columns 5 and 6 of Table 5 suggest the answer is yes, but our estimates of the impact of formal sector care on ACT access are imprecise. Column 5 uses the full sample, while column 6 reports results conditional on obtaining any malaria medication. Both estimates suggest a large impact (just under 30 percentage points), but both are not estimated with precision. We interpret these results as suggestive evidence on access to higher quality malaria medication for children in the formal sector.

4.3.2 Medication-Related Behaviors—Policymakers often lament the “Last Mile” problem in global health: even when essential medicines are made accessible, promoting acceptance and proper usage in the population—the “last mile”—is often ignored or not incentivized. Since even the most effective medicine will not have its full effects if it is not taken up and used properly, solving this problem may have high health returns.

Accordingly, in addition to access to medications, we examine two important “last mile” behaviors: adherence to malaria therapy and delay before medicating. For each antimalarial reported as having been used as treatment, we construct indicators for whether the individual adhered to the full regimen for that particular antimalarial. We construct adherence measures based on self-reported number of doses of each therapy that was taken. We use the following variables: an indicator for adherence to at least one antimalarial; the number of antimalarials to which the patient adhered; and an indicator for adherence to all antimalarials taken. Details of the construction of these variables are provided in the Appendix.

We begin by examining adherence not conditional on receipt of antimalarial. That is, we combine the indicators for receipt of antimalarial with adherence to the particular antimalarial(s) received. Thus, the dependent variables in these regressions are only equal to 1 (or greater than 0 in the case of the count variable) if an antimalarial was received *and* the dosing regimen was completed.

The second-stage results on the effects of formal sector care on these unconditional adherence measures are reported in columns 1–3 of Table 6. Across the adherence measures, the coefficient estimates on formal sector are positive, fairly large, and statistically significant. For example, formal sector care increases the probability of receiving *and* adhering to at least one malaria therapy by nearly 60 percentage points, or nearly 100 percent of the mean of this adherence measure. In columns 4–6, we run the same specifications as in the previous columns but restrict the sample to children who took at least one antimalarial. In these conditional regressions, the coefficient estimates on formal sector care are positive and of similar magnitude to the previous (unrestricted sample) estimates, though not statistically significant.

Finally, we estimate the effect of formal sector care on delay (number of days) before receiving malaria medication. We define this variable as the number of days after falling ill the child begins taking malaria therapy. Of course, the delay variable is only defined for children who eventually got at least one antimalarial. The coefficient estimate, reported in column 7, is negative and statistically significant: formal sector reduces delay to medication by about 1 day.

Overall, while we cannot conclude with confidence that access to malaria therapy matters in mediating the impact of formal sector care on health, the results described above indicate that formal care does improve access *and* appropriate usage of antimalarials, and that those who use the formal health care sector are less likely to delay malaria medication. Although data limitations restrict our ability to explore additional mechanisms, these results suggest that information regarding optimal medication-related behaviors is an important mediating mechanism for the impact of formal care on health outcomes.

These results seem sensible in light of two important facts. First, malaria is a known quantity in this context, and thus information regarding its causes, treatment, and prevention are more or less well known. Yet, there is large heterogeneity here. For example, Bjorkman-Nyqvist, Svensson, and Yanagizawa-Drott (2013) report, from a study in Uganda, that 34 percent of the households have “severe” misconceptions about the transmission of malaria. Extrapolating a bit to the East African context, the data from Bjorkman-Nyqvist et al. (2013) suggest that while most people have a basic understanding of malaria, a nuanced understanding of the disease, and of its treatment, may not be common. Second, while it does seem unlikely that people with frequent interaction with antimalarials forget how to take medication, it seems plausible that 1) individuals do not know what the age-specific treatment regimen is for the particular antimalarial they take (e.g., a smaller dose is prescribed for younger children, and different combination therapies (ACTs) have different regimens) (Kachur et al. 2004); and 2) salience might matter, so that a reminder from the

doctor to finish the regimen may improve adherence even if the patient is familiar with the regimen already (see, e.g., Pop-Eleches et al. 2011).

4.3.3 Effects on General Health Knowledge—Table 7 reports estimates from second stage IV regressions on indices measuring the mother's information about the existence of tuberculosis and sexually transmitted diseases including HIV/AIDS, the transmission of these diseases, and the existence of treatments for these diseases.²³

Column 1 of Table 7 reports estimates of the effect of a visit to formal care on the composite index of all this disease related information. We find no significant effects on health-related knowledge of the mother. Though we lack precision, the point estimate is quite small (less than .5) in absolute value, as compared to a mean of 14. Columns 2–4 show results on sub-indices that measure information on the specific topics of disease existence, transmission, and treatment, respectively. We do not find significant effects on information by topic, with similarly small point estimates. Overall, the evidence presented here indicates that formal sector care does not appear to affect mothers' knowledge about common diseases and their treatment.

4.4 Instrument Checks

Finally, we assess the importance of several threats to the validity of our empirical strategy. We describe these analyses in brief here and in more detail in the Appendix.

4.4.1 Differential Effects of Rainfall by Remoteness—We address the concern that fluctuations in rainfall might have differential effects across remote versus non-remote locales, and this disparity may drive differences in health outcomes across the two types of areas. Since distance to the nearest health facility is likely to be correlated with general remoteness, the interaction of distance with rainfall may not be excludable if the differential effects of rainfall by remoteness are not adequately accounted for. To address this problem, we include the interaction of rainfall with a proxy for remoteness (as well including its main effect)—distance to the nearest marketplace. We are thus absorbing the variation in health facility distance that is associated with this proxy for remoteness.

As further evidence, we report in the first column of Table A3 results from the first stage regression including additional controls of interactions of rainfall with all other covariates of the household. If the interaction instrument is merely picking up on some unobservable remoteness or lack of access to resources of the household, which is then exacerbated by rainfall and ultimately predictive of health outcomes irrespective of health care choice, we should expect that controlling for the interaction of rainfall with demographic covariates such as wealth, education, household size, etc. would attenuate the coefficient on the instrument in the first stage regression. We find, however, that the first stage, and in fact most of the second stage results reported in columns 2–6, are robust to these additional controls.

²³The corresponding first stage regression results from this mother-level sample are reported in the second column of Table 2.

4.4.2 Selection into Self-reporting of Acute Illness—We address the concern that the instrument could change the self-reporting of acute illness. Perhaps in remote areas, high rainfall could result in differentially more sick individuals, or more individuals *perceiving* themselves to be acutely ill. This compositional shift could then account for part of the observed impact on health outcomes.

We measure the extent of this problem by regressing a dummy for reporting recent fever (the sample inclusion criterion) on the instrument, the main effects, and all controls as described above. The results are reported in column 1 of Table A4. We estimate a small coefficient tightly bound around zero, indicating that the instrument does not appear to induce greater reporting of recent acute illness.

4.4.3 Falsification Tests—We perform a variety of falsification tests. We answer the following questions: 1) does the instrument drives changes in malaria for the non-acutely ill sample (no fever in the past 2 weeks); and 2) does past (1 month before acute illness) or future (1 month after acute illness) rainfall have a differential effect on health care choice by health facility distance?

We report the results of these empirical tests in columns 2–4 of Table A4, and discuss the specifications and results more in depth in the Appendix. The results suggest that the answers to both the above questions are in the negative, lending further support to our identifying assumptions.

4.4.4 Robustness Tests—We perform a variety of additional robustness tests. First, we ask if the coefficient on the instrument in the first stage is affected by the inclusion of interactions of past and future rainfall with health facility distance (Table A5, column 1). Second, we ask whether the second stage results are affected by the inclusion of interactions of past and future rainfall with health facility distance (in both stages) (Table A5, columns 2–6). Third, we test whether the first and second stage results are robust to the inclusion of non-linear health facility distance terms (Table A6). We find in each case that the results are robust to the inclusion of additional controls. A more in-depth discussion is contained in the Appendix.

4.5 Correlation of the Instrument with Severity

Lastly, we check whether children living farther away from a health facility are more severely ill on average and require more rainfall to discourage their health-seeking. In order to address this concern, we control for measures of severity and their interaction with rainfall and check whether our main results are preserved. Table A7 reports results from first stage and select second stage regressions including additional controls of rainfall in the month of reported illness interacted with anemia and a dummy for whether the child is under the age of 1 as measures of severity as well as the main effects of these measures. The results suggest that the results are not driven by a correlation between the instrument and severity of illness.

