ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 994

Healthcare Choices, Information and Health Outcomes

Achyuta Adhvaryu
Yale University

Anant Nyshadham
Yale University

March 2011

Notes: Center discussion papers are preliminary materials circulated to stimulate discussion and critical comments.
We thank Abhijit Banerjee, Prashant Bharadwaj, Michael Boozer, Joe Doyle, James Fenske, Jason Fletcher, T. Paul Schultz, Tavneet Suri, Chris Udry and participants at the Yale Development Lunch Seminar and the 2010 Pacific Development Conference for helpful comments and suggestions. All errors are our own.
Adhvaryu: achyuta.adhvaryu@yale.edu. web: http://www.yale.edu/adhvaryu.
Nyshadham: e-mail: anant.nyshadham@yale.edu.

Healthcare Choices, Information and Health Outcomes

Achyuta Adhvaryu and Anant Nyshadham*

September 29, 2010

Abstract

Self-selection into healthcare options biases estimates of the effects of healthcare on health outcomes. We exploit exogenous variation in the cost of formal-sector care to show that the use of such care improves short-term health outcomes for acutely ill children in Tanzania. Better treatment-specific information, rather than greater access to medicines, appears to be the primary mechanism for this effect: children who use formal-sector care are as a result more likely to get timely treatment and adhere to their medications.

JEL: I10, I18, O10, O12
Keywords: healthcare, information, child health, Tanzania

1 Introduction

Healthcare choices and health outcomes following an acute shock to one's health are likely to be simultaneously determined, not only by preferences and information but also by the severity of the shock (Maureen Cropper 1977; Thomas Selden 1993; Fwu-Ranq Chang 1996; Michael Grossman 2000). Self-selection into healthcare options based on severity of illness renders estimating the effects of healthcare choice on health outcomes difficult. In particular, comparing health outcomes across high- and low-quality healthcare options would underestimate the beneficial effects of quality on outcomes, since sicker individuals (with more severe illnesses) are both

*Adhvaryu: Yale University, e-mail: achyuta.adhvaryu@yale.edu. web: http://www.yale.edu/adhvaryu. Nyshadham: Yale University, e-mail: anant.nyshadham@yale.edu. We thank Abhijit Banerjee, Prashant Bharadwaj, Michael Boozer, Joe Doyle, James Fenske, Jason Fletcher, T. Paul Schultz, Tavneet Suri, Chris Udry and participants at the Yale Development Lunch Seminar and the 2010 Pacific Development Conference for helpful comments and suggestions. All errors are our own.

more likely to select into higher-quality options and have worse health outcomes on average (Gautam Gowrisankaran and Robert Town 1999; David Cutler et al. 2004).

Several recent studies in the United States have exploited "natural experiments" to estimate the returns to healthcare net of this bias. This has been done in the context of hospital choice (John Geweke et al. 2003; Thomas Buchmueller et al. 2006), emergency care (Joseph Doyle 2005, 2010), physician choice (Doyle et al. 2010), and postnatal and postpartum care (Douglas Almond and Doyle 2009; Almond et al. 2010).

We study the effects of formal-sector healthcare on the incidence of fever and malaria in young children in Tanzania. Our study advances the aforementioned literature in two main ways. First, despite a large number of studies in the developing country context on the effects of health interventions–e.g. nutritional supplements, preventive technology and treatments[1]–on outcomes, little is known about the returns to choosing formal-sector healthcare (over self-treatment or traditional care, for example) following an illness. We provide what is to our knowledge the first assessment of these returns in a developing country setting using methods which account for the bias induced by self-selection into healthcare options.[2]

Second, though disparities in quality of care (formal vis-a-vis informal healthcare options) have been well documented, the particular *mechanisms* through which improvements in health outcomes may be induced, in a causal sense, have not to our knowledge been explored. Our empirical strategy allows us to answer the following question: if we do observe improvements in health outcomes, *why* do these improvements occur? Does seeking care at a formal-sector facility enable greater access to medicines, as studies in public health have often suggested (Hans Hogerzeil 2004; WHO 2004; Minmin Zhu et al. 2008)? Or does formal-sector care induce beneficial changes in treatment-specific information and behaviors?[3] Our evidence suggests the latter

[1]See, 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).

[2]Victor Lavy et al. (1996) study the effects of accessibility of health services on health outcomes, but their the validity of their exclusion restriction is questionable. William Dow et al. (1997) study the effect of healthcare prices on health and labor outcomes, but as a result of their experimental design, the mechanism for the induced effects cannot be adduced.

[3]See, for example, Wen Chern et al. (1995), Maglosia Madajewicz et al. (2007), Julie Downs et al. (2009), Pascaline Dupas (2010), and Rebecca Thornton (2010).

is predominantly at play.

As shown in a simple model of healthcare choice, the structural parameter we estimate is the returns to formal-sector healthcare in the health production function. We identify these effects using an instrumental variables (IV) strategy which exploits exogenous variation in the costs of traveling to formal-sector health facilities. Since insurance mechanisms are rarely in place in developing countries and many governments heavily subsidize payments for care at health facilities and hospitals, the costs of healthcare to individuals are in large part related to the opportunity costs of traveling to the care option (Paul Gertler et al. 1987; Germano Mwabu et al. 1995). These costs can be economically substantial, especially in remote areas of the developing world.[4] Moreover, recent studies (Gertler 2004; Michael Kremer and Edward Miguel 2007; Kremer and Alaka Holla 2008; Jessica Cohen and Pascaline Dupas 2010) find that the price elasticity of health-related investment is very high in developing countries, suggesting that shifts in the costs of traveling to health facilities can plausibly induce large changes in the probability of choosing this option for care.

We use variation in this cost, generated by distance (to the nearest health facility) and rainfall, to predict the choice of healthcare in response to acute illness. We argue that distance alone, while it clearly affects the costs of choosing formal-sector care, is likely to be correlated with unobserved determinants of health outcomes. Health facilities in developing countries are often built or expanded where they are needed most (for example, in areas more prone to severe illnesses). We thus interact distance with rainfall in the month the child became sick, and use the interaction as the instrumental variable, while controlling for both main effects in the first- and second-stage specifications. The intuition behind the instrument is simple: rainfall generates random variation in the cost (or disutility) of traveling a given distance. In other words, heavier rains should discourage individuals who live farther away *more* than individuals living closer to the nearest facility.

We find that the instrument is strongly predictive in the first stage. It is robust to a variety

[4] In 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.

of additional controls and passes various falsification tests.[5] Using a sample of young children who were sick with fever and/or cough in the two weeks preceding survey, we focus on two main health outcome variables: the incidences of fever and malaria on the day of survey. We find that ordinary least squares (OLS) estimates of the effects of formal-sector healthcare use are small (and precisely estimated). In contrast, the IV estimates are several times as large in magnitude, consistent with self-selection based on the severity of illness. Our overall finding is that formal-sector care greatly reduces the incidence of fever and malaria amongst children who reported being acutely ill.

Given the large magnitudes of the IV estimates, next we ask *why* formal-sector care is more effective. Greater access to medicines at formal-sector health facilities does not appear to be the primary mechanism. IV estimates show that using formal-sector care does not significantly increase the number of medicines received for treatment. Moreover, we can test whether formal-sector care enables greater access to *appropriate* medicines, since guidelines for the treatment of fever in young children in malaria-endemic countries require that anti-malarial drugs be prescribed presumptively to this vulnerable group. We find no evidence that using formal-sector care increases access to malaria medication.

On the other hand, we do find evidence that formal-sector healthcare improves individuals' treatment-related information and behavior. In particular, we find that children using formal-sector care begin antimalarial treatment with substantially less delay and are much more likely to adhere to their medications. Further, using mother-level data on questions related to health perceptions, we construct a variety of indices measuring mothers' general knowledge of important diseases and their transmission and treatment. IV estimates show no evidence that bringing their children to formal-sector care increases mothers' health-related information *in general*, in-

[5] Additional 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 healthcare 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.

dicating that improvements in access to information at the formal healthcare sector are focused on the use of immediately relevant treatments, rather than on broad health concepts. Taken together, these results show, at least for the Tanzanian context, that the outcome gradient across the formal and informal healthcare sectors is driven by quality differences not related to access to medicines, but rather to access to specific, treatment-related information which spurs changes in patients' behaviors.

Measuring these effects, and understanding the mechanisms underlying them, is relevant in the developing country context for at least two reasons. First, there is substantial heterogeneity in the quality of care, particularly across formal and informal options (Mwabu et al. 1993; Jishnu Das et al. 2008; Heather Klemick et al. 2009). This potentially exposes individuals facing healthcare choices to different outcomes depending on the type of care they choose. In this sense, healthcare choice at the individual level matters perhaps to a greater degree than in the developed world.