5 Conclusion

This paper estimates the returns to treatment in the formal health care sector in a nationally representative sample of children in Tanzania. Using geographic and temporal variation in the cost of formal sector care to isolate exogenous variation in health care choice, we estimate large reductions in fever and malaria for sick children receiving treatment in the formal health care sector. These reductions are at least in part due to timely access to and appropriate usage of antimalarial therapy.

Our results underscore the importance of public sector health facilities in the management of fever for children in resource-poor environments. Policymakers and academics alike have, perhaps rightfully, levied criticism on the inadequate quality of health care at formal sector facilities. Nevertheless, we find that, at least for the population on the margin of treatment, these facilities do much better than self-treatment or traditional remedies.

Our results also speak to the importance of tackling the Last Mile Problem in global health. When we examine mechanisms of impact, we find the strongest effects of formal sector care on the receipt of *and adherence to* antimalarial therapy. Even the most effective treatments need to be taken properly to realize their full effectiveness: incentivizing patients to adhere to treatment regimens, either through the transfer of knowledge or through monetary incentives, may have large health returns.

In related work (Adhvaryu and Anant Nyshadham 2012, 2013), we estimate the schooling and labor supply impacts of formal sector care access for children with fever, and examine how households' self-employment patterns change in response to treatment in the formal health care sector. Given the potential gains to improved access, future research should rigorously test the impacts of policy solutions that improve access to care. For example, governments could invest in transportation voucher programs in rural areas to subsidize the cost of traveling to the facility on common transportation forms, e.g., buses or rented bicycles. In particularly remote populations, where health facilities are effectively too far to travel to, policy could focus on equipping community health workers with proper diagnostics and treatment for common illnesses, and providing training that emphasizes timely treatment and adherence to treatment regimen.

References

1. Adhvaryu, Achyuta. Learning, Misallocation, and Technology Adoption: Evidence from New Malaria Therapy in Tanzania. *Review of Economic Studies*. 2014 forthcoming.
2. Adhvaryu, Achyuta; Nyshadham, Anant. Schooling, Child Labor and the Returns to Healthcare in Tanzania. *Journal of Human Resources*. 2012; 47(2):364–396. [PubMed: 24353348]
3. Adhvaryu, Achyuta; Nyshadham, Anant. Health, Enterprise, and Labor Complementarity in the Household. Working paper. 2013
4. Arrow, KJ.; Panosian, C.; Gelband, H., et al. Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance. National Academies Press; 2004.
5. Almond, Douglas; Doyle, Joseph. After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays. *American Economic Journal: Economic Policy*. 2011; 3(3):1–34.
6. Almond, Douglas; Doyle, Joseph; Kowalski, Amanda; Williams, Heidi. Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns. *Quarterly Journal of Economics*. 2010; 125(2):591–634. [PubMed: 20634927]

7. Bjorkman-Nyqvist, Martina; Svensson, Jakob; Yanagizawa-Drott, David. The Market for (Fake) Antimalarial Medicine: Evidence from Uganda. Working paper. 2013
8. Bobonis, Gustavo; Miguel, Edward; Sharma, Charu Puri. Iron Deficiency Anemia and School Participation. *Journal of Human Resources*. 2006; 41(4):692–721.
9. Buchmueller, Thomas C.; Jacobson, Mireille; Wold, Cheryl. How far to the hospital?: The effect of hospital closures on access to care. *Journal of Health Economics*. 2006; 25(4):740–761. [PubMed: 16356570]
10. Chang, Fwu-Ranq. Uncertainty and investment in health. *Journal of Health Economics*. 1996; 15(3):369–376. [PubMed: 10159447]
11. Cohen, Jessica; Dupas, Pascaline; Schaner, Simone. Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment. Working paper. 2013
12. Cropper, Maureen. Health, investment in health, and occupational choice. *Journal of Political Economy*. 1977; 85:1273–1294.
13. Das, Jishnu; Hammer, Jeffrey; Leonard, Kenneth. The Quality of Medical Advice in Low-Income Countries. *Journal of Economic Perspectives*. 2008; 22(2):93–114. [PubMed: 19768841]
14. Doyle, Joseph. Returns to Local-Area Emergency Health Care Spending: Using Health Shocks to Patients Far From Home. *American Economic Journal: Applied Economics*. 2011; 3(3):221–243. [PubMed: 23853699]
15. Doyle, Joseph; Ewer, Steven; Wagner, Todd. Returns to Physician Human Capital: Analyzing Patients Randomized to Physician Teams. *Journal of Health Economics*. 2010; 29(6):866–882. [PubMed: 20869783]
16. Gertler, Paul; Locay, Luis; Sanderson, Warren. Are user fees regressive? The welfare implications of health care financing proposals in Peru. *Journal of Econometrics*. 1987; 36:6788.
17. Geweke, John; Gowrisankaran, Gautam; Town, Robert. Bayesian Inference for Hospital Quality in a Selection Model. *Econometrica*. 2003; 71(4):1215–1238.
18. Grossman, Michael. The human capital model. In: Culyer, AJ.; Newhouse, JP., editors. *Handbook of Health Economics*. 1. Vol. 1. Elsevier; 2000. p. 347-408.chapter 7
19. Hogerzeil, Hans. The concept of essential medicines: lessons for rich countries. *British Medical Journal*. 2004; 329:1169–1172. [PubMed: 15539676]
20. Kachur, S Patrick; Khatib, Rashid; Kaizer, Ellen; Fox, Susan; Abdulla, Salim; Bloland, Peter. Adherence to Antimalarial Combination Therapy with Sulphadoxine-Pyrimethamine and Artesunate in Rural Tanzania. *American Journal of Tropical Medicine and Hygiene*. 2004; 7(6)
21. Klemick, Heather; Leonard, Kenneth L.; Masatu, Melkiory C. Defining access to health care: evidence on the importance of quality and distance in rural Tanzania. *American Journal of Agricultural Economics*. 2009; 91(2):347–358.
22. Leonard, Kenneth. Improving Health Outcomes By Choosing Better Doctors: Evidence of Social Learning about Doctor Quality from Rural Tanzania. Working paper. 2007
23. Leonard, Kenneth; Masatu, Melkiory. Variations in the quality of care accessible to rural communities in Tanzania. *Health Affairs*. 2007; 26(3):380–392.
24. Mboera LEG, Fanello CI, Malima RC, Talbert A, Fogliati P, Bobbio F, Molteni F. Comparison of the Paracheck-Pf test with microscopy, for the confirmation of *Plasmodium falciparum* malaria in Tanzania. *Annals of Tropical Medicine and Parasitology*. 2006; 100(2):115–122. [PubMed: 16492359]
25. Miguel, Edward; Kremer, Michael. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*. 2004; 72(1):159–217.
26. Mwabu, Germano; Mwanzia, James; Liambila, Wilson. User charges in government health facilities in Kenya: effect on attendance and revenue. *Health Policy and Planning*. 1995; 10(2): 164–170. [PubMed: 10143454]
27. Pop-Eleches, Cristian, et al. Mobile Phone Technologies Improve Adherence to Antiretroviral Treatment in Resource-Limited Settings: A Randomized Controlled Trial of Text Message Reminders. *Aids*. 2011; 25:825–834. [PubMed: 21252632]
28. Selden, Thomas. Uncertainty and health care spending by the poor: the human capital model revisited. *Journal of Health Economics*. 1993; 12:109–115. [PubMed: 10126487]

29. Strauss, John; Thomas, Duncan. Health, Nutrition, and Economic Development. *Journal of Economic Literature*. 1998; 36:766–817.
30. Thirumurthy, Harsha; Zivin, Joshua Graff; Goldstein, Markus. The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya. *Journal of Human Resources*. 2009; XLIII(3):511–552.
31. Thomas, Duncan; Frankenburg, Elizabeth; Friedman, Jed; Habicht, Jean-Pierre; Hakimi, Mohammed; Ingwersen, Nicholas; Jaswadi; Jones, Nathan; McKelvey, Christopher; Peltó, Gretel; Sikoki, Bondan; Seeman, Teresa; Smith, James P.; Sumantri, Cecep; Suriastini, Wayan; Wilopo, Siswanto. California Center for Population Research working paper, no. CCPR-070-06. 2006. Causal Effect of Health on Labor Market Outcomes: Experimental Evidence.
32. World Health Organization. Equitable access to essential medicines: a framework for collective action. *Essential Drug Monitor*. 2004; 4:813.
33. Zhu, Minmin; Dib, Hassan; Zhang, Xinping; Tang, Shengchun; Liu, Liqun. The influence of health insurance towards accessing essential medicines: The experience from Shenzhen labor health insurance. *Health Policy*. 2008; 88(2–3):371–380. [PubMed: 18514354]

A Checks and Additional Results

A.1 Additional Plots

Figure A6 below checks that the rainfall distributions, both during the survey period and historically, do not vary systematically by distance to health facility. We repeat the exercise from Figure 3 for subsamples of households with above and below median distance to nearest health facility.