Second, since insurance mechanisms do not exist in the developing world to the degree they do in developed countries, healthcare choice is more of an acute decision, in which an individual responds to a one-off shock (recognizing, of course, that one's stock of health determines the probability of shock and its ensuing severity). In developed-country settings, since much of healthcare is regulated by insurance providers and the choice of insurance is endogenous, it is difficult to find examples of shocks for which an individual's set of possible healthcare choices is exogenous.[6]

The remainder of the paper is structured as follows. Section 2 describes a model of healthcare choice and outcomes to motivate the empirical analysis. Section 3 describes the data we use and the construction of important variables. Section 4 explains the empirical strategy and provides evidence for the validity of the instrument. Section 5 presents the results, and Section 6 concludes.

[6]Work by Doyle (2005, 2010), as an exception, finds convincing solutions to this endogeneity problem as relates to emergency care in the United States.

2 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.

2.1 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 $D^g < D^b$.[7] Healthcare choice and severity combine to determine the probability that the good health outcome occurs. Denote $p(s, h)$ as the function which maps severity and healthcare choice into a probability: $p(s, h) = \Pr(D = D^g|s, h) \in [0, 1]$. By definition of severity s, $p(s, h)$ 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: $p(s, 1) - p(s, 0)$ 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.

[7] In our empirical setting, since we analyze a dataset of children, we can interpret D as the number of days a parent must spend in caring for her sick child.

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 $p(s,1)$ and $p(s,0)$ are likely to be close to 0). If this were true, then $p(s,1) - p(s,0)$ 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 $p(s,1) > p(s,0)$. 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 of formal-sector care is substantially larger than informal care, which establishes the tradeoff in price.

2.2 Utility maximization

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

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

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 ($p(s,h)$), when the budget constraint binds we can write the maximization problem as

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

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

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

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

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

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 $g(s) = p(s,1) - p(s,0)$) has an inverse (g^{-1}). Thus, the utility maximization problem can be expressed as a simple cutoff rule:

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

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 (Δu), for example due to an increase in the wage w, lowers the cutoff and thus increases the probability of choosing $h = 1$.

2.3 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

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

We can rewrite this quantity as $\Big(D^g - D^b\Big) \int_{-\infty}^{\infty} \Big(p(s,1) - p(s,0)\Big) f(s)ds$. Since, $D^g < D^b$, and $p(s,1) - p(s,0) > 0$ 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:

$$\begin{aligned} &\int_{K}^{\infty} \Big(\Big(p(s,1)D^g + (1 - p(s,1))D^b \Big) f(s)ds \\ &- \int_{-\infty}^{K} \Big(p(s,0)D^g + (1 - p(s,0))D^b \Big) \Big) f(s)ds \end{aligned} \tag{5}$$

Again, we can rewrite this quantity as $\Big(D^g - D^b\Big)\Big(\int_{K}^{\infty} p(s,1)f(s)ds - \int_{-\infty}^{K} p(s,0)f(s)ds \Big)$.

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):

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

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 and of p. For example, if we assume that severity s is uniformly distributed between 0 and 1, we can show that the bias is negative if and only if

$$\int_K^1 p(s,0)ds < \int_0^K p(s,1)ds. \tag{7}$$

Whether this inequality holds depends on K as well as the shape of p. Severity generates a more negative (or less positive) bias the larger the difference between $p(s,1)$ and $p(s,0)$, and the larger the value of K. The results presented in this paper are consistent with an 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 4. 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 ($\Delta P' > \Delta P$). 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

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

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 $p(s, h)$, 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 $p(s,1)$ and $p(s,0)$ 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 $p(s,1)$ over $s \in (K, 100]$ with the average of $p(s,0)$ over $s \in [0, K]$. On the other hand, the LATE estimate from the second stage IV regression compares the average of $p(s,1)$ over $s \in [K, K']$ with the average of $p(s,0)$ over $s \in [K, K']$.

3 Data

We use the 2007-2008 Tanzania HIV/AIDS and Malaria Indicator Survey (THMIS), which is part of the Demographic and Health Surveys (DHS) collected by MEASURE DHS (in conjunction with the Tanzanian National Bureau of Statistics), a project funded by the United States Agency for International Development. 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].

3.1 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

[8] The 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).

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.

We use two main health outcome variables. The first is a fever dummy 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 a dummy 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. Please refer to the Appendix for more details on construction of the variables.

3.2 Mother- and household-level variables

In subsequent analysis, we collapse the child-level data to the mother level, and analyze the level of healthcare-related information help by the sample of mothers who had at least one sick child. 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 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. Please refer to the Appendix for details on the construction of these indices.[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: dummies for age of the child (in years); region dummies; wealth dummies; mother's educational attainment in years; mother's age at marriage; year in which mother was married; dummies for the month of survey; household size; number of living children in household; number of children under age of 5 in household; a dummy for urban clusters; gender of the child; and household's altitude.

3.3 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-

[9] We 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.

[10] Note 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).

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.

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.

3.4 Summary Statistics

Table I 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 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.

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

The second and third columns of Table 1 compare means in 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.[12]

3.4.1 Current health status

Children who did and did not receive formal-sector care appear similar in terms of probability of 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.

3.4.2 Demographic characteristics

Across healthcare choices, sick children appear generally similar in terms of age, gender and rural residence. The two groups differ on mean household size, with those not receiving formal care coming from larger households.

3.4.3 Medications and medicine-related behaviors

Those who received formal sector care received more medicines on average, and the difference is apparent for antimalarials as well as non-antimalarials. Adherence, on the other hand, appears to be similar across healthcare choice groups.

3.4.4 Information indices

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

[12]We 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 4.

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. Please see the data appendix for more information on the creation of these indices.

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, though the differences are fairly small (as compared to the means).

4 Empirical strategy

In this section, we discuss the aim of our empirical analysis, describe our instrumental variables strategy, and discuss and present evidence related to the validity of our instrument. Our primary aim is to obtain unbiased estimates of the effects of 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, which is omitted from the regression and a determinant of both healthcare choice and health outcomes. While the model does not explicitly allow for heterogeneity in preferences for health, it is possible that these unobserved preferences also influence both choice of care and outcomes.

Let O_{ij} denote a health outcome for (sick) child i in cluster j; let h_{ij} denote the healthcare choice made for the child; and let $\mathbf{X}_{ij}$ denote a vector of child-, mother- and household-level characteristics. The following is the OLS specification, which as we discussed above, likely results in a biased estimate of the effect of h on O, due to unobserved determinants of outcomes in the error term ϵ that are correlated with healthcare choice:

$$O_{ij} = \beta h_{ij} + \mathbf{X}'_{ij}\gamma + \epsilon_{ij}. \tag{9}$$

4.1 Endogeneity of distance and rainfall

We propose instrumenting for healthcare choice using exogenous variation in the relative price of formal-sector care. The most salient costs to healthcare 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, individuals' largest costs are the opportunity costs and general disutility associated with traveling to the source of care. Gertler et al. (1987) find that the most important determinant of these costs is the distance to health facilities. 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 in public transportation for many hours; individuals must incur the opportunity cost of this time and the disutility of strenuous travel if they choose to visit a formal-sector healthcare provider. On the other hand, informal healthcare 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.

While distance to the nearest health facility has clearly been established as a shifter of the demand for formal-sector care, it is likely that this variable is correlated with unobserved elements of the error term in equation 9, and is thus invalid as an instrument. Specifically, an important issue is one of endogenous placement: governments often choose to locate health facilities where they will have the greatest impact on the surrounding population. As a result, facilities may be more likely to be placed in areas with poor average health, where illnesses are more likely to manifest as severe, and preferences for health are likely to be different than in places with generally better health. Conversely, it may also be that health facilities, as they are able to provide more quality healthcare to the local population, are positively correlated with the average level of health stock in their surrounding areas. In this case, distance to the nearest facility would be negatively correlated with unobserved severity of illness.

While rainfall is clearly exogenous, and so does not suffer the same critique as was given above regarding distance to health facility, we argue that it is, nevertheless, not a valid candidate (as a main effect) for an instrument. Rainfall does indeed generate variation in healthcare choices (as is shown in the last column of Table A.2). However, it surely also affects, amongst other things, the general extent and severity of sickness in a given locality (e.g. Frank Tanser et al. 2003); agricultural incomes for farm households (e.g. Christina Paxson 1992); and local prices (e.g. Esther Duflo and Christopher Udry 2004). All of these variables are 1) unavailable in our data or commonly unobservable, and 2) likely determinants of healthcare choice and health outcomes, and thus will be in the error term in the second stage. Therefore, rainfall as a main effect is not valid as an instrument.