We see from Figure A6 that the rainfall distributions for locations close to and far from health facilities are nearly identical both during the survey period and historically. Rainfall in February appears to differ by distance to facility; however, these differences also exist in the historical distributions and ought to be sufficiently controlled for by the inclusion of historical means and standard deviations as well as month of survey fixed effects.

In Figure A7 a contour plot of fraction of households seeking care against both distance to nearest health facility and rainfall deviations from the historical mean. Figure A7 also shows that distance and rainfall interact to determine health-seeking.

A.2 First Stage Robustness

Column 1 of table A1 reports results from the regression of just the main effects of rainfall, distance to nearest health facility, and distance to nearest market on healthcare choice, including all the usual controls. That is, the specification in column 1 does not include the interaction of rainfall and distance to nearest health facility nor the interaction of rainfall and distance to nearest market. As expected, the main effects of both distance and rainfall are negative and significant.

Columns 2–4 of table A1 present results from the estimation of alternate specifications of the first stage equation. The specification in column 2 includes no additional controls beyond those listed; the specification in column 3 includes only age, region, wealth, education, and month of survey group effects; and the specification in column 4 is a reproduction of the main specification in column 1 of table 2 which includes all controls and

group effects. Note that the estimates with and without controls are not statistically significantly different from each other at conventional levels.

A.3 Reduced Form Effects

Table A2 reports estimates of the reduced form effects of the interaction of rainfall and distance to the nearest health facility on the incidence of fever and malaria, the number of medicines received, adherence to malaria medicines, and the number of days delayed before receiving malaria medication conditional on receiving at least one malaria medication. Given that the point estimates in the reduced form regressions are arithmetically equivalent to the point estimates from the second stage IV regression multiplied by those from the first stage, the signs and magnitudes of the reduced form results are exactly as expected. Significance is also preserved across all regressions, excepting incidence of malaria.

A.4 Robustness to Additional Controls of Interactions of Rainfall with Other Covariates

Table A3 reports results from first stage and select second stage regressions including additional controls of rainfall in the month of reported illness interacted with all other covariates. The specifications reported in table A3 are identical to those reported in the main results except for these additional controls. Again, we find that the first stage is robust to the inclusion of these additional controls. Furthermore, in columns 2–6 we see that the main results are entirely preserved. The coefficients of interest are all of the same sign and magnitude as in the main results; they are also all significant at the same level, except for in the regression of incidence of malaria.

A.5 Selection into sickness

We test whether the instrument drives selection into sickness. The dependent variable in this test is whether the mother reported that her child under 5 had been sick at some time in the two weeks before date of survey. Results are reported in table A4 column 1. We estimate a tightly bound 0 on the effect of the interaction instrument on self-report of fever, indicating the instrument did not drive selection into the sick sample.

A.6 Falsification test among children who were not sick

We show next that the interaction term *does not* significantly predict prevalence of malaria. In particular, we regress the result of the malaria test given to children on the interaction of distance and rainfall, the distance and rainfall main effects, and the same controls as in the analysis above. We do this for the non-sick sample only, noting that the same regression run on the 40 sick sample is just the reduced form regression reported in column 2 of table A2.²⁴ We should *expect* to find an effect for the reduced form regression on the sick sample, if the instrument is driving formal-sector healthcare choice, and these changes are driving

²⁴Of course, we cannot replicate this analysis using fever (at the time of survey) as the dependent variable, since the question related to fever is only asked of individuals who answered “yes” to having been acutely ill in the two weeks prior to survey.

changes in outcomes. The coefficient estimate of this regression, reported in table A2 column 2, is indeed positive (though insignificant).

On the other hand, we should expect to find no effect of the interaction in the sample of non-sick children. The results of this estimation are reported in table A4 column 2. The results indicate that for the non-sick sample of children, the interaction term does not predict the local prevalence of malaria. This serves as evidence in support of the interaction instrument's exclusion from second stage regressions of health outcomes on healthcare choice.

A.7 Falsification Test with Past and Future Rainfall

Another concern we must deal with is the transience of the rainfall variation. Should this variation in the cost of formal-sector care be insufficiently transitory (that is, should it persist from last period or into next period), the interaction term could affect health outcomes outside of its effect on contemporaneous acute healthcare choice and therefore violate the exclusion restriction. To rule out this possibility, we explore the predictive power of the interaction of distance to nearest health facility with rainfall in the months prior to and after the individual was surveyed.

Columns 3 and 4 of table A4 report results from the estimation of specifications identical to the first stage regressions reported in table 2, but with rainfall in the month of interview replaced by rainfall in the month following the month of survey and rainfall in the month preceding the month of survey, respectively. These replacements are made in the interaction terms as well, of course. The results show that the interaction of future or past rainfall and distance to nearest health facility is not predictive of healthcare choice. These serve as falsification exercises for the validity of the instrument, verifying that the predictive power of the instrument is in fact coming from a transient shock to the cost of traveling long distances to formal-sector care.

A.8 Robustness of Results to Future and Past Rain Interaction Controls

Table A5 reports results from first stage and select second stage regressions including additional controls of rainfall in the month before and after the month of reported illness and their interaction with distance to nearest health facility. As in the previous table, the specifications reported in table A5 are identical to those reported in the main results except for these additional controls. Again, we find that the first stage is robust to the inclusion of these additional controls.

Furthermore, in columns 2–6 we see that the main results are, with some exceptions, preserved. Results from regressions of the incidence of fever and malaria on formal care are of the same sign and magnitude as the main results, and are statistically significant. The estimate of the effect of formal care on the number of medicines received is still insignificant. Results from the regressions of adherence to malaria medicines and the number of days delayed before receiving malaria medication on formal care are no longer significant at conventional levels, but are of the same sign.

A.9 Robustness of Results to Nonlinear Distance Terms

Here, we discuss the specific functional form requirements for the exclusion of the interaction term given that we have established the possibility that distance to nearest health facility is endogenous. In particular, if distance to nearest health facility enters the true data generating process of the health outcome nonlinearly, the interaction of distance with rainfall might correlate with these nonlinear terms and cause a violation of the exclusion restriction. That is, the interaction of rainfall and distance to nearest health facility might simply pick up the more extreme effects of living far away from a health facility on health outcomes if we do not control for these nonlinearities.

To explicitly account for this possibility, we include 2nd and 3rd degree polynomial terms in distance and dummies for the deciles of the distance distribution in the specifications reported in table A6. For the sake of parsimony, we chose not to include these terms in the main specifications of the analysis. However, it is clear from table A6 that the first stage and main second stage results are robust to the inclusion of nonlinear terms in distance.

Table A6 reports results from first stage and select second stage regressions including nonlinear terms in distance as additional controls. In particular, the specifications reported in table A6 are identical to those reported in the main results above except for the inclusion of 2nd and 3rd degree polynomials in distance to nearest health facility and a vector of dummies representing the deciles of the distance distribution. It is clear that the first stage is robust to the inclusion of these additional controls. Furthermore, in columns 2–6 of table A6 we see that the main results on the incidence of fever and malaria, the number of medicines received, adherence to malaria medicines, and the number of days delayed before receiving malaria medication conditional on receiving at least one malaria medication are preserved.

A.10 Robustness to Additional Controls of Interactions of Rainfall with Measures of Severity

Table A7 reports results from first stage and select second stage regressions including additional controls of rainfall in the month of reported illness interacted with anemia and a dummy for 42 whether the child is under the age of 1 as measures of severity as well as the main effects of these measures. The specifications reported in table A7 are identical to those reported in the main results except for these additional controls. Again, we find that the first stage is robust to the inclusion of these additional controls, though the F-statistic is a bit smaller. Furthermore, in columns 2–6 we see that the main results are again entirely preserved. The coefficients of interest are all of the same sign and magnitude as in the main results; they are also all significant at the same level, except for in the regression of incidence of malaria.

A.11 Comparison of Self-Selection Bias in Estimates

Table A8 reproduces the main health outcomes results from this study in columns 1 and 2. The remaining columns present results from other studies which also attempt to overcome the issue of bias due to self-selection on severity. The results show that severity bias

significantly attenuates estimates of the effects of healthcare quality on health outcomes. The degree of bias appears to be on a similar order of magnitude as the results found in this study.

B Construction of Variables

The following variables were constructed for use in the analysis:

- *formalhealthcare* = 1 if child visited a government or private hospital or health centre; *formalhealthcare* = 0 if child did not visit a government or private hospital or health centre
- *histmean* of rainfall and temperature are calculated using average rainfall or temperature over the month of survey in the year of survey back to the year 1949
- *histsd* is calculated as the standard deviation from historical mean for average rainfall or temperature in month of survey from 2007-1949
- $\text{shocknorm} = (\text{rain} - \text{histmean})/(\text{histsd})$
- *infodisease* = *TB* + *STD* + *HIV=AIDS* + *otherSTD*, where each variable on the RHS is a multinomial variable taking values 1 and -1 for the right and wrong answer, respectively, and value 0 for a response of “don’t know” in response to questions of whether the respondent had heard of the disease
- *infotransmission* = *TBair*+*TBtensils*+*TBtouch*+*TBfood*+*TBsex*+*TBunknown*+*AIDSabst*+*AIDScondom*+*AIDS1prtnr*+*healthyAIDS* +*AIDSmsqto*+*AIDSfood*+*AIDSswitch*, where each variable on the RHS is a multinomial variable taking values 1 and -1 for the right and wrong answer, respectively, and value 0 for a response of “don’t know” in response to questions of whether the disease could be transmitted in that particular way
- *infotreatment* = *TBcure* + *AIDSdrugs* + *AIDSdrugsbaby*, where each variable on the RHS is a multinomial variable taking values 1 and -1 for the right and wrong answer, respectively, and value 0 for a response of “don’t know” in response to questions regarding the existence of a cure for TB, drugs to help HIV/AIDS infected individuals live a long life, and drugs to block transmission of HIV/AIDS from an infected mother to her baby.
- *info* = *infodisease* + *infotransmission* + *infotreatment*
- *nummeds* is the number of medications the child received of any type, including 0 if the child did not receive any medication
- *nummedscon* is the number of medications the child received of any type, it is missing if the child did not receive any medication
- *nummalmeds* is the number of malaria medications the child received, including 0 if the child did not receive any malaria medication
- *numnonmalmeds* is the number of non-malaria medications the child received, including 0 if the child did not receive any non-malaria medication

- $adhereone = 1$ if the child received and adhered to at least one malaria medication; $adhereone = 0$ if the child did not adhere to any of the medications received or did not receive any medications
- $numadhere$ is the number of malaria medications the child received and adhered to, including 0 if the child did not receive any malaria medication or if he received malaria medications but did not adhere to any of them
- $adhereall = 1$ if the child adhered to all of the malaria medications received; $adhereall = 0$ if the child did not adhere to all of the malaria medications received or did not receive any medications
- $adhereonecon = 1$ if the child received and adhered to at least one malaria medication; $adhereonecon = 0$ if the child did not adhere to any of the medications received
- $numadherecon$ is the number of malaria medications the child received and adhered to, including 0 if he received malaria medications but did not adhere to any of them; it is missing if the child did not receive any malaria medication
- $adhereallcon = 1$ if the child adhered to all of the malaria medications received; $adhereall = 0$ if the child did not adhere to all of the malaria medications received
- $meddelay = \min(delay)$, where $delay$ is measured in days for each malaria medication observed and is missing if the child did not receive that malaria medication at all.

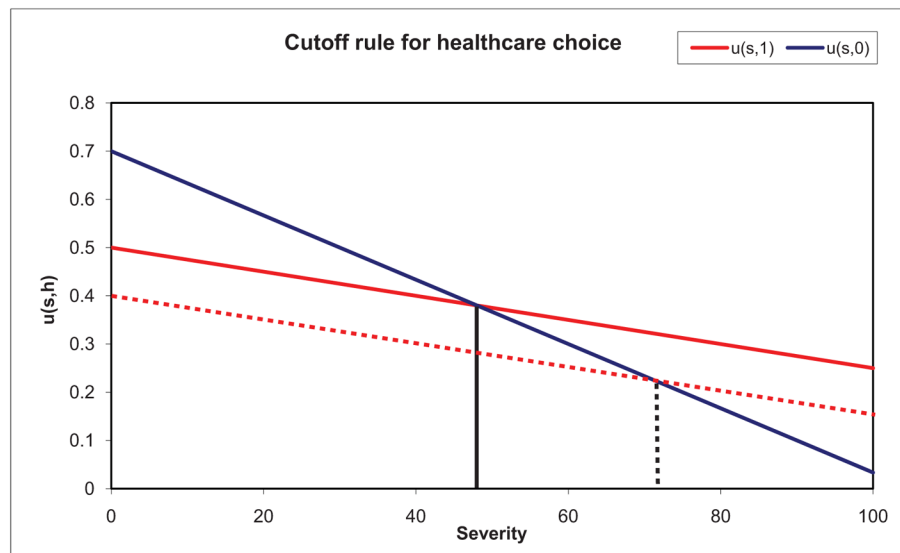


Figure 1.
Price increase induces rightward shift in cutoff value

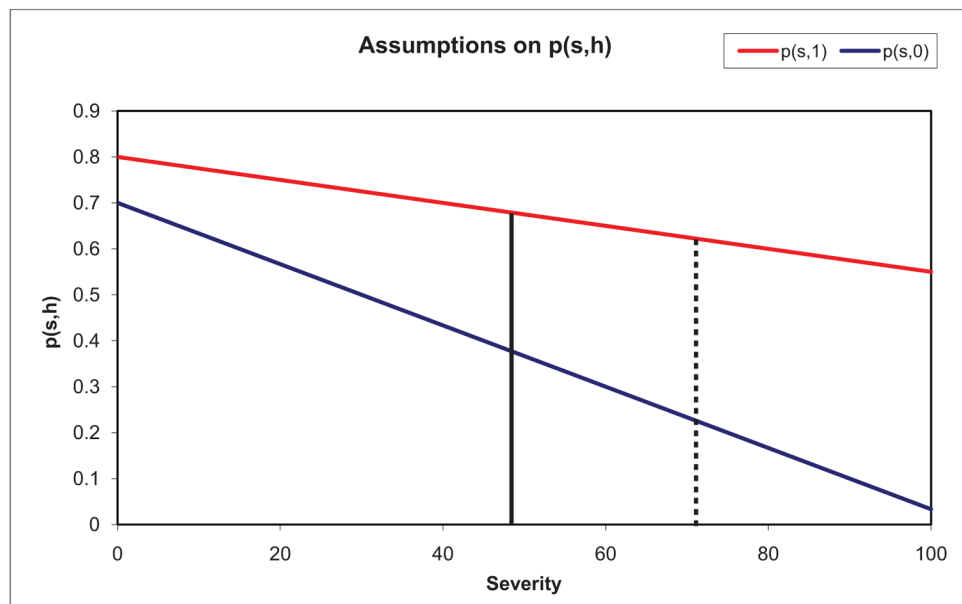
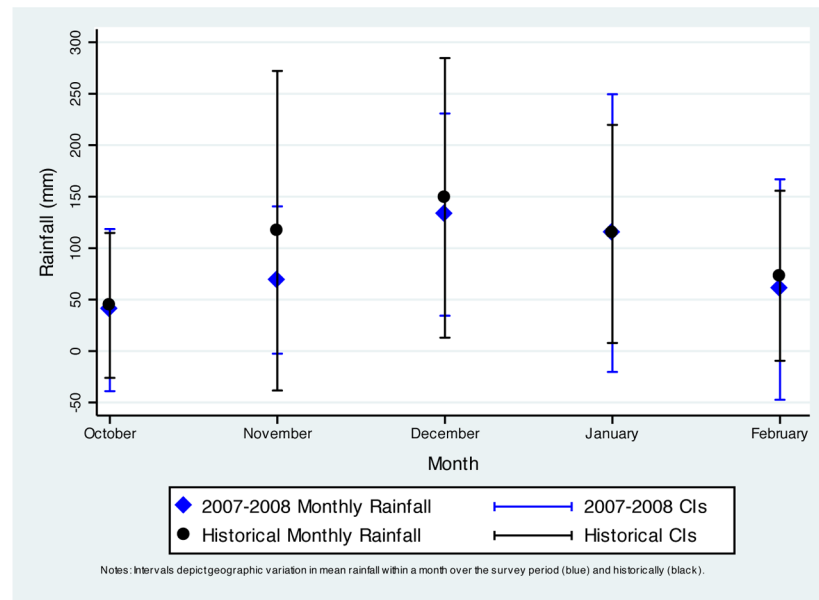


Figure 2.
Price shift, LATE and ATE

**Figure 3.**

Survey Period and Historical Monthly Rainfall Variation

Notes: Intervals depict geographic variation in mean rainfall within a month over the survey period (blue) and historically (black).