4.2 An interaction instrument using distance

We introduce a new instrument for healthcare choice which does not suffer from the endogeneity issues discussed above, to which the distance to health facility variable falls prey. We *interact* distance to nearest health facility with the amount of rainfall at the time of sickness, and exclude only this interaction from the second stage, while controlling for the main effects of distance and rainfall in the first and second stages of a two-stage instrumental variables estimator. The two stages of analysis are specified as follows. Define d_{ij} to be the distance to the closest health facility, and R_{ij} to be the quantity of rainfall in cluster j at the time of the individual's sickness. The specification is:

$$\text{1st stage:} \quad h_{ij} = \alpha_1 \left(d_{ij} \text{ x } R_{ij}\right) + \alpha_2 d_{ij} + \alpha_3 R_{ij} + \mathbf{X}'_{ij}\alpha_4 + \zeta_{ij} \tag{10}$$

$$\text{2nd stage:} \quad O_{ij} = \beta_1 h_{ij} + \beta_2 d_{ij} + \beta_3 R_{ij} + \mathbf{X}'_{ij}\beta_4 + \epsilon_{ij} \tag{11}$$

The intuition behind the instrument is simple. The main effects of distance and rain should both be negative; that is, traveling a greater distance and being exposed to heavier rains should both discourage formal-sector health facility usage for individuals seeking care. Moreover, heavier rains should discourage individuals who live farther away *more* than individuals living closer to the nearest facility. Imagine, in the extreme case, one household is located 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 rains, the rain should incrementally deter the farther household *more* than the one just next door.

4.3 First Stage Results

The first two columns of Table II present 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. 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.

Insofar as distance to the nearest health facility is correlated with distance (or general access) to other resources, the first stage estimates may reflect the effect of differential access to non-healthcare resources on health outcomes. In order to avoid (at least in part) such bias, we control for the direct effects of rainfall and distance to the nearest health facility along with distance to the nearest market and the interaction of this distance with rainfall. 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.

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 A.2 in the appendix presents results from the first stage specifications with various sets of controls omitted.

It is important to note that the model presented above implicitly assumes that sickness is random; that is, whether or not an individual falls ill is not in any way determined by the distribution of health outcomes, nor therefore by the ex-post healthcare choice. In practice, all we need for the empirical application is that sickness, or more specifically the reporting of sickness, is not in some way driven by the instrument. To explicitly address any concerns that reduced access to healthcare (increased cost of travel to a health facility as proxied by the interaction of rainfall and distance) may lead to some systematic selection into or out of the sample of children with reported acute illnesses in the two weeks prior to survey, we present results in the third column of Table II from the regression of a binary for fever and/or cough in the last two weeks on the instrument and full set of controls. We find no evidence of selection into sickness on the basis of the instrument.

4.4 Instrument Validity: Satisfying the exclusion restriction

Given the importance of the proposed instrument to the empirical strategy, we undertake a careful discussion of its validity and conduct several checks before proceeding. Specifically, in addition to being strongly predictive of healthcare choice as established in the first stage results discussed in the previous section, the instrument should be uncorrelated with unobserved determinants of health outcomes, such as preferences for health and the severity of illness, or more completely, should be excludable from the second stage regressions.

In terms of the model parameters, the proposed second stage regressions of health outcomes on healthcare choice correspond to a regression of $p(s, h)$ on $h(s, \Delta P)$ to recover $\partial p/\partial h$. The major obstacle we attempt to overcome with the proposed empirical strategy is the presence of unobserved severity on both the left and right hand sides of the regression. We can imagine several ways in which severity, and even other parameters of the model, depend on rainfall and distance to nearest health facility. Given that we control for the direct effects of rainfall and distance in all of the linear regression specifications, all that is required for the exclusion restriction of the interaction instrument to hold is that severity and any other parameters which might enter both p and h are additively separable functions of rainfall and distance, and ultimately that p is itself additively separable in rainfall and distance; whereas ΔP, and hence h, are nonseparable in rainfall and distance. Below we present several checks as evidence in support of this assumption.

4.4.1 Instrument does not predict illness amongst non-acutely ill

We begin by applying the same strategy used earlier with regards to showing the endogeneity of distance to demonstrate 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 sick sample is just the reduced form regression reported in Column 2 of Table A.1.[13] 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 changes in outcomes. The coefficient estimate of this regression, reported in column 2 of Table A.1, 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 Column 1 of Table III. 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.

[13]Of 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.

4.4.2 Non-effect for faraway facilities

The instrument is posited to generate variation in healthcare choices by exogenously shifting the cost of formal-sector care, but only for individuals on the margin between choosing formal versus informal care. Thus, we should not expect to see an effect for individuals away from this margin, whose costs of travel are either too low or too high for a fluctuation to matter. We test this implication by running the baseline first-stage specification restricting the sample to individuals living more than 10km from the nearest health facility. The results are reported in column 2 of Table III. We find, consistent with the above hypothesis, that there is no predictive power of the instrument in the first stage for the restricted sample.

4.4.3 Using transitory rainfall variation

Our next concern 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 5 of Table III report results from the estimation of specifications identical to the first stage regressions reported in Table II, 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.

Columns 4 and 6 report first stage results controlling for past and future rainfall and their interactions with distance to nearest health facility and distance to nearest market. The results show that the predictive power of the instrument is very robust to the inclusion of these additional controls.

4.4.4 Controlling for general remoteness

Next, it is likely that the distance to the nearest health facility is correlated with the distance to other care options or other amenities that influence health outcomes. Fluctuations in rainfall might affect individuals who live farther away from these amenities differently than individuals who live closer, and this disparity may drive resulting changes in health outcomes.

If this were the case, our interaction instrument would not be validly excludable from the second stage. We address this issue by controlling for the distance to the nearest market, which is a proxy for the distance to other amenities, and its interaction with rainfall. As shown in Table A.2, even after including the distance-to-market variable and its interaction with rainfall along with numerous other household-level controls, the predictive power of our interaction instrument remains strong, and in fact increases.

As further evidence, we report in the first column of Table A.6 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 healthcare choice, we should expect that controlling for the interaction of rainfall with demographic covariates such as wealth, education, household size, etc. will greatly 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. Note that the point estimate of the coefficient on the instrument in column 1 of Table A.6 is nearly identical to that in the main specification reported in column 1 of Table II.

4.4.5 Nonlinear effects of endogenous distance

Lastly, 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 A.3 of the appendix. 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 A.3 that the first stage and main second stage results are robust to the inclusion of nonlinear terms in distance.

5 Results

5.1 Health Outcomes

Table IV presents the main results from regressions of healthcare choice on the incidence of fever and malaria amongst the sample of children who were reportedly ill with fever and/or cough in the two weeks prior to survey. The first two columns of Table IV report the OLS endogenous regression estimates. We find that formal healthcare has small, negative and precisely estimated effects on the incidences of fever and malaria. The point estimates suggest that a child is roughly 9 percentage points less likely to still have fever or malaria at the time of survey if they visit a health facility. Both estimates are significant at the 1 percent level.

The third and fourth columns in Table IV report second stage instrumental variables estimates from the estimations of equations (10) and (11) with the incidence of fever and malaria as outcomes, respectively. We find that the estimates of the effects on fever and malaria are still negative and significant, but the magnitudes are roughly 7 and 4 times as large as their OLS counterparts with point estimates of approximately 62 percentage points and 40 percentage points, respectively. Therefore, we find strong evidence of beneficial effects of formal sector healthcare on health outcomes following acute health shocks.

The underestimation of OLS is consistent with bias from the omission of severity from the specification. Given that the assumptions hold, the model predicts that unobserved severity will reduce the estimated efficacy of formal sector healthcare in the OLS regressions. That is, provided that more severe illnesses require longer recovery times, or more generally correspond to a lower probability of the child no longer being sick at the time of survey, and provided that the benefit of formal sector care is increasing in severity, OLS will find that the children who are selected to visit the formal sector health facility are more severely ill and therefore are less likely to be well by the time of survey. This will lead to systematic underestimation of the beneficial effects of formal sector care on recovery from acute illness.

The extent to which the second stage IV and OLS estimates on incidence of fever and malaria differ is consistent with that found in similar studies. Table A.7 reports a comparison of the estimates of severity bias found in this paper with those found in other work.

5.2 Receipt of Medication

Given the large and significant effects on health outcomes, we now explore whether access to medication is a potential mechanism by which formal-sector care improves health outcomes following acute illness. The medicine specifications reported in Table V are exactly the same as those reported in Table IV, excepting the dependent variables. This table only reports second-stage instrumental variables coefficient estimates.

Columns 1 and 2 of Table V show results for the number of medications received and number of medications conditional on receipt of at least one, respectively. We find that neither the number of medications received, nor the number conditional on receipt of any medication, increase significantly when an individual is driven exogenously to formal-sector healthcare.