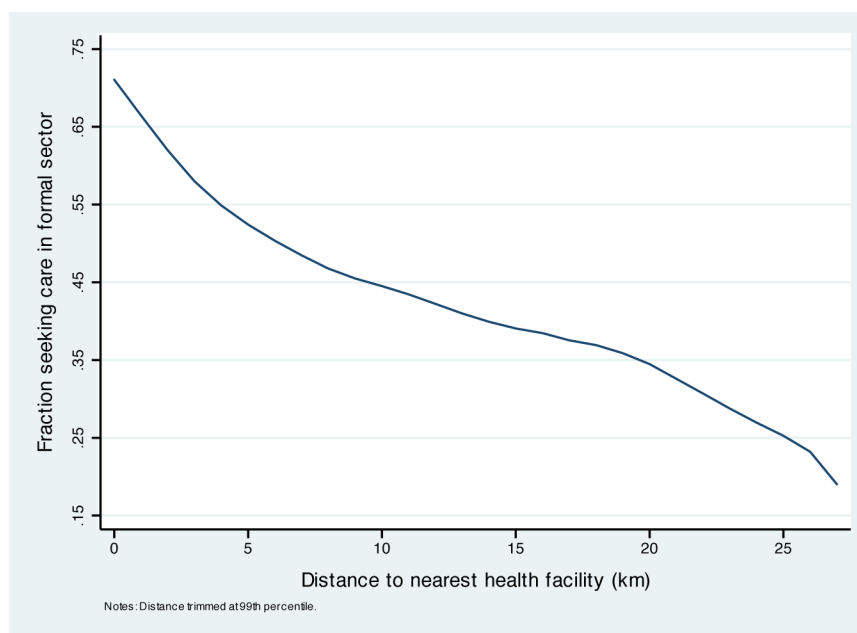
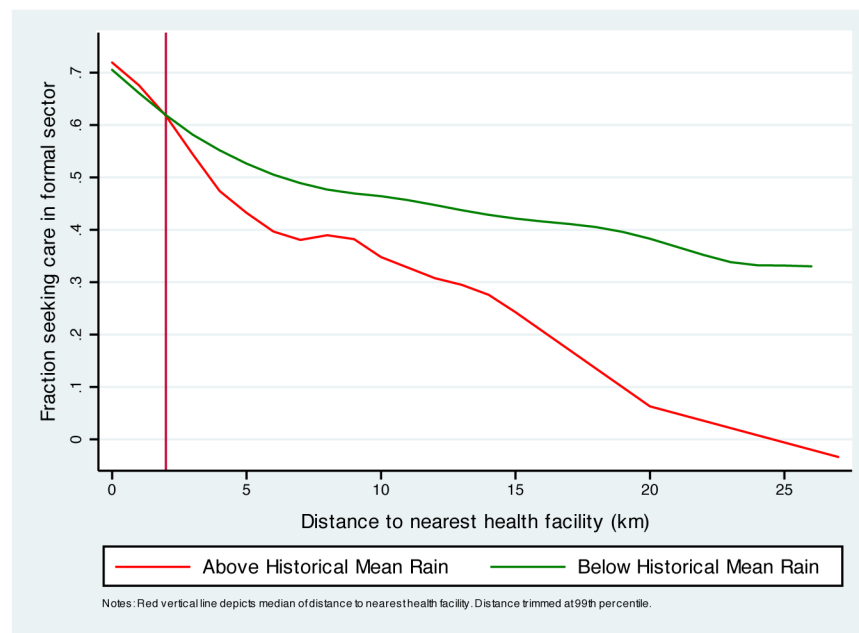


Figure 4.
Health Care Choice by Distance to Facility
Notes: Distance trimmed at 99th percentile.

**Figure 5.**

Health Care Choice by Distance to Facility and Rainfall

Notes: Red vertical line depicts median of distance to nearest health facility. Distance trimmed at 99th percentile.

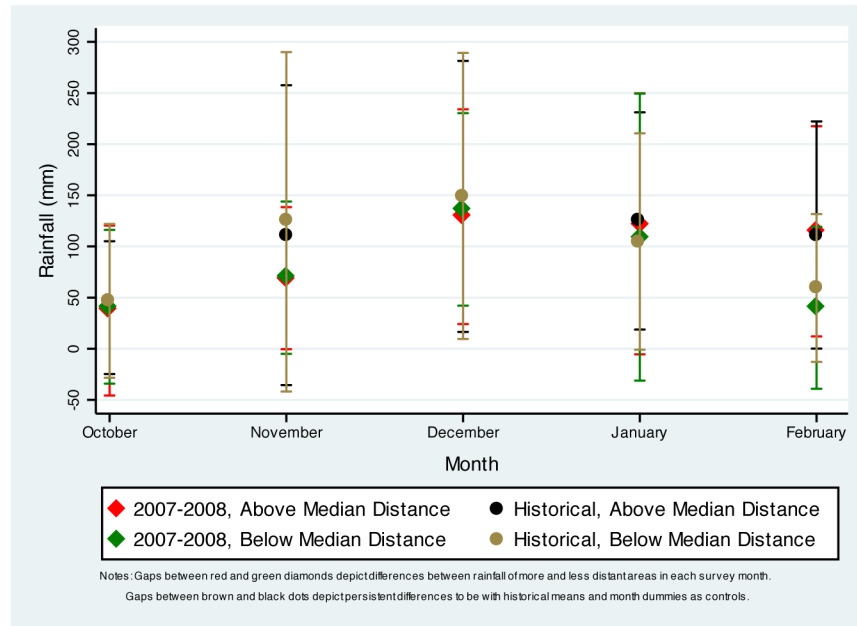


Figure 6.

Survey Period and Historical Monthly Rainfall Variation by Distance

Notes: Gaps between red and green diamonds depict differences between rainfall of more and less distant areas in each survey month. Gaps between brown and black dots depict persistent differences to be with historical means and month dummies as controls.

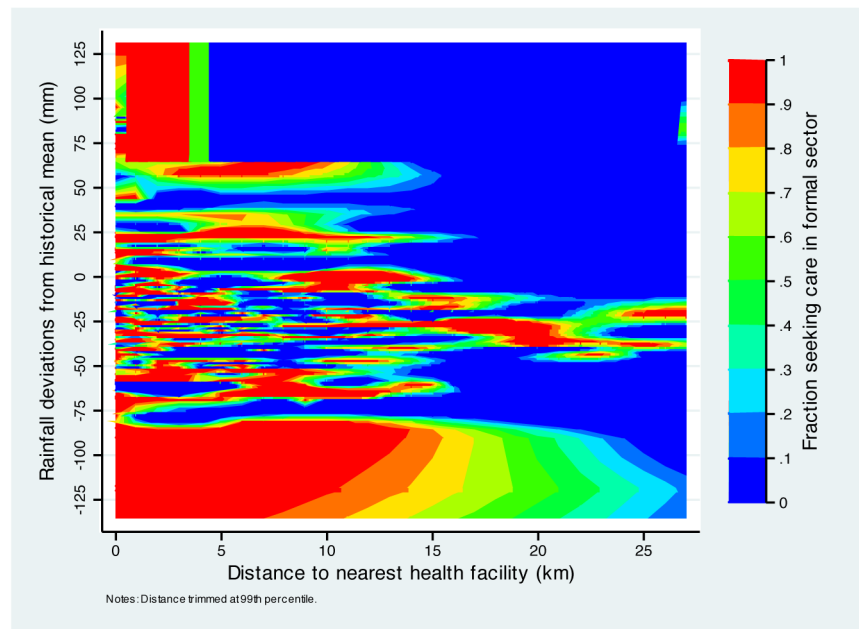


Figure 7.
Contour Plot of Health Care Choice, Distance to Facility, and Rainfall
Notes: Distance trimmed at 99th percentile.

Table 1

Summary Statistics

Number of Observations:									
All children	7502								
Children reporting sickness in 2 wks prior to survey	1200								
All mothers	4910								
Mothers reporting at least 1 child (<5) as acutely ill	1071								
Sick Children									
All	Formal Care		No Formal Care		Differences by Care				
	Mean	SD	Mean	SD	Mean	SD	t-statistic	p-value	
Sought care at a formal-sector healthcare facility	0.573	0.495							
Instrument (cost-of-care shifters):									
Rain (mm/100)	0.918	0.610	0.881	0.631	0.967	0.579	2.41**	0.0161	
Distance to nearest formal-sector health facility (km)	4.418	6.027	3.352	4.789	5.853	7.131	7.21***	0.0000	
Distance to nearest market (km)	26.293	29.585	22.127	27.652	31.927	31.168	5.64***	0.0000	
Health status:									
Tested positive for mild to severe anemia	0.770	0.421	0.777	0.417	0.761	0.427	-0.59	0.5538	
Tested positive for malaria (at time of survey)	0.232	0.422	0.165	0.372	0.319	0.467	5.91***	0.0000	
Fever now (self-reported, at time of survey)	0.272	0.445	0.213	0.409	0.351	0.478	5.37***	0.0000	
Demographic characteristics:									
Age	1.603	1.301	1.489	1.273	1.754	1.324	3.51***	0.0005	
Female	0.496	0.500	0.483	0.500	0.513	0.500	1.01	0.3138	
Rural	0.819	0.385	0.783	0.412	0.867	0.339	3.77***	0.0002	
Household size	7.254	3.856	6.815	3.427	7.842	4.298	4.60***	0.0000	
Medications:									
No. of medicines received	1.491	0.863	1.761	0.756	1.122	0.863	-13.42***	0.0000	
No. of medicines conditional on # > 0	1.694	0.708	1.795	0.722	1.511	0.644	-6.28***	0.0000	
No. of antimalarials	0.551	0.521	0.703	0.492	0.343	0.488	-12.39***	0.0000	
No. of non-antimalarials	0.940	0.762	1.058	0.808	0.779	0.661	-6.28***	0.0000	
Medication-related behaviors:									

	Sick Children							
	All	Formal Care		No Formal Care		Differences by Care		
		Mean	SD	Mean	SD	Mean	SD	t-statistic
Adhered to at least 1 medication	0.629	0.483	0.639	0.481	0.602	0.491	-0.83	0.4095
No. of medications adhered to	0.634	0.492	0.643	0.489	0.608	0.502	-0.77	0.4398
Adhered to all medications	0.619	0.486	0.628	0.484	0.596	0.492	-0.71	0.4767
No. of days delayed before receiving medication	1.180	1.079	1.150	1.071	1.264	1.099	1.15	0.2505
Health information indices:								
Composite	14.173	3.559	14.307	3.496	13.971	3.650	-1.42	0.1565
General disease-related	3.481	1.093	3.463	1.105	3.507	1.078	0.65	0.5156
Transmission-related	8.507	2.749	8.611	2.701	8.357	2.815	-1.48	0.1403
Treatment-related	1.909	1.298	1.996	1.275	1.778	1.324	-2.55**	0.0110

Notes: Please see data appendix for details on the construction of variables.

Table 2

First Stage: Interaction of Rain and Distance to Health Facility

Dependent var: 1(Sought care at a formal-sector healthcare facility)		
	Child Level	Mother Level
Rain x Distance	−0.0204*** (0.00516)	−0.0192*** (0.00534)
Distance	0.00569 (0.00489)	0.00348 (0.00506)
Market	−0.00167 (0.00130)	−0.00177 (0.00131)
Rain	−0.139* (0.0706)	−0.175** (0.0724)
Rain x Market	0.000776 (0.00118)	0.000912 (0.00122)
F-test: Rain x Distance=0	15.60	12.98
Prob>F	9.37e-05	0.000358
Observations	1081	963
R-squared	0.178	0.177

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, mother's age at marriage, year in which mother was married, and a dummy for whether the household is located in a rural or urban area. For the sake of parsimony, all coefficients are not reported here, but are available upon request. Results from linear probability model estimations shown.

Table 3

Effects of Healthcare Choice on Health Outcomes

	OLS		Second Stage IV	
	Fever	Malaria	Fever	Malaria
Formal Healthcare	−0.0945*** (0.0315)	−0.0961*** (0.0270)	−0.619** (0.266)	−0.398* (0.236)
Distance	0.00132 (0.00295)	0.00228 (0.00236)	−0.00442 (0.00385)	−0.00102 (0.00286)
Market	−0.00123 (0.00111)	0.000902 (0.000940)	−0.00191 (0.00130)	0.000706 (0.00105)
Rain	0.0640 (0.0855)	0.0577 (0.0599)	−0.0191 (0.106)	0.0152 (0.0702)
Rain x Market	0.000494 (0.000943)	−0.000885 (0.000779)	0.000612 (0.00108)	−0.000958 (0.000928)
Observations	1073	934	1073	934
R-squared	0.110	0.294	−0.173	0.189

Notes: Robust standard errors in parentheses (***) $p < 0.01$, (**) $p < 0.05$, (*) $p < 0.1$. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table 4

First Stage Heterogeneity and Health Outcome Means by Subsample

Dependent var: 1(Sought care at a formal-sector healthcare facility)					
	Household Wealth		Child's gender		
	Higher wealth >= 3rd quintile	Lower wealth < 3rd quintile	Female	Male	
Rain x Distance	-0.0194* (0.0112)	-0.0247*** (0.00591)	-0.0180** (0.00827)	-0.0215*** (0.00659)	
Distance	0.00933 (0.0127)	0.00841* (0.00500)	0.000553 (0.00699)	0.0109* (0.00637)	
Rain	-0.0365 (0.0973)	-0.225** (0.109)	-0.113 (0.0983)	-0.197** (0.0872)	
Observations	663	418	536	545	
R-squared	0.138	0.328	0.189	0.257	
Means within Subsample, No Formal Care					
Mean of Fever	0.303	0.408	0.314	0.384	
SD of Fever	0.460	0.493	0.465	0.487	
Mean of Malaria	0.226	0.434	0.319	0.311	
SD of Malaria	0.419	0.497	0.467	0.464	

Notes: Robust standard errors in parentheses (***) $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table 5

Effects of Healthcare Choice on Number and Types of Medications Obtained

	Any Medication		Medications By Type		Quality of Malaria Medication	
	No. of Meds	No. of Meds (Conditional on No. > 0)	No. of Malaria Meds	No. of Non-Malaria Meds	Artemisinin-based combination therapy (ACT) obtained	ACT conditional on any malaria medication
Formal Healthcare	0.572 (0.410)	0.218 (0.345)	0.283 (0.279)	0.290 (0.356)	0.299 (0.245)	0.308 (0.273)
Distance	-0.000531 (0.00619)	-0.00519 (0.00549)	0.000886 (0.00437)	-0.00142 (0.00546)	-0.00368 (0.00387)	-0.00760 (0.00582)
Market	-0.00227 (0.00243)	-0.00242 (0.00210)	-0.00231 (0.00143)	3.12e-05 (0.00187)	0.000107 (0.000969)	0.00144 (0.00166)
Rain	0.0546 (0.133)	0.122 (0.119)	-0.0473 (0.0837)	0.102 (0.125)	-0.133* (0.0786)	-0.171* (0.0911)
Rain x Market	0.000743 (0.00185)	0.000493 (0.00153)	0.000697 (0.00121)	4.60e-05 (0.00131)	-8.84e-05 (0.000650)	3.05e-05 (0.00103)
Observations	1049	931	1049	1049	1,078	565
R-squared	0.205	0.117	0.290	0.146	0.209	0.237

Notes: Robust standard errors in parentheses (***) $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Sample reductions in columns 1, 3 and 4 are due to missing observations in the outcomes. Further sample reduction in column 2 is due to conditionality in outcome. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table 6