The argument is often made that one important benefit of expanding the formal healthcare sector is enhanced access to appropriate medicines. While we find significant health improvements as a result of formal-sector care, we find no evidence that the mechanism by which these health improvements are achieved involves improved access to medications. This conclusion is distinctly at odds with the common assumption, made by academics and policymakers alike, that bad health outcomes following acute sickness in underserved populations are due to a lack of access to appropriate medicines (Hogerzeil 2004; WHO 2004; Zhu et al. 2008).

While these results are quite interesting, we must be careful here to make the distinction between our context–in which many drugs are available in the informal sector as well as at formal health facilities–and other contexts in which the informal drug sector is less active. In Tanzania, a fairly well-developed informal health infrastructure provides access to many medications outside of formal health facilities; this is not necessarily true in other developing settings, including other areas of sub-Saharan Africa (Catherine Goodman et al. 2004; S. Patrick Kachur et al. 2006; Goodman et al. 2007).

We must also contend with the possibility that formal-sector care, though it may not be associated with increased receipt of total medications, may nevertheless improve health outcomes through improved access to better quality or more appropriate medications. We test for this possibility by noting the fact that in much of Tanzania, and in general in malaria-endemic areas, international health policy standards recommend treating fever in young children presumptively as malaria (so as to avoid the potential onset of more severe forms of malaria) (Hugh Reyburn et al. 2004). Thus in our context, antimalarials specifically represent appropriate medication.

In columns 3 and 4 of Table V, we report results for the number of antimalarial and non-antimalarial medications. We find similar conclusions to the ones made from columns 1 and 2: for neither dependent variable does health facility usage generate increases in the number of medicines received. We therefore do not find evidence that access to more appropriate medicines is a primary mechanism through which formal-sector care improves health outcomes. Accordingly, we next explore the effects of formal-sector care on efficacy-enhancing behaviors related to medications as a possible mechanism.

5.3 Medication-Related Behaviors

Table VI reports second-stage instrumental variables results from regressions of adherence to medicine regimens and the number of days delayed before receiving medication on formal-sector care. Columns 1-3 of Table VI show respectively estimates of effects on a binary for whether the individual adhered to at least one malaria medication, the number of malaria medications to which the individual adhered, and a binary for adherence to all malaria medications received. The regressions reported in columns 1-3 are conducted on unrestricted samples; that is, sick individuals who did not receive any malaria medications are still included and treated as zeros in all three dependent variables.

Across all regressions on the unconditional sample, we find evidence that formal-sector care improves adherence to medication regimens. Acutely ill individuals that were exogenously driven to formal-sector care are 56 percentage points more likely to adhere to at least one malaria medication; adhere to .5 medications more than those who chose informal care or no care; and are 58 percentage points more likely to adhere to all malaria medications they received.

Columns 4-6 of Table VI report the same adherence regressions as in columns 1-3, but conducted on a conditional sample of children who received at least one malaria medication. Though the estimates are no longer significant at conventional levels, the signs and magnitudes of the point estimates are similar to those from the unconditional sample. Column 7 reports a regression conducted on the conditioned sample of days delayed before receipt of medication on formal-sector care. We find a large and significant reduction in days delayed induced by a visit to a formal health facility. This result suggests that amongst children who received at least one malaria medication, those who visited a health facility delayed in receiving malaria medications nearly a day and a half *less* than those who did not.

The fact that delay greatly reduces the efficacy of malaria treatment is well documented in the health literature (see, for example, Susan McCombie 1996). Similarly, full adherence to medicine regimens is crucial to their effectiveness. The large and significant reduction in the days delayed to treatment and increases in the probability of adhering to malaria medications and the number of malaria medications to which the child adhered are likely channels through which use of formal healthcare affects the incidence of malaria and fever (a primary symptom of malaria).

It is unlikely that these results derive from more prompt and consistent access to medicines through formal-sector care because, in the Tanzanian context, malaria medicines are commonly available in the informal sector as well. If individuals wish to purchase malaria medications, they can in general do so at drug stores, convenience stores, kiosks, etc. (with the exception of artemisinin-based combination therapies, whose distribution is often regulated through formal sector health facilities (see Achyuta Adhvaryu (2010)). We therefore interpret the results on delay and adherence as working primarily through an informational channel: the prescription of medicines is usually accompanied at health facilities with instructions and guidance from health workers on their optimal usage, which includes an emphasis on prompt and complete treatment.

Following this suggestive evidence of the transmission of information about effective treatment for malaria as a result of formal-sector care use, we explore the degree to which visits to formal care facilities are associated with the transmission of information about other illnesses, their transmission, and treatments.

5.4 Effects on General Health Knowledge

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

Column 1 of table VII 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 the standard error is large, the point estimate is quite small at less than .5 in absolute value, as compared to a mean of more than 14. Columns 2-4 show results on sub-indices which measure information on the specific topics of disease existence, transmission, and treatment, respectively. We do not find significant effects on information by topic either, with similarly small point estimates as well.

These results indicate that, though the evidence suggests that treatment at formal-sector facilities is accompanied by information which improves medication-related behaviors and subsequently improves treatment efficacy, there does not seem to be any transmission of information about other illnesses.

5.5 Heterogeneous Returns Across Severity of Illness

Lastly, we explore heterogeneity across the severity of illness in the returns to healthcare. Table VIII shows results from the 2SLS regressions of the interaction of formal healthcare and anemia status of the child on incidence of fever and malaria. To the extent that anemia, at least in part, proxies for severity, the third assumption from the model requires that the effects of formal healthcare on health outcomes be larger for children who are anemic. We find, in fact, large and significant effects of the interaction of formal healthcare and anemia on the incidence of fever and malaria, verifying this assumption in the model. We unfortunately do not have information on other co-morbidities, but we believe anemia status is an appropriate measure in this context given that it is a slightly longer term stock variable measuring some degree of general health.

6 Conclusion

Estimating the effects of formal-sector care on health outcomes is an important undertaking in both developed and developing contexts, the results of which are particularly germane to policy-makers' decisions regarding resource allocation and infrastructural investment. Consistent estimation of these effects has proven difficult due to the endogeneity of healthcare choice.

This paper improves on previous work by developing a new identification strategy, particularly appropriate to the developing country context, which exploits exogenous variation in the cost of formal-sector care, and in doing so overcomes the well-known problem of self-selection into healthcare options. We find large and significant reductions in the incidence of fever and malaria among children with acute illness as a result of visiting formal-sector facilities.

The paper makes further contributions to the literature in its exploration of the mechanisms by which formal-sector care improves health outcomes. We find that enhanced access to medications is not the primary mechanism by which health facilities improve health outcomes. Rather we find that visiting health facilities improves medication-related behaviors. Specifically, visits to formal-sector care reduce the delay in receiving medication and increase adherence to medication regimens, both of which are known to greatly improve the efficacy of treatment.

In contrast, we find no evidence that general health information is improved by formal health facility use. We find that mothers who are exogenously driven to take their acutely ill children to formal health facilities are no better informed about disease existence, transmission, and transmission than those who are not. Altogether, the results suggest that improvements in health outcomes following formal-sector care are driven by transfers of immediately relevant, treatment-specific information, rather than access to medicines or improvements in general health-related knowledge.

Consistent estimation of returns to quality healthcare in a developing country context and exploration of corresponding mechanisms by which these returns are achieved are, to the best of our knowledge, novel contributions to the literature. More generally, the identification strategy used in this paper contributes to the vast literature regarding self-selection issues in applied economics studies. The concept of self-selection has been addressed in the literature on treatment effects (Anders Bjorklund and Robert Moffitt 1987), technology adoption (Tavneet Suri 2010), and returns to schooling (David Card 2001). This paper proposes a robust identification strategy for dealing with similar self-selection issues in healthcare choice, as relates particularly to the developing country context.

References

[1] Adhvaryu, Achyuta. 2010. "Learning, Misdiagnosis and the Adoption of New Malaria Therapy." *working paper*.

[2] Adhvaryu, Achyuta and Anant Nyshadham. 2010. "Labor Supply, Schooling and the Returns to Healthcare in Tanzania." *working paper*.

[3] Almond, Douglas and Joseph Doyle. 2009. "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays." *working paper*.

[4] Almond, Douglas, Joseph Doyle, Amanda Kowalski, and Heidi Williams. 2010. "Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns." *Quarterly Journal of Economics,* forthcoming.

[5] Bjorklund, Anders and Robert Moffitt. 1987. "The estimation of wage gains and welfare gains in self-selection models." *The Review of Economics and Statistics* Vol. 69 no. 1: pp. 42-49.

[6] Bobonis, Gustavo, Edward Miguel and Charu Puri Sharma. 2006. "Iron Deficiency Anemia and School Participation." *Journal of Human Resources,* vol. 41, no. 4: pp. 692-721.

[7] Buchmueller, Thomas C., Mireille Jacobson, and Cheryl Wold. 2006. "How far to the hospital?: The effect of hospital closures on access to care," *Journal of Health Economics*. Vol. 25 no. 4: pp. 740-761.

[8] Card, David. 2001. "Estimating the return to schooling: Progress on some persistent econometric problems." *Econometrica* Vol. 69 no. 5: pp. 1127-1160.

[9] Chang, Fwu-Ranq. 1996. "Uncertainty and investment in health." *Journal of Health Economics,* Vol. 15, no. 3: pp. 369-376.

[10] Chern, Wen S., Edna T. Loehman, and Steven T. Yen. 1995. "Information, Health Risk Beliefs, and the Demand for Fats and Oils." *The Review of Economics and Statistics*, vol. 77, no. 3: pp. 555-564.

[11] Cohen, Jessica and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *Quarterly Journal of Economics*, vol. 125, no. 1: pp. 1-45.

[12] Cropper, Maureen. 1977. "Health, investment in health, and occupational choice." *Journal of Political Economy*, Vol. 85: pp. 1273-1294.

[13] Cutler, David, Robert Huckman, and Mary Beth Landrum. 2004. "The role of information in medical markets: an analysis of publicly reported outcomes in cardiac surgery." *American Economic Review, Papers and Proceedings*, Vol. 94, no. 2: pp. 342-346.

[14] Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. "The Quality of Medical Advice in Low-Income Countries," *Journal of Economic Perspectives*, Vol. 22, no. 2: pp. 93-114.

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

[16] Downs, Julie S., George Loewenstein, and Jessica Wisdom. 2009. "Strategies for Promoting Healthier Food Choices." *American Economic Review, Papers and Proceedings*, Vol. 99, no. 2: pp. 159-164.

[17] Doyle, Joseph. 2005. "Health Insurance, Treatment and Outcomes: Using Automobile Accidents as Health Shocks." *Review of Economics and Statistics*, Vol. 87, no 2: pp. 256-270.

[18] Doyle, Joseph. 2010. "Returns to Local-Area Emergency Health Care Spending: Using Health Shocks to Patients Far From Home." *working paper.*

[19] Doyle, Joseph, Steven Ewer, and Todd Wagner. 2010. "Returns to Physician Human Capital: Analyzing Patients Randomized to Physician Teams." *working paper.*

[20] Duflo, Esther and Christopher Udry. 2004. "Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices." *working paper.*

[21] Dupas, Pascaline. 2010. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics,* forthcoming.

[22] Gertler, Paul. 2004. "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *The American Economic Review,* Vol. 94, No. 2: pp. 336-341.

[23] Gertler, Paul, Luis Locay, and Warren Sanderson. 1987. "Are user fees regressive? The welfare implications of health care financing proposals in Peru." *Journal of Econometrics,* Vol. 36: pp. 6788.

[24] Geweke, John, Gautam Gowrisankaran and Robert Town. 2003. "Bayesian Inference for Hospital Quality in a Selection Model," *Econometrica,* Vol. 71, no. 4: pp. 1215-1238.

[25] Catherine Goodman, S. Patrick Kachur, Salim Abdulla, Eleuther Mwageni, Joyce Nyoni, Joanna A. Schellenberg, Anne Mills, and Peter Bloland. 2004. "Retail supply of malaria-related drugs in rural Tanzania: risks and opportunities." *Tropical Medicine and International Health,* vol. 9, no. 6: pp. 655663.

[26] Catherine Goodman, S. Patrick Kachur, Salim Abdulla, Peter Bloland, and Anne Mills. 2007. "Drug shop regulation and malaria treatment in Tanzania–why do shops break the rules, and does it matter?" *Health Policy and Planning,* vol. 22, no. 6: pp. 393-403.

[27] Gowrisankaran, Gautam, and Robert Town. 1999. "Estimating the Quality of Care in Hospitals Using Instrumental Variables." *Journal of Health Economics,* Vol. 18: pp. 747-767.

[28] Grossman, Michael. 2000. "The human capital model." In: A. J. Culyer and J. P. Newhouse (ed.), *Handbook of Health Economics,* edition 1, volume 1, chapter 7, pp. 347-408. Elsevier.

[29] Hogerzeil, Hans. 2004. "The concept of essential medicines: lessons for rich countries." *British Medical Journal*, vol. 329: pp. 1169-1172.

[30] S. Patrick Kachur, Jeffrey Schulden, Catherine A. Goodman, Herry Kassala, Berty Farida Elling, Rashid A. Khatib, Louise M. Causer, Saidi Mkikima, Salim Abdulla, and Peter B. Bloland. 2006. "Prevalence of malaria parasitemia among clients seeking treatment for fever or malaria at drug stores in rural Tanzania 2004." *Tropical Medicine and International Health,* vol. 11, no. 4: pp. 441451.

[31] Klemick, Heather, Kenneth Leonard, and Melkiory Masatu. 2006. "Understaffed or under-qualified? Health care staffing and patient access to quality care in rural Tanzania." *working paper.*

[32] Kremer, Michael and Alaka Holla. 2008. "Pricing and Access: Lessons from Randomized Evaluations in Education and Health." *working paper.*

[33] Kremer, Michael and Edward Miguel. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics,* Vol. 112, no. 3: pp. 1007-1065.

[34] Lavy, Victor, John Strauss, Duncan Thomas and Philippe de Vreyer. 1996. "Quality of health care, survival and health outcomes in Ghana." *Journal of Health Economics,* vol. 15, no. 3: pp. 333-357.

[35] Leonard, Kenneth. 2007. "Improving Health Outcomes By Choosing Better Doctors: Evidence of Social Learning about Doctor Quality from Rural Tanzania." *working paper.*

[36] Madajewicz, Malgosia, Alexander Pfaff, Alexander van Geena, Joseph Grazianoa, Iftikhar Husseinb, Hasina Momotajb, Roksana Sylvi and Habibul Ahsan. 2007. *Journal of Development Economics,* vol. 84, no. 2: pp. 731-754.

[37] Mboera, L.E.G., C.I. Fanello, R.C. Malima, A. Talbert, P. Fogliati, F. Bobbio, and F. Molteni. 2006. "Comparison of the Paracheck-Pf test with microscopy, for the confirmation of Plasmodium falciparum malaria in Tanzania." *Annals of Tropical Medicine and Parasitology*, Volume 100, Number 2: pp. 115-122.

[38] McCombie, S. C. 1996. "Treatment seeking for malaria: a review of recent research." *Social Science and Medicine* Vol. 43 no. 6: pp. 933-945.

[39] Miguel, Edward and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica,* vol. 72, no. 1: pp. 159-217.

[40] Mwabu, Germano, James Mwanzia and Wilson Liambila. 1995. "User charges in government health facilities in Kenya: effect on attendance and revenue." *Health Policy and Planning,* Vol. 10, no. 2: pp. 164-170.

[41] Mwabu, Germano, Martha Ainsworth and Andrew Nyamete. 1993. "Quality of Medical Care and Choice of Medical Treatment in Kenya: An Empirical Analysis." *Journal of Human Resources,* Vol. 28, no. 3: pp. 838-862.

[42] Paxson, Christina. "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." *American Economic Review,* Vol. 82, No. 1: pp. 15-33.

[43] Reyburn, Hugh, Redepmta Mbatia, Chris Drakeley, Ilona Carneiro, Emmanuel Mwakasungula, Ombeni Mwerinde, Kapalala Saganda, John Shao, Andrew Kitua, Raimos Olomi, Brian M Greenwood, and Christopher J M Whitty. 2004. "Overdiagnosis of malaria in patients with severe febrile illness in Tanzania: a prospective study." *British Medical Journal,* vol. 329, no. 7476.

[44] Selden, Thomas. 1993. "Uncertainty and health care spending by the poor: the human capital model revisited." *Journal of Health Economics,* Vol. 12: pp. 109-115.

[45] Strauss, John and Duncan Thomas. 1998. "Health, Nutrition, and Economic Development." *Journal of Economic Literature,* vol. XXXVI: pp. 766-817.

[46] Suri, Tavneet. 2010. "Selection and comparative advantage in technology adoption." *Econometrica,* forthcoming.

[47] Tanser, Frank, Brian Sharp and David le Sueur. 2003. "Potential effect of climate change on malaria transmission in Africa." *The Lancet*, vol. 362, no. 9398: pp. 1792-1798.

[48] Thirumurthy, Harsha, Joshua Graff Zivin and Markus Goldstein. 2009. "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya." *Journal of Human Resources*, vol. XLIII, no. 3: pp. 511-552.