Effects of Healthcare Choice on Adherence and Delay to Medication

	Unconditional			Conditional on Receipt of Malaria Medication			
	Adhered to at least 1 Med	No. of Meds Adhered to	Adhered to All Meds	Adhered to at least 1 Med	No. of Meds Adhered to	Adhered to All Meds	Days Delayed Before Medicating
Formal Healthcare	0.560** (0.277)	0.506* (0.285)	0.579** (0.278)	0.427 (0.314)	0.343 (0.322)	0.440 (0.315)	-1.328** (0.578)
Distance	0.00287 (0.00477)	0.00279 (0.00476)	0.00321 (0.00471)	0.00481 (0.00583)	0.00420 (0.00588)	0.00530 (0.00577)	-0.00618 (0.0119)
Market	0.000777 (0.00139)	0.000771 (0.00141)	0.000967 (0.00136)	0.00446** (0.00179)	0.00453** (0.00179)	0.00471*** (0.00179)	-0.00447 (0.00417)
Rain	-0.0659 (0.0948)	-0.0485 (0.100)	-0.0669 (0.0950)	-0.0844 (0.103)	-0.0528 (0.113)	-0.0916 (0.104)	0.0551 (0.238)
Rain x Market	-0.000731 (0.00122)	-0.000761 (0.00128)	-0.000877 (0.00122)	-0.00232* (0.00140)	-0.00249* (0.00146)	-0.00255* (0.00140)	0.000279 (0.00328)
Observations	1049	1049	1049	565	565	565	554
R-squared	0.065	0.096	0.042	0.036	0.073	0.024	-0.003

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Sample reductions in columns 1 through 3 are due to missing observations in the outcomes. Further sample reduction in columns 4–7 are due to conditionality in outcomes as well as missing values. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table 7

Effects of Healthcare Choice on Health-related Information

	Information by Topic		
	Composite	Information About Disease	Information About Treatment
Formal Healthcare	-0.421 (2.086)	-0.0276 (0.575)	0.358 (0.754)
Distance	-0.0510 (0.0321)	-0.00666 (0.0111)	0.00664 (0.0134)
Market	-0.0188** (0.00816)	-0.000181 (0.00310)	-8.97e-05 (0.00332)
Rain	-0.457 (0.693)	0.0653 (0.207)	-0.0253 (0.262)
Rain x Market	0.0113* (0.00582)	-0.00178 (0.00224)	2.53e-05 (0.00251)
Observations	845	963	854
R-squared	0.258	0.157	0.170

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table A1

Incremental Addition of Controls to First Stage

	Main Effects Only	No Controls	Fixed Effects	Main Specification
Rain x Distance		-0.0184*** (0.00543)	-0.0176*** (0.00513)	-0.0204*** (0.00516)
Distance	-0.0108*** (0.00310)	0.00161 (0.00529)	0.00390 (0.00502)	0.00569 (0.00489)
Market	-0.000794 (0.000756)	-0.00253** (0.00123)	-0.00182 (0.00130)	-0.00167 (0.00130)
Rain	-0.148** (0.0662)	-0.00756 (0.0409)	-0.0322 (0.0557)	-0.139* (0.0706)
Rain x Market		0.00110 (0.00106)	0.000851 (0.00114)	0.000776 (0.00118)
F-test: Rain x Distance=0		11.52	11.72	15.60
Prob>F		0.000761	0.000688	9.37e-05
Observations	1130	1130	1130	1081
R-squared	0.166	0.062	0.150	0.178

Notes: Robust standard errors in parentheses (***) $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, mother's age at marriage, year in which mother was married, and a dummy for whether the household is located in a rural or urban area. For the sake of parsimony, all coefficients are not reported here, but are available upon request. Results from linear probability model estimations shown.

Reduced Form Estimates

Table A2

	Health Outcomes			Medications		
	Fever	Malaria	No. of Meds	Adherence Days	Delayed Before Medicating	
Rain x Distance	0.0127** (0.00559)	0.00769 (0.00496)	-0.0125 (0.00941)	-0.0122* (0.00633)	0.0334** (0.0152)	
Distance	-0.00786 (0.00500)	-0.00300 (0.00424)	0.00323 (0.00871)	0.00655 (0.00634)	-0.0132 (0.0156)	
Market	-0.000845 (0.00113)	0.00110 (0.000948)	-0.00304 (0.00225)	2.62e-05 (0.00129)	-0.00316 (0.00453)	
Rain	0.0663 (0.0838)	0.0612 (0.0585)	-0.0250 (0.119)	-0.144* (0.0779)	0.159 (0.225)	
Rain x Market	0.000119 (0.000983)	-0.00105 (0.000767)	0.00108 (0.00166)	-0.000405 (0.00101)	0.000797 (0.00312)	
Observations	1073	934	1049	1049	554	
R-squared	0.106	0.286	0.122	0.119	0.216	

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table A3

Robustness to Adding Interactions of Rainfall with Other Covariates

First Stage		Health Outcomes			Medications	
	Formal Healthcare	Fever	Malaria	No. of Meds	Adherence	Days Delayed Before Medicating
Formal Healthcare						
Rain x Distance	−0.0190*** (0.00557)	−0.771** (0.309)	−0.399 (0.265)	0.568 (0.462)	0.522* (0.291)	−1.403** (0.591)
Distance	0.00500 (0.00523)	−0.00569 (0.00417)	−8.56e−05 (0.00319)	−0.000368 (0.00645)	0.00210 (0.00463)	−0.00648 (0.0107)
Market	−0.00177 (0.00143)	−0.00259* (0.00151)	0.000700 − (0.00116)	0.00273 (0.00252)	0.00137 (0.00136)	−0.00628 (0.00421)
Rain	0.294 (0.577)	−0.641 (0.649)	−0.179 (0.490)	−1.005 (0.890)	−1.435*** (0.474)	0.902 (1.877)
Rain x Market	0.00109 (0.00132)	0.00169 (0.00132)	−0.000455 (0.00110)	0.00116 (0.00206)	−0.00136 (0.00111)	0.00285 (0.00357)
F-test: Rain x Distance=0	11.63					
Prob>F	0.000718					
Observations	1081	1073	934	1049	1049	554
R-squared	0.196	−0.328	0.215	0.224	0.115	0.024

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Specifications used in this table are identical to those reported in the main results, except for the inclusion of interaction of contemporaneous rainfall with all covariates listed in Table IV (as well as the main effects of all controls). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table A4**Instrument Checks**

	Selection on Sick	Reduced-form falsification	Falsification with Future Rain	Falsification with Past Rain
	Fever in last 2 weeks	Malaria (if no fever in last 2 weeks)	Formal Healthcare	Formal Healthcare
Future or Past Rain x Dist			−0.00563 (0.00506)	−0.000504 (0.00652)
Future or Past Rain			0.116** (0.0519)	−0.0428 (0.0859)
Future or Past Rain x Market			0.000694 (0.00102)	−5.87e-05 (0.00146)
Rain x Distance	0.000988 (0.00185)	0.00151 (0.00206)		
Distance	−0.000503 (0.00173)	0.000226 (0.00183)	−0.00382 (0.00693)	−0.00943 (0.00598)
Market	−1.72e-05 (0.000493)	−0.000149 (0.000517)	−0.00169 (0.00139)	−0.000770 (0.00123)
Rain	−0.0162 (0.0271)	−0.0409 (0.0307)		
Rain x Market	0.000299 (0.000409)	0.000826* (0.000482)		
F-test: Future or Past Rain x Distance=0			1.24	0.01
Prob>F			0.267	0.938
F-test: Rain x Distance=0				
Prob>F				
Observations	6267	4177	1130	1130
R-squared	0.048	0.157	0.167	

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, mother's age at marriage, year in which mother was married, and a dummy for whether the household is located in a rural or urban area. For the sake of parsimony, all coefficients are not reported here, but are available upon request. Results from linear probability model estimations shown.