[49] Thomas, Duncan, Elizabeth Frankenburg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Jaswadi, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini and Siswanto Wilopo. 2006. "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence." *California Center for Population Research working paper,* no. CCPR-070-06.

[50] Thornton, Rebecca. 2010. "Circumcision, Information, and HIV Prevention." *working paper.*

[51] World Health Organization. 2004. "Equitable access to essential medicines: a framework for collective action." *Essential Drug Monitor,* vol. 4: pp. 813.

[52] Zhu, Minmin, Hassan Dib, Xinping Zhang, Shengchun Tang, and Liqun Liu. 2008. "The influence of health insurance towards accessing essential medicines: The experience from Shenzhen labor health insurance." *Health Policy,* vol. 88, no. 2-3: pp. 371-380.

A Additional Results

A.1 Reduced Form Effects

Table A.1 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.2 First Stage Robustness

The first 3 columns of Table A.2 present results from the estimation of alternate specifications of the first stage equation. The specification in column 1 includes no additional controls beyond those listed; the specification in column 2 includes only age, region, wealth, education, and month of survey group effects; and the specification in column 3 is a reproduction of the main specification in column 1 of Table II 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.

Column 4 of Table A.2 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 4 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.

A.3 Robustness of Results to Nonlinear Distance Terms

Table A.3 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 A.3 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 A.3 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.4 Robustness of Results to Future and Past Rain Interaction Controls

Table A.4 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 Table A.3, the specifications reported in Table A.4 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 of Table A.3 we see that the main results are, in general, 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 and are of diminished magnitude, but are of the same sign.

A.5 Reduced Form Checks on Future/Past Rain

In Table A.5 we check reduced form effects of future and past rain interactions with distance on incidence of fever and malaria and on selection into sickness. We omit rainfall in the month of reported illness and its interaction with distance and check the predictive power of the non-contemporaneous rainfall interactions. The future rain interaction is not predictive and the past rain interaction coefficient is only weakly significant at the 10 percent level. In column 3 of Table A.5, we find no evidence that non-contemporaneous rainfall nor its interaction with distance are predictive of selection into sickness. This result provides further evidence that the interaction of rainfall and distance to health facility only effects health temporarily through its effect on healthcare choice in times of acute illness and not otherwise through additional channels.

A.6 Reduced to Additional Controls of Interactions of Rainfall with Other Covariates

Table A.6 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 A.4 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 of Table A.3 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.7 Comparison of Self-Selection Bias in Estimates

Table A.7 reproduces the main health outcomes results from this study in columns 1 and 2. The remaining columns of Table A.7 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. Moreover, the degree of severity 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

- $shocknorm = (rain - histmean)/(histsd)$

- $infodisease = TB + STD + HIV/AIDS + otherSTD$, 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

- $infotransmission = TBair + TButensils + TBtouch + TBfood + TBsex + TBunknown + AIDSabst + AIDScondom + AIDS1prtnr + healthyAIDS + AIDSmsqto + AIDSfood + AIDSwitch$, 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

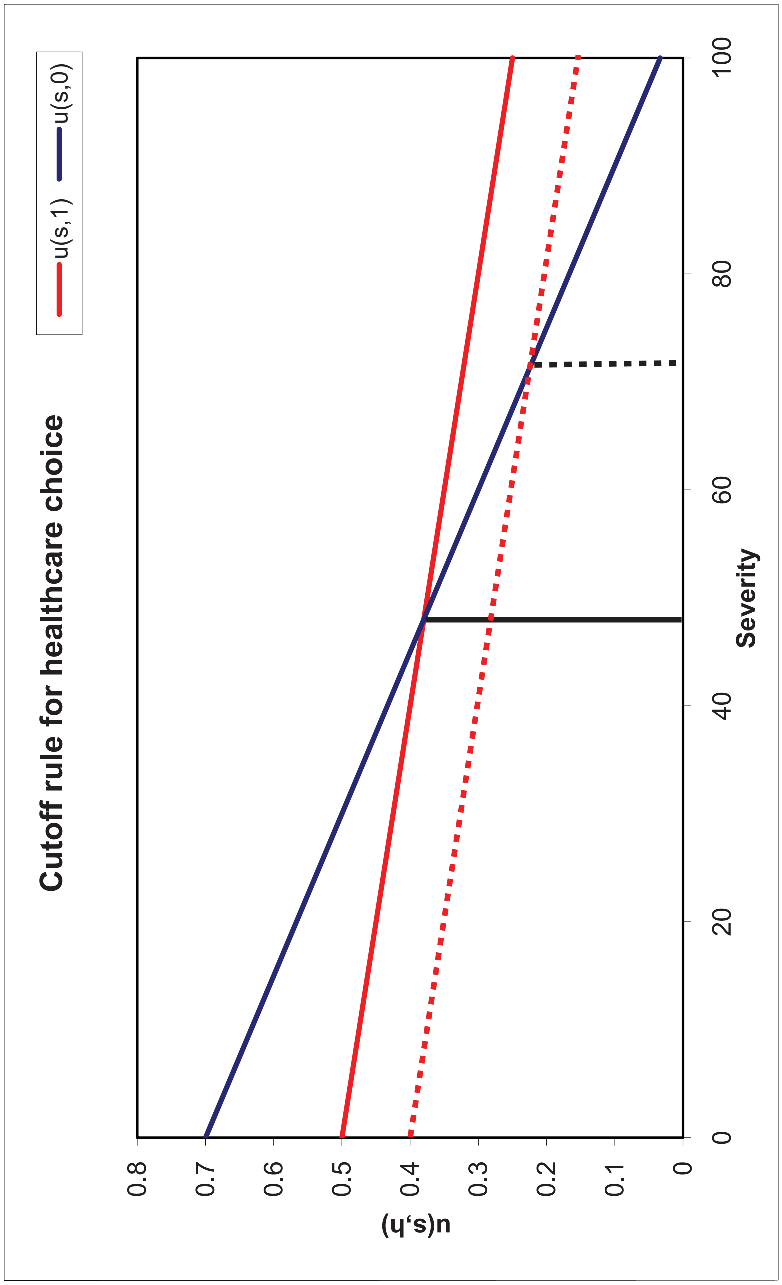

FIGURE 2: PRICE SHIFT, LATE AND ATE

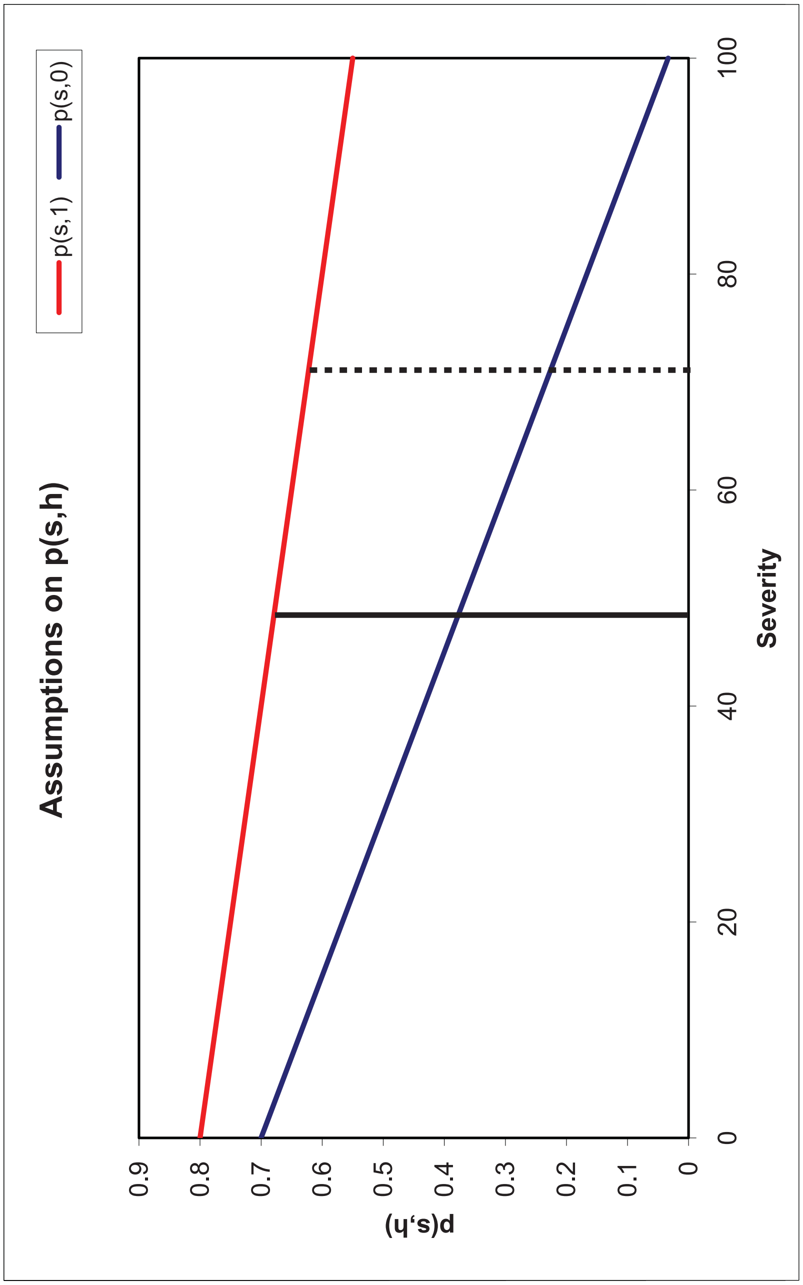

Table I: Summary Statistics

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	
	Mean	SD	Mean	SD	Mean	SD
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
Distance to nearest formal-sector health facility (km)	4.418	6.027	3.352	4.789	5.853	7.131
Distance to nearest market (km)	26.293	29.585	22.127	27.652	31.927	31.168
Health status:						
Tested positive for mild to severe anemia	0.770	0.421	0.777	0.417	0.761	0.427
Tested positive for malaria (at time of survey)	0.232	0.422	0.165	0.372	0.319	0.467
Fever now (self-reported, at time of survey)	0.272	0.445	0.213	0.409	0.351	0.478
Demographic characteristics:						
Age	1.603	1.301	1.489	1.273	1.754	1.324
Female	0.496	0.500	0.483	0.500	0.513	0.500
Rural	0.819	0.385	0.783	0.412	0.867	0.339
Household size	7.254	3.856	6.815	3.427	7.842	4.298
Medications:						
No. of medicines received	1.491	0.863	1.761	0.756	1.122	0.863
No. of medicines conditional on # > 0	1.694	0.708	1.795	0.722	1.511	0.644
No. of antimalarials	0.551	0.521	0.703	0.492	0.343	0.488
No. of non-antimalarials	0.940	0.762	1.058	0.808	0.779	0.661
Medication-related behaviors:						
Adhered to at least 1 medication	0.629	0.483	0.639	0.481	0.602	0.491
No. of medications adhered to	0.634	0.492	0.643	0.489	0.608	0.502
Adhered to all medications	0.619	0.486	0.628	0.484	0.596	0.492
No. of days delayed before receiving medication	1.180	1.079	1.150	1.071	1.264	1.099
Health information indices:						
Composite	14.173	3.559	14.307	3.496	13.971	3.650
General disease-related	3.481	1.093	3.463	1.105	3.507	1.078
Transmission-related	8.507	2.749	8.611	2.701	8.357	2.815
Treatment-related	1.909	1.298	1.996	1.275	1.778	1.324

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

Table II: First Stages and Selection Check

Interaction of Rain and Distance to Health Facility on Healthcare Choice			Selection on Sick
	Child Level	Mother Level	Fever Last 2 Weeks
Rain x Distance	**-0.0204*** **	**-0.0192*** **	**0.000988**
	(0.00516)	**(0.00534)**	**(0.00185)**
Distance	0.00569	0.00348	-0.000503
	(0.00489)	(0.00506)	(0.00173)
Market	-0.00167	-0.00177	-1.72e-05
	(0.00130)	(0.00131)	(0.000493)
Rain	-0.139*	-0.175**	-0.0162
	(0.0706)	(0.0724)	(0.0271)
Rain x Market	0.000776	0.000912	0.000299
	(0.00118)	(0.00122)	(0.000409)
F-test: Rain x Distance=0	**15.60**	**12.98**	
Prob>F	**9.37e-05**	**0.000358**	
Observations	1081	963	6267
R-squared	0.178	0.177	0.048

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 inlcude age, region, wealth, eduction, 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.

Table III: Instrument Checks

	Evidence of Exogeneity	Falsification Using Faraway Facilities	Falsification with Future Rain	Robust to Future Rain Inclusion	Falsification with Past Rain	Robust to Past Rain Inclusion
	Malaria (if no fever in last 2 weeks)	Formal Healthcare	Formal Healthcare		Formal Healthcare	
Future or Past Rain x Dist			**-0.00563**	0.00241	**-0.000504**	0.00673
			(0.00506)	(0.00528)	**(0.00652)**	(0.00686)
Future or Past Rain			0.116**	0.0640	-0.0428	-0.108
			(0.0519)	(0.0642)	(0.0859)	(0.0855)
Future or Past Rain x Market			0.000694	0.000396	-5.87e-05	-8.74e-05
			(0.00102)	(0.00114)	(0.00146)	(0.00132)
Rain x Distance	0.00151	-0.0201		**-0.0203*** **		**-0.0214*** **
	(0.00206)	(0.0193)		**(0.00630)**		**(0.00611)**
Distance	0.000226	0.0228	-0.00382	0.00259	-0.00943	0.00145
	(0.00183)	(0.0139)	(0.00693)	(0.00644)	(0.00598)	(0.00557)
Market	-0.000149	-0.00545	-0.00169	-0.00202	-0.000770	-0.00116
	(0.000517)	(0.00370)	(0.00139)	(0.00167)	(0.00123)	(0.00144)
Rain	-0.0409	0.638		-0.0713		-0.132*
	(0.0307)	(0.534)		(0.0784)		(0.0732)
Rain x Market	0.000826*	0.00507		0.000806		0.000513
	(0.000482)	(0.00489)		(0.00115)		(0.00106)
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		**1.09**		**10.33**		**12.29**
Prob>F		**0.301**		**0.00142**		**0.000511**
Observations	4177	135	1130	1130	1130	1130
R-squared	0.157	0.429	0.167	0.176	0.162	0.177

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). "Faraway facilities" are those more than 10km from the individual's dwelling. See Table II for other notes.

Table IV: Health Outcomes

Effects of Healthcare Choice on Health Outcomes

	OLS		Second Stage IV	
	Fever	Malaria	Fever	Malaria
Formal Healthcare	**-0.0945*** **	**-0.0961*** **	**-0.619** **	**-0.398***
	(0.0315)	**(0.0270)**	**(0.266)**	**(0.236)**
Distance	0.00132	0.00228	-0.00442	-0.00102
	(0.00295)	(0.00236)	(0.00385)	(0.00286)
Market	-0.00123	0.000902	-0.00191	0.000706
	(0.00111)	(0.000940)	(0.00130)	(0.00105)
Rain	0.0640	0.0577	-0.0191	0.0152
	(0.0855)	(0.0599)	(0.106)	(0.0702)
Rain x Market	0.000494	-0.000885	0.000612	-0.000958
	(0.000943)	(0.000779)	(0.00108)	(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 inlcude age, region, wealth, eduction, 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.

Table V: Receipt of Medication

Second Stage IV: Effects of Healthcare Choice on Number of Medications Obtained

	Any Medication		Medications By Type	
	No. of Meds	No. of Meds (Conditional on One Med)	No. of Malaria Meds	No. of Non-Malaria Meds
Formal Healthcare	**0.572**	**0.218**	**0.283**	**0.290**
	(0.410)	**(0.345)**	**(0.279)**	**(0.356)**
Distance	-0.000531	-0.00519	0.000886	-0.00142
	(0.00619)	(0.00549)	(0.00437)	(0.00546)
Market	-0.00227	-0.00242	-0.00231	3.12e-05
	(0.00243)	(0.00210)	(0.00143)	(0.00187)
Rain	0.0546	0.122	-0.0473	0.102
	(0.133)	(0.119)	(0.0837)	(0.125)
Rain x Market	0.000743	0.000493	0.000697	4.60e-05
	(0.00185)	(0.00153)	(0.00121)	(0.00131)
Observations	1049	931	1049	1049
R-squared	0.205	0.117	0.290	0.146

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). See Table IV for other notes.

Table VI: Medication-Related Behaviors

Second Stage IV: Effects of Healthcare Choice on Adherence and Delay to Medication

	Uncondtional			Conditional on Reciept 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****	**0.506***	**0.579****	**0.427**	**0.343**	**0.440**	**-1.328****
	(0.277)	**(0.285)**	**(0.278)**	**(0.314)**	**(0.322)**	**(0.315)**	**(0.578)**
Distance	0.00287	0.00279	0.00321	0.00481	0.00420	0.00530	-0.00618
	(0.00477)	(0.00476)	(0.00471)	(0.00583)	(0.00588)	(0.00577)	(0.0119)
Market	0.000777	0.000771	0.000967	0.00446**	0.00453**	0.00471***	-0.00447
	(0.00139)	(0.00141)	(0.00136)	(0.00179)	(0.00179)	(0.00179)	(0.00417)
Rain	-0.0659	-0.0485	-0.0669	-0.0844	-0.0528	-0.0916	0.0551
	(0.0948)	(0.100)	(0.0950)	(0.103)	(0.113)	(0.104)	(0.238)
Rain x Market	-0.000731	-0.000761	-0.000877	-0.00232*	-0.00249*	-0.00255*	0.000279
	(0.00122)	(0.00128)	(0.00122)	(0.00140)	(0.00146)	(0.00140)	(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). See Table IV for other notes.