Robustness to Future and Past Rain Interaction Controls

Table A5

	First Stage		Health Outcomes			Medications	
	Formal Healthcare	Fever	Malaria	No. of Meds	Adherence	Days Delayed Before Medicating	
Formal Healthcare							
Rain x Distance	−0.0225*** (0.00689)	−0.731** (0.291)	−0.505* (0.293)	0.652 (0.515)	0.112 (0.288)	−0.931 (0.798)	
Distance	0.000585 (0.00697)	−0.00680 (0.00734)	−0.00393 (0.00522)	−0.0131 (0.0120)	0.0107 (0.00771)	−0.0144 (0.0151)	
Market	−0.00224 (0.00188)	−0.00179 (0.00182)	0.000525 (0.00167)	0.00184 (0.00371)	0.000913 (0.00197)	−0.00556 (0.00537)	
Rain	−0.0863 (0.0821)	−0.0189 (0.110)	0.0751 (0.0830)	−0.0133 (0.161)	−0.252*** (0.0907)	0.0573 (0.252)	
Rain x Market	0.000673 (0.00115)	0.000405 (0.00110)	−0.000415 (0.00114)	0.000747 (0.00197)	−0.00106 (0.00107)	−0.00317 (0.00271)	
Future Rain x Dist	0.00263 (0.00535)	−0.00499 (0.00599)	−0.00544 (0.00404)	0.0142* (0.00842)	−0.00434 (0.00454)	0.0145 (0.0131)	
Future Rain	0.0421 (0.0627)	0.119* (0.0620)	0.133** (0.0599)	−0.139 (0.102)	−0.0540 (0.0617)	−0.179 (0.185)	
Future Rain x Market	0.000652 (0.00117)	−0.00138 (0.00107)	−0.000797 (0.000918)	−0.000729 (0.00200)	−0.000771 (0.00101)	0.00107 (0.00321)	
Past Rain x Dist	0.00465 (0.00680)	0.00834 (0.00594)	0.0109** (0.00526)	−0.00539 (0.0111)	−0.00991 (0.00642)	−0.00464 (0.0219)	
Past Rain	−0.107 (0.0869)	−0.202** (0.0834)	−0.0298 (0.0757)	0.151 (0.120)	0.0111 (0.0765)	−0.485** (0.203)	
Past Rain x Market	5.22e−05 (0.00130)	0.00260** (0.00118)	0.000267 (0.00120)	−0.00356* (0.00194)	0.000911 (0.00136)	0.00789** (0.00382)	
F-test: Rain x Distance=0	10.65						
Prob>F	0.00120						
Observations	1130	1122	973	1098	1098	582	
R-squared	0.179	−0.274	0.111	0.205	0.150	0.116	

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Specifications used in this table are identical to those reported in the main results, except for the inclusion of rainfall in the month directly prior to and following the month of sickness and their interactions with distance to nearest health facility and distance to nearest market. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Robustness to Nonlinear Distance Terms

Table A6

	First Stage	Health Outcomes			Medications	
	Formal Healthcare	Fever	Malaria	No. of Meds	Adherence	Days Delayed Before Medicating
Formal Healthcare						
Rain x Distance	−0.0187*** (0.00502)	−0.584** (0.291)	−0.362 (0.254)	0.551 (0.456)	0.642** (0.289)	−1.206** (0.568)
Distance	0.00870 (0.0551)	−0.143** (0.0641)	−0.108* (0.0591)	0.106 (0.126)	0.155** (0.0786)	−0.105 (0.256)
Distance^2	0.000948 (0.00217)	0.00594** (0.00236)	0.00407* (0.00219)	−0.00424 (0.00564)	−0.00711** (0.00362)	0.00532 (0.0116)
Distance^3	−1.83e-05 (2.55e-05)	−7.40e-05*** (2.64e-05)	−4.54e-05* (2.43e-05)	5.49e-05 (7.70e-05)	9.52e-05* (5.19e-05)	−8.69e-05 (0.000161)
2nd Distance Decile	0.0181 (0.0724)	0.116 (0.0785)	0.126* (0.0686)	−0.266* (0.152)	−0.207** (0.0926)	0.000781 (0.294)
4th Distance Decile	0.0395 (0.116)	0.273** (0.129)	0.192* (0.116)	−0.290 (0.251)	−0.288* (0.154)	0.376 (0.501)
6th Distance Decile	−0.0236 (0.161)	0.365* (0.188)	0.307* (0.172)	−0.386 (0.360)	−0.400* (0.217)	0.370 (0.678)
7th Distance Decile	0.0103 (0.209)	0.577** (0.240)	0.452** (0.209)	−0.612 (0.448)	−0.639** (0.274)	0.499 (0.886)
8th Distance Decile	−0.120 (0.253)	0.481 (0.304)	0.506* (0.272)	−0.459 (0.571)	−0.596* (0.338)	0.409 (1.090)
9th Distance Decile	−0.171 (0.340)	0.780* (0.409)	0.675* (0.355)	−0.671 (0.751)	−0.710 (0.447)	0.489 (1.483)
10th Distance Decile	−0.307 (0.452)	0.951* (0.551)	0.827* (0.481)	−0.975 (0.922)	−0.939* (0.557)	0.709 (1.795)
Market	−0.00166 (0.00131)	−0.00199 (0.00131)	0.000662 (0.00105)	−0.00245 (0.00240)	0.000955 (0.00141)	−0.00430 (0.00412)
Rain	−0.123* (0.0708)	−0.0252 (0.104)	0.0126 (0.0700)	0.0835 (0.133)	−0.0563 (0.0949)	0.0208 (0.239)
Rain x Market	0.000857 (0.00118)	0.000729 (0.00104)	−0.000849 (0.000907)	0.000722 (0.00178)	−0.000961 (0.00122)	0.000754 (0.00318)
F-test: Rain x Distance=0	13.91					
Prob>F	0.0002					
Observations	1081	1073	934	1049	1049	554
R-squared	0.189	−0.120	0.219	0.213	0.030	0.043

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Specifications used in this table are identical to those reported in the main results, except for the inclusion of 2nd and 3rd degree polynomials in distance to nearest health facility, and a vector of dummies representing the deciles of the distance distribution. Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table A7

Robustness to Adding Measures of Severity and Their Interaction with Rainfall as Controls

First Stage		Health Outcomes			Medications	
	Formal Healthcare	Fever	Malaria	No. of Meds	Adherence	Days Delayed Before Medicating
Formal Healthcare						
Rain x Distance	−0.0173*** (0.00613)	−0.776** (0.340)	−0.390 (0.255)	0.379 (0.518)	0.460 (0.314)	−1.550** (0.636)
Distance	0.00360 (0.00583)	−0.00570 (0.00450)	0.000112 (0.00319)	−0.00219 (0.00726)	0.000675 (0.00482)	−0.00998 (0.0122)
Market	−0.00106 (0.00139)	−0.00271* (0.00143)	0.000611 (0.00115)	−0.00221 (0.00247)	0.000494 (0.00143)	−0.00566 (0.00436)
Rain	0.353 (0.606)	−0.910 (0.630)	−0.290 (0.482)	−1.133 (0.906)	−1.076** (0.487)	2.771 (1.974)
Rain x Market	0.000658 (0.00135)	0.00194 (0.00129)	−0.000376 (0.00107)	0.00116 (0.00203)	−0.000451 (0.00119)	0.00104 (0.00364)
F-test: Rain x Distance=0	8.00					
Prob>F	0.00					
Observations	941	935	934	914	914	492
R-squared	0.210	−0.265	0.235	0.228	0.163	0.018

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Specifications used in this table are identical to those reported in the main results, except for the inclusion of interaction of contemporaneous rainfall with all covariates listed in Table IV (as well as the main effects of all controls). Standard errors are clustered at the sampling cluster level. All specifications include age, region, wealth, education, and month of survey group effects. Other controls include temperature, historical means and standard deviations of both rainfall and temperature, altitude, household size, number of living children, number of children under 5, gender, and a dummy for whether the household is located in a rural or urban area. Sample sizes are reduced by some observations in these specifications due to missing values in the outcomes. Results from linear probability model estimations shown.

Table A8

Comparison of Self-Selection Bias in Estimates

Adhvaryu and Nyshadham (2012)		Gowrisankaran and Town (1999)	Doyle (2012)	
KHDS Kagera, Tanzania (1991–1994)		United States	United States	
Dependent Variable	Fever	Malaria	Fever	Mortality
Formal Healthcare		Formal Healthcare		Area-level Healthcare Spending
Independent Variable of Interest	Formal Healthcare		Not-for-profit Hospitals (vs. For-profit)	
Adjusted Estimate	IV −0.619** (0.266)	IV −0.398* (0.236)	IV −0.0055* (0.003)	OLS (Florida Visitors) −0.028*** (0.0082)
Biased Estimate	OLS −0.0945*** (0.0315)	OLS −0.0961*** (0.0270)	OLS −0.00016 (0.00047)	OLS (Florida Locals) −0.003 (0.0039)
Observations	1123	973	844	37185 (Visitors), 749762 (Locals)

Notes: Columns 1–2: see notes in table 3; column 3: see Adhvaryu and Nyshadham (2012) table II; column 4: see Gowrisankaran and Town (1999) table 4; column 5: See Doyle (2012) table 2.