Table VII: Information Externalities

Effects of Healthcare Choice on Health-related Information

	Composite	Information by Topic		
		Information About Disease	Information About Transmission	Information About Treatment
Formal Healthcare	**-0.421**	**-0.0276**	**-0.351**	**0.358**
	(2.086)	**(0.575)**	**(1.473)**	**(0.754)**
Distance	-0.0510	-0.00666	-0.0546**	0.00664
	(0.0321)	(0.0111)	(0.0219)	(0.0134)
Market	-0.0188**	-0.000181	-0.0110*	-8.97e-05
	(0.00816)	(0.00310)	(0.00572)	(0.00332)
Rain	-0.457	0.0653	-0.262	-0.0253
	(0.693)	(0.207)	(0.451)	(0.262)
Rain x Market	0.0113*	-0.00178	0.00821**	2.53e-05
	(0.00582)	(0.00224)	(0.00401)	(0.00251)
Observations	845	963	943	854
R-squared	0.258	0.157	0.215	0.170

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). See Table IV for other notes.

Table VIII: Heterogeneous Effects

Effects of Healthcare Choice on Health Outcomes by Initial Health

	Fever	Malaria
Anemic x Formal Healthcare	**-1.497***	**-1.819****
	(0.854)	**(0.795)**
Formal Healthcare	0.568	1.036
	(0.754)	(0.679)
Anemic	0.916	1.351**
	(0.642)	(0.599)
Joint F-test for Anemic x Formal Healthcare	**5.17**	**5.6**
Prob>F	**0.006**	**0.004**
Joint F-test for Formal Healthcare	**6.11**	**6.36**
Prob>F	**0.003**	**0.002**
Observations	935	934
R-squared	-0.512	-0.481

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). See Table IV for other notes. Instruments are (Anemic x Rain x Distance) and (Rain x Distance).

Table A.1: Reduced Form Estimates

Effects of Healthcare Choice on Health Outcomes and Medication-Related Behaviors

	Health Outcomes		Medications		
	Fever	Malaria	No. of Meds	Adherence (Unconditional)	Days Delayed Before Medicating
Rain x Distance	**0.0127****	**0.00769**	**-0.0125**	**-0.0122***	**0.0334****
	(0.00559)	**(0.00496)**	**(0.00941)**	**(0.00633)**	**(0.0152)**
Distance	-0.00786	-0.00300	0.00323	0.00655	-0.0132
	(0.00500)	(0.00424)	(0.00871)	(0.00634)	(0.0156)
Market	-0.000845	0.00110	-0.00304	2.62e-05	-0.00316
	(0.00113)	(0.000948)	(0.00225)	(0.00129)	(0.00453)
Rain	0.0663	0.0612	-0.0250	-0.144*	0.159
	(0.0838)	(0.0585)	(0.119)	(0.0779)	(0.225)
Rain x Market	0.000119	-0.00105	0.00108	-0.000405	0.000797
	(0.000983)	(0.000767)	(0.00166)	(0.00101)	(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). See Table IV for other notes.

Table A.2: First Stage Robustness

	First Stage			Main Effects
	No Controls	Group Effects	Main Specification	No Rain Interactions
Rain x Distance	**-0.0184*****	**-0.0176*****	**-0.0204*****	
	(0.00543)	**(0.00513)**	**(0.00516)**	
Distance	0.00161	0.00390	0.00569	**-0.0108*****
	(0.00529)	(0.00502)	(0.00489)	**(0.00310)**
Market	-0.00253**	-0.00182	-0.00167	-0.000794
	(0.00123)	(0.00130)	(0.00130)	(0.000756)
Rain	-0.00756	-0.0322	-0.139*	**-0.148****
	(0.0409)	(0.0557)	(0.0706)	**(0.0662)**
Rain x Market	0.00110	0.000851	0.000776	
	(0.00106)	(0.00114)	(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	1081	1130
R-squared	0.062	0.150	0.178	0.166

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

Table A.3: Robustness to Nonlinear Distance Terms

Effects of Healthcare Choice on Health Outcomes and Medication-Related Behaviors

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

Table A.4: Robustness to Both Future and Past Rain Interaction Controls

Effects of Healthcare Choice on Health Outcomes and Medication-Related Behaviors

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

Table A.5: Reduced Form Checks on Future/Past Rain

Effects of Future/Past Rain and Interactions on Health Outcomes

	Health Outcomes		Selection on Sick
	Fever	Malaria	Fever Last 2 Weeks
Future Rain x Dist	**-0.00171**	**5.86e-05**	**-0.00113**
	(0.00506)	**(0.00159)**	**(0.00138)**
Future Rain	0.0569	0.0449*	0.00171
	(0.0525)	(0.0272)	(0.0217)
Future Rain x Market	-0.00179**	-0.00113***	0.000132
	(0.000908)	(0.000422)	(0.000413)
Past Rain x Dist	**0.0110***	**0.00554***	**0.00236**
	(0.00581)	**(0.00301)**	**(0.00206)**
Past Rain	-0.152**	-0.0665*	-0.00993
	(0.0641)	(0.0365)	(0.0277)
Past Rain x Market	0.00243**	0.000509	0.000379
	(0.00115)	(0.000591)	(0.000492)
Distance	-0.00493	-0.00172	0.000162
	(0.00665)	(0.00251)	(0.00228)
Market	-0.000340	0.000871	-0.000365
	(0.00158)	(0.000695)	(0.000600)
Rain	0.0832	-0.0227	-0.00279
	(0.0802)	(0.0323)	(0.0296)
Rain x Market	0.000187	0.000698	0.000148
	(0.000883)	(0.000632)	(0.000469)
Observations	1122	5312	6502
R-squared	0.115	0.183	0.046

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). See Table IV for other notes.

Table A.6: Robustness to Additional Controls of Interactions of Rainfall with Other Covariates

Effects of Healthcare Choice on Health Outcomes and Medication-Related Behaviors

	First Stage	Health Outcomes		Medications		
	Formal Healthcare	Fever	Malaria	No. of Meds	Adherence (Unconditional)	Days Delayed Before Medicating
Formal Healthcare		**-0.771****	**-0.399**	**0.568**	**0.522***	**-1.403****
		(0.309)	**(0.265)**	**(0.462)**	**(0.291)**	**(0.591)**
Rain x Distance	**-0.0190*****					
	(0.00557)					
Distance	0.00500	-0.00569	-8.56e-05	-0.000368	0.00210	-0.00648
	(0.00523)	(0.00417)	(0.00319)	(0.00645)	(0.00463)	(0.0107)
Market	-0.00177	-0.00259*	0.000700	-0.00273	0.00137	-0.00628
	(0.00143)	(0.00151)	(0.00116)	(0.00252)	(0.00136)	(0.00421)
Rain	0.294	-0.641	-0.179	-1.005	-1.435***	0.902
	(0.577)	(0.649)	(0.490)	(0.890)	(0.474)	(1.877)
Rain x Market	0.00109	0.00169	-0.000455	0.00116	-0.00136	0.00285
	(0.00132)	(0.00132)	(0.00110)	(0.00206)	(0.00111)	(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 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. See Table IV for other notes.

Table A.7: Comparison of Self-Selection Bias in Estimates

Comparison of Self-Selection Bias in Estimates with Previous Studies			Adhvaryu and Nyshadham (2010)	Gowrisankaran and Town (1999)	Doyle (2010)
Data/Location	DHS Tanzania (2007-08)		LSMS Kagera, Tanzania (1991-1994)	United States	United States
Dependent Variable	Fever	Malaria	Fever	Mortality	Mortality
Independent Variable of Interest	*Formal Healthcare*		*Formal Healthcare*	*Not-for-profit Hospitals (vs. For-profit)*	*Area-level Healthcare Spending*
Adjusted Estimate	IV **-0.619**** **(0.266)**	IV **-0.398*** **(0.236)**	IV **-0.595**** **(0.291)**	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 **0.0300** **(0.0236)**	OLS **-0.00016** **(0.00047)**	OLS (Florida Locals) **-0.003** **(0.0039)**
Observations	1123	973	5203	844	37185 (Visitors), 749762 (Locals)

Notes: Columns 1-2: See notes in table IV; Column 3: See Adhvaryu and Nyshadham (2010) Table IV; Column 4: See Gowrisankaran and Town (1999) Table IV; Column 5: See Doyle et al. (2010) Table II